

V. Hutson · S. Martinez · K. Mischaikow · G.T. Vickers

The evolution of dispersal

Received: 17 August 2002 / Revised version: 8 January 2003 /
Published online: 15 May 2003 – © Springer-Verlag 2003

Abstract. A non-local model for dispersal with continuous time and space is carefully justified and discussed. The necessary mathematical background is developed and we point out some interesting and challenging problems. While the basic model is not new, a ‘spread’ parameter (effectively the width of the dispersal kernel) has been introduced along with a conventional rate parameter, and we compare their competitive advantages and disadvantages in a spatially heterogeneous environment. We show that, as in the case of reaction-diffusion models, for fixed spread slower rates of diffusion are always optimal. However, fixing the dispersal rate and varying the spread while assuming a constant cost of dispersal leads to more complicated results. For example, in a fairly general setting given two phenotypes with different, but small spread, the smaller spread is selected while in the case of large spread the larger spread is selected.

1. Introduction

The spatial dispersal of cells or organisms is clearly central to biology, and the evolution of dispersal itself is consequently of great importance. Although the latter issue has only fairly recently received much attention, it is now a major focus of theoretical interest. In outline, it is probably fair to say that the models for this have usually fallen into two categories. The first consists of models based on discrete time and space (patch models); these are notoriously intractable theoretically and are usually studied by numerical simulation. We do not discuss the validity of these models, suffice it to say that it is far from clear what their relationship is with models based upon continuous space (see [19] p. 503 for some remarks on this issue). Extensive references to the evolution of dispersal may be found in the reviews [20], [3] and [8]. The patch models are discussed in [5], [17], [24] and [37]. Here we shall consider models continuous in both time and space.

V. Hutson: Department of Applied Mathematics, The University of Sheffield, Sheffield, S3 7RH, U.K.

S. Martinez: Departamento de Ingeniería Matemática, Universidad de Chile, Casilla 170 Correo 3, Santiago, Chile

K. Mischaikow: Department of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, U.S.A.

G.T. Vickers: Department of Applied Mathematics, The University of Sheffield, Sheffield, S3 7RH, U.K. e-mail: g.vickers@shef.ac.uk

S. Martinez was partially supported by Fondecyt 1020126 and Fondecyt Líneas Complementarias 8000010. K. Mischaikow was supported in part by NSF Grant DMS 0107396.

Key words or phrases: Non-local dispersal – Integral kernel – Evolution of dispersal

The continuous model which has been most studied in the context of dispersal is based upon reaction-diffusion equations. In particular, for discussions on the evolution of diffusion in this setting see [4], [13], [27], [28] and [19]. For a spatially heterogeneous but temporally homogeneous (i.e. stationary) environment, there is a clear conclusion ([4] and [13]): *In a very strong sense, that is under very weak conditions, the phenotype with the lowest diffusion rate will be selected.* For an environment which is heterogeneous but time periodic, this conclusion no longer holds and the situation is much more complex, see [19]. This conclusion is nevertheless somewhat surprising from a biological point of view, as one might be inclined to think that, although there are a whole range of possible mechanisms, spatial heterogeneity would be very important in driving (at least under some circumstances) an increase in the rate of diffusion. For example, it is possible that this would reduce intra-specific competition.

The exact nature of dispersal and its theoretical treatment are coming more and more into question, see [2] and [31] for lengthy discussions and a huge variety of articles cited. It is, therefore, natural to ask whether the class of reaction-diffusion models, with its fundamental assumption that motion is governed by a random walk, is too restricted for the more subtle question of dispersal selection. It is possible to ask whether certain dispersal mechanisms may be preferred over others. This is a rather broad question and, although interesting, appears to be hard to formulate precisely.

Our approach is to ask a more precise but simpler question. In Section 2 we derive a class of models more general than diffusion, but still moderately tractable analytically, in order to examine whether, in a controlled context, one can gain information and insight in this difficult area. Variants of the particular model we shall adopt have been discussed in [12] and [32] and is there termed a ‘position-jump process’. We use the continuous-time version of this model. Similar discrete-time versions have been long current and we give only three references, one mathematical, [39], and the others biological, [21] [22].

Let x denote position in space and t time. The basis of our model are equations of the form

$$\frac{\partial u}{\partial t} = Du + uf(u, x), \quad (1.1)$$

where the dispersal operator D is based upon an integral kernel. To illustrate, consider the simplest example,

$$(Du)(x) = (D_{\rho, L}u)(x) = \rho \left\{ \frac{1}{L} \int_0^L k \left(\frac{|x-y|}{L} \right) u(y) dy - u(x) \right\} \quad (1.2)$$

where the spatial domain is $[0, l] \subset \mathbb{R}$ and k is a non-negative function satisfying the normalisation condition:

$$\int_{-\infty}^{\infty} k(y) dy = 1. \quad (1.3)$$

Notice that there are two parameters governing D : the *dispersal rate* ρ , which represents the total number of the dispersing organisms per unit time, and the *dispersal*

spread L , which characterises the dispersal distance. As will become clear through the analysis presented in the paper, these two parameters have quite different qualitative implications concerning the relative competitive advantage.

With regard to the reaction term, the following assumptions are made:

$$f(0, x) > 0 \text{ and non-constant, } \frac{\partial f}{\partial u} < 0. \quad (1.4)$$

The simplest example is provided by the classical logistic growth function

$$f(u, x) = a(x) - u. \quad (1.5)$$

Here $a(x)$ models the effect of a spatially variable, but temporally constant, environment.

The goal of our investigation is to understand how competition drives selection of ρ and L . Ideally we would do this by considering a competitive system of the form

$$\frac{\partial u}{\partial t} = D_{\rho_0, L_0} u + u f(u + v, x) \quad (1.6)$$

$$\frac{\partial v}{\partial t} = D_{\rho_1, L_1} v + v f(u + v, x) \quad (1.7)$$

Observe that in this system the reaction terms are identical, i.e. the interaction of the different phenotypes with the environment is identical. The only distinction is in the dispersal terms, thus understanding the global dynamics of this system allows one to decide to what extent either dispersal strategy is advantageous.

Before proceeding to describe our investigation, we wish to persuade the reader that the answers are unlikely to be obvious. To motivate this statement, we consider the simplest model of this nature (a reaction-diffusion model) and point out that, in our view, the obvious argument is false. Suppose for simplicity that the environment a , determining the per capita growth rate, is such that $a(x) \geq 0$ for all x . The ‘obvious’ argument referred to goes as follows. *It is best for individuals to live where the environment $a(x)$ is ‘best’, that is greatest, and it is an advantage to the species to have small diffusion since this will ensure that individuals stay in the best regions and do not wander into poor regions. Thus small diffusion leads to the maximum carrying capacity and is selected.* Unfortunately it is easy to prove that the carrying capacity is maximised by an *intermediate* diffusion rate. Of course the above argument is not correct for it is the question of invasion by small numbers of a new phenotype (with the environment represented by $a - \tilde{u}$, where \tilde{u} is the density of the resident species) which must be settled. It is well known that it is not the case that all successful mutations are beneficial in the sense that they increase the carrying capacity.

Easy intuitive arguments are not readily found and we shall therefore rely mainly on the mathematical results rather than on (possibly fallible) intuition. Unfortunately, results concerning the global dynamics of the above mentioned system appear to be quite difficult to obtain. Therefore, we adopt the following approach. Assume

that for the parameters (ρ_0, L_0) the phenotype u is at a positive equilibrium $\tilde{u}(x)$, i.e. $\tilde{u}(x)$ satisfies,

$$0 = D_{\rho_0, L_0} \tilde{u} + \tilde{u} f(\tilde{u}, x),$$

and suppose that a mutant with one parameter changed, e.g. (ρ_1, L_0) or (ρ_0, L_1) , is introduced. We then ask whether the mutant (which is always introduced with very small numbers) will or will not invade successfully. Mathematically, this corresponds to determining the stability of the semi-trivial equilibrium $(\tilde{u}, 0)$. This in turn requires an understanding of the principal eigenvalue associated with the linearization at $(\tilde{u}, 0)$; if it is positive the invasion attempt is successful; if it is negative the mutants die out.

The technical background for this approach is, in the case of reaction-diffusion equations, extremely well known. For dispersal of the form (1.2) this is not the case. The analysis of the dynamics of such equations seems largely to have taken place in the context of physically motivated problems (see [10], [11], [38] and [1]) and some results are available; see also [25] for a discussion and references from a different point of view. Although the fact that the dispersal is represented by a bounded operator may seem at first sight to simplify matters, as can be seen from [6], under some circumstances quite the opposite is the case. At any rate, it is convenient (and safer) to present briefly some basic results in the form needed here and this is done in Sections 3, 4 and 5. For example, using the results of Section 3, we prove the following.

Suppose (1.1) has a positive solution \tilde{u} . Then, \tilde{u} is the global attractor for all solutions with initial conditions consisting of non-trivial, non-negative functions.

As we have remarked above, the dispersal is controlled by two parameters. The rate ρ is analogous to the diffusion coefficient in the setting of reaction diffusion models. We conjecture that if the spread L is held fixed, i.e. $L_0 = L_1$ and $\rho_0 < \rho_1$, then even in the more general context of (1.6) and (1.7) the same result as for the reaction diffusion holds: slow dispersal is always selected. More precisely, the semi-trivial equilibrium $(\tilde{u}, 0)$ is the global attractor for the cone of initial conditions consisting of non-trivial, non-negative functions. In Section 5 we prove a partial result along these lines, see further remarks in Sections 5 and 8.1.

These results indicate that the selection appears to be a remarkably robust phenomenon: if the resource is spatially heterogeneous, but temporally constant, then the slowest disperser is *always* selected, this result being independent of the form of the kernel. However, the overall picture changes radically if the rate is fixed and invasion by phenotypes with different spreads are considered; this is the main focus of the discussion of the non-local dispersal in this paper.

For the remainder of the paper, we fix the rate and consider the invasion of phenotypes with different spread. The results in this setting are considerably more varied, and substantially more difficult to obtain. The first point that needs to be addressed is how to choose an interesting 1-parameter family of kernels with comparable spreads. This is done in Section 2.4 where we introduce the concept of a dispersal budget and use this to restrict our attention to kernels of equal cost.

Even with this restriction, comparing different spreads is difficult. Thus, as a first step we limit our analysis to the case where $|L_1 - L_0| \ll 1$. With spread it

appears that the behaviour near the boundaries of the domain may play an important role. We essentially consider two cases. In Section 6 we study the most obvious case which we term ‘hostile surroundings’; here individuals which disperse outside a finite region, $[0, l]$, die. In Section 7, we imagine an infinite but periodic environment. A variant of this is to say that the boundary reflects back all individuals which reach it. The advantage of these latter approaches is that the boundary does not play a crucial role.

An interesting phenomena is observed. For L_0, L_1 sufficiently small, the smaller spread is preferred. More precisely, the semi-trivial equilibrium $(\bar{u}, 0)$ corresponding to the phenotype with smaller spread is stable. In other words, the phenotype with larger spread cannot invade.

On the other hand, if the spread is sufficiently large then under certain circumstances the selection of larger spread is possible; these questions are treated in Sections 6 and 7. A significant conclusion is that no simple intuitive view of selection seems possible. We discuss these matters in Section 8, where we review the mathematical and biological results of this investigation.

2. Derivation of the model

2.1. Infinite environment

Consider a single species in an n -dimensional habitat where it is presumed that the population can be adequately modelled by a single function, $u(x, t)$, which is the density at position x at time t . We will derive a continuous model for the population dynamics for this species by considering in detail a situation discrete in both space and time, and then letting the size and time intervals become small. Although the approach has some similarities with the classic derivation of the Laplacian via a random walk, here it is not presumed that individuals move from a given patch via a binomial distribution. For a derivation of the diffusion model, we cite [30] and [29] among the huge number of references on this topic.

For clarity of exposition, we derive the continuous model for the case $n = 1$, and claim that it is straightforward to generalize to arbitrary dimension. Divide \mathbb{R} (the habitat) into contiguous sites, each of length Δx . Discretize time into steps of size Δt . Let $u(i, t)$ be the density of individuals in site i at time t . We wish to derive the change in the number in this site during the next time interval. The first assumption is that the rate at which individuals are leaving site i and going to site j is constant. Thus the total number should be proportional to: the population in the interval i , which is $u(i, t)\Delta x$; the size of the target site, which is Δx ; and the amount of time during which the transit is being measured, Δt . Let $\alpha(j, i)$ be the proportionality constant. Then, the number of individuals leaving site i during the interval $[t, t + \Delta t]$ is

$$\sum_{\substack{j=-\infty \\ j \neq i}}^{\infty} \alpha(j, i)u(i, t)(\Delta x)^2 \Delta t. \quad (2.1)$$

It is biologically reasonable to insist that the mean and variance of the distances moved are finite. Hence

$$\sum_{\substack{j=-\infty \\ j \neq i}}^{\infty} \alpha(j, i) \Delta x, \quad \sum_{\substack{j=-\infty \\ j \neq i}}^{\infty} |j - i| \alpha(j, i) (\Delta x)^2 \quad \text{and} \quad \sum_{\substack{j=-\infty \\ j \neq i}}^{\infty} |j - i|^2 \alpha(j, i) (\Delta x)^3 \tag{2.2}$$

will all be assumed finite. Later these sums will be interpreted as integrals.

During this same time interval, the number of arrivals to site i from elsewhere is

$$\sum_{\substack{j=-\infty \\ j \neq i}}^{\infty} \alpha(i, j) u(j, t) (\Delta x)^2 \Delta t. \tag{2.3}$$

Finally, within each site we allow for the birth and death of individuals. Let $f(u(i, t), i)$ denote the per capita net reproduction rate at site i at the given population density. Again, we assume that this rate is constant over the time interval in question. Then the number of new individuals at site i is

$$f(u(i, t), i) u(i, t) \Delta x \Delta t. \tag{2.4}$$

Combining these expressions, we deduce that the population density at location i and time $t + \Delta t$ is given by

$$\begin{aligned} u(i, t + \Delta t) = & u(i, t) + \sum_{\substack{j=-\infty \\ j \neq i}}^{\infty} \alpha(i, j) u(j, t) \Delta x \Delta t \\ & - \sum_{\substack{j=-\infty \\ j \neq i}}^{\infty} \alpha(j, i) u(i, t) \Delta x \Delta t + f(u(i, t), i) u(i, t) \Delta t. \end{aligned} \tag{2.5}$$

There are several different limits which can be taken here, each of which results in a different model. Our interest is in the integro-differential equation obtained by allowing both $\Delta t \rightarrow 0$ and $\Delta x \rightarrow 0$, viz.

$$\begin{aligned} \frac{\partial u}{\partial t}(x, t) = & \int_{-\infty}^{\infty} [\alpha(x, y) u(y, t) - \alpha(y, x) u(x, t)] dy \\ & + f(u(x, t), x) u(x, t). \end{aligned} \tag{2.6}$$

It is worth observing that, unlike the derivation of the Laplacian, we are not imposing a constraint on the relative rates at which $\Delta t \rightarrow 0$ and $\Delta x \rightarrow 0$.

We shall assume throughout that the rate of transition between the various patches, $\alpha(x, y)$, is homogenous and only depends on the distance between patches i.e. upon $|x - y|$. Equation (2.6) may now be written as

$$\frac{\partial u}{\partial t} = f(u, x) u + \rho \left[\int_{-\infty}^{\infty} k(x - y) u(y) dy - u(x) \right] \tag{2.7}$$

where k is an even function with

$$\int_{-\infty}^{\infty} k(s) ds = 1, \quad (2.8)$$

and

$$\rho := \int_{-\infty}^{\infty} \alpha(|y|) dy. \quad (2.9)$$

2.2. Parameter dependence

As indicated in the introduction, we are concerned with how evolution might select dispersal parameters; particularly the *spread* and *rate* of diffusion. When we want α to explicitly depend upon a parameter, say τ , then we write α_τ . It is easy to quantify the notion of rate, it is just the multiplicative factor ρ already introduced in equations (2.9) and (2.7). To treat the idea of ‘spread’, it is reasonable to preserve the shape of the function α but allow it to be stretched or contracted. Thus, for a given function k , we are led to define the spread parameter L by the manner in which it modifies k to k_L . Specifically, we define

$$k_L(x) = \frac{1}{L} k\left(\frac{x}{L}\right). \quad (2.10)$$

If, for example, L is doubled, then the mean distance between natal and parental nest sites is doubled. Later, in Sections 2.3.2 and 2.3.3, we shall have a kernel function $K(x, y)$ derived from k and when this is required to depend explicitly upon L we shall write $K_L(x, y)$. The dependence of K upon L may be quite complicated but will always be based on the assumption that it is derived from the corresponding k_L given by equation (2.10), see equation (2.13) for example.

2.3. Finite region

Thus far the habitat has been considered to be infinite in extent. Without some qualifications, this leads to technical difficulties. Furthermore, it is biologically unrealistic. With this in mind we modify equation (2.7) to apply to a habitat in \mathbb{R} of length l which occupies the region $[0, l]$. There are two natural ways in which this may be done, and each is considered in turn in the following subsections.

2.3.1. Hostile surroundings

Suppose that the habitat outside the interval $[0, l]$ is so hostile that all individuals which land there immediately die. Then the density may be set to zero outside this region and so equation (2.7) becomes

$$\frac{\partial u}{\partial t} = \rho \left\{ \int_0^l k(x-y)u(y) dy - u \right\} + f(u, x)u. \quad (2.11)$$

Notice that $\int_0^l k(x-y) dy$ is less than unity for each x unless the support of k is contained in $[-l/2, l/2]$.

When the kernel k depends upon the spread parameter L as in equation (2.10) then equation (2.11) becomes

$$\frac{\partial u}{\partial t} = \rho \left\{ \frac{1}{L} \int_0^L k\left(\frac{x-y}{L}\right) u(y) dy - u \right\} + f(u, x)u. \tag{2.12}$$

2.3.2. *Periodic environment*

It is easy to believe that with hostile surroundings the behaviour near the boundary may play an important role in the dynamics. The simplest way to exclude these effects is to assume that the environmentally dependent reproduction rate $f(u, x)$ and the density $u(x)$ are periodic in x with period l . In this case

$$\begin{aligned} \int_{-\infty}^{\infty} k(x-y)u(y) dy &= \sum_{i=-\infty}^{\infty} \int_{il}^{(i+1)l} k(x-y)u(y) dy \\ &= \sum_{i=-\infty}^{\infty} \int_0^l k(x-y-il)u(y+il) dy \\ &= \int_0^l u(y) \left[\sum_{i=-\infty}^{\infty} k(x-y+il) \right] dy. \end{aligned}$$

When the spread parameter is included, k must be replaced by k_L as in equation (2.10) yielding the kernel K_L where

$$K_L(x) = \frac{1}{L} \sum_{i=-\infty}^{\infty} k\left(\frac{x+il}{L}\right). \tag{2.13}$$

The model equation may be written as

$$\frac{\partial u}{\partial t} = f(u, x)u + \rho \left\{ \int_0^l K_L(x-y)u(y) dy - u(x) \right\}, \quad (0 \leq x \leq l). \tag{2.14}$$

Note that the kernel function K_L is non-negative, periodic (with period l) and satisfies

$$\int_0^l K_L(y) dy = 1. \tag{2.15}$$

Data suggests that there are a variety of different plausible forms for the transport function (see [22] and references therein). Following [34] we shall sometimes choose k to be a back-to-back exponential. Its dependence upon the spread parameter L , see equation (2.10), is then given by

$$k_L(x) = \frac{\exp(-|x|/L)}{2L} \quad (x \in \mathbb{R}). \tag{2.16}$$

It is now easy to show that

$$K_L(x) = \frac{\cosh \frac{|x| - l/2}{L}}{2L \sinh \frac{l}{2L}} \quad (|x| \leq l). \tag{2.17}$$

2.3.3. *Reflecting boundary*

Instead of the environment being periodic, there is a different way in which periodicity may occur. Suppose that the animals (or plants) are confined to the region $x \in [0, l]$ and that whenever they reach a boundary they are reflected back into this region. Then, provided only that

$$\frac{\partial f}{\partial x}(u, 0) = 0 = \frac{\partial f}{\partial x}(u, l), \tag{2.18}$$

the problem is equivalent to one in which the environment has period $2l$ and is symmetrical about $x = 0$. This model has the merit of preventing any loss of animals at the boundaries. The dispersal kernel is obtained in a similar manner to that for the periodic environment and the result is

$$\frac{\partial u}{\partial t} = \rho \left\{ \int_0^l K(x, y)u(y) dy - u \right\} + f(u, x)u \tag{2.19}$$

where

$$K(x, y) = \sum_{i=-\infty}^{\infty} [k(x + y + 2il) + k(x - y + 2il)].$$

2.4. *The dispersal budget*

Suppose that the amount of energy per individual that the organisms can expend upon dispersal is fixed (because of environmental and developmental constraints). As an example, consider a tree producing and dispersing seed. Evolutionarily speaking, the species can ‘choose’ to disperse a few seeds over a long distance or many over a short distance or some compromise. Similar considerations may be applied to animals or birds moving away from their natal site and finding their own site to reproduce. We assume that the costs involved in dispersal are proportional to:

- the number of individuals dispersed,
- a nondecreasing, even function F of the distance moved.

Now the amount of seed transferred from site i to site j in time Δt is

$$\alpha(j, i)u(i, t)(\Delta x)^2 \Delta t.$$

Thus, the cost associated with this is

$$F(i - j)\alpha(j, i)u(i, t)(\Delta x)^2 \Delta t.$$

With the above assumptions, the total cost (or energy consumption) in time Δt for a typical site is

$$\int_{-\infty}^{\infty} F(y)\alpha(y) dy [u(x, t)\Delta x]$$

and so if the amount of energy expended per individual is fixed then

$$\int_{-\infty}^{\infty} F(y)\alpha_L(y) dy = \text{constant}. \quad (2.20)$$

Our ultimate goal is to compare different dispersal strategies. To make a fair comparison it is reasonable to assume that the cost function F is fixed, but allow the transport function α (and hence k) to depend upon a parameter L in the manner shown in equation (2.10). Thus

$$\alpha_L(x) = \rho(L) \frac{k(x/L)}{L}. \quad (2.21)$$

where, as in equation (2.8), the integral of k is unity. Let us choose

$$F(y) \propto |y|^m \quad (m \geq 0) \quad (2.22)$$

as the simplest possibility. Equations (2.20), (2.21) and (2.22) imply that

$$\rho(L) = \rho_0 L^{-m}, \quad (2.23)$$

where ρ_0 is a constant.

Fixing the energy budget means that, evolutionarily speaking, the species has the choice between large L , in which the dispersal rate is small but the seeds move large distances, and small L (larger dispersal rate but the seeds are spread over smaller distances). Our objective will be to determine which option is successful, or whether some intermediate value of L is selected.

The dispersal will always depend upon the two parameters ρ_0 and L . Additionally it may depend upon m , the cost parameter, but often this will be taken to be zero. As mentioned in the introduction, ρ (or ρ_0 when $m = 0$) is called the *rate* of dispersal. It measures the rate of flow of the animals (or seeds) at any point. As explained in Section 2.2, L is a measure of the *spread* of the animals.

2.5. Summary of the model

The governing equation for the population dynamics of a single species is taken to be

$$\frac{\partial u}{\partial t} = uf(u, x) + Du, \quad (2.24)$$

where D is the dispersal operator. We shall concentrate on two models subsequently; the hostile surroundings case, where Du is given in (2.12), and the periodic case, where Du is given in (2.14). In each case $\rho = \rho_0 L^{-m}$ from equation (2.23).

The reflecting boundary case, see Section 2.3.3, may also be treated very similarly. In both cases we have

$$\int_0^l Du \, dx \leq 0. \quad (2.25)$$

For the periodic (and reflecting boundary) cases the number of individuals is preserved by dispersal and there is equality in (2.25). For the hostile surroundings, if $k > 0$ the inequality is strict. There will be certain similarities with the classical Laplacian operator with zero Dirichlet or mixed boundary conditions, whereas the periodic case will be more similar to the case where there is a zero Neumann condition.

Although it is possible to treat a more general case in which f satisfies the conditions (1.4), we shall, for simplicity, assume from here on that the reaction term is of the logistic form (1.5).

It is convenient to list here certain conditions which are assumed in the sequel unless there is a specific indication to the contrary.

(C1) $a \in C^1[0, l]$ is strictly positive and, in the periodic case, $a \in C^1(\mathbb{R})$.

(C2) $k \in C^1[0, \infty)$ is even and strictly positive.

(C3) $k(x)$ and $x^2k(x) \in L^2(\mathbb{R}^+)$. These conditions are required by (2.2).

Relaxing the condition $a(x) > 0$ for all $x \in \Omega$ still allows much progress to be made, but this condition is assumed for simplicity. We need to develop some theory in the next section and in order to do this, without unnecessarily restricting the generality, we write the dispersal operator in terms of a fairly general integral operator. We define

$$D = \rho E = \rho\{X - I\} \quad (2.26)$$

where

$$(Xu)(x, t) = \int_0^l \beta(x, y)u(y, t) \, dy. \quad (2.27)$$

3. Some basic theory for non-local dispersal

In Section 2 we derived various parameterized classes of kernel, a particular choice of parameter, say τ , representing a dispersal phenotype. We shall eventually consider a τ -phenotype at equilibrium and investigate whether a τ' -phenotype, introduced in small numbers (say by mutation) into the population, can invade. For a dispersal process modelled by the Laplacian, the technicalities needed are extremely well known, see [4] for example. However, non-local processes governed by integral operators differ in some important respects and we must now outline certain basic elements of the theory needed. In this section we discuss the scalar case and in Section 4 the system. In order not to interrupt too much the continuity of the account, we have placed some of the proofs in an appendix. We first consider the analogue of the idea of a principal eigenvalue. It is well known that there is a

maximum principle, see [9] for example. We next derive a convenient form of this and use it to give conditions for the existence of an equilibrium.

For illustration, let $\Omega = [0, l]$ where $0 < l < \infty$ but note that the theorems and proofs in this section extend almost exactly to an n -dimensional cube in \mathbb{R}^n . For $1 \leq p \leq \infty$, let $\|\cdot\|_p$ denote the L^p norm and let \mathcal{H} be the Hilbert space L^2 with inner product (\cdot, \cdot) . Let C, C^1 be the usual Banach spaces of continuous and once differentiable functions $\Omega \rightarrow \mathbb{R}$.

We list the following conditions, with X defined by equation (2.26):

- (H1) $\beta : \Omega \times \Omega \rightarrow \mathbb{R}$ is continuous, symmetric and non-negative.
- (H2) $X : L^p \rightarrow C^1$ is continuous. (Note that this condition holds for the kernels discussed in Section 2, including $\exp(-|x|)$, or if β is C^1 .)
- (H3) $\int_{\Omega} \beta(x, y) dy \leq 1$ ($x \in \Omega$).
- (H4) $\beta(x, y) > 0$ ($x, y \in \Omega$).

The concept of asymptotic stability (i.e. local stability) of an equilibrium is crucial in the present investigation. The key fact is that there is a close analogy with the idea of a principal eigenvalue (PEV) for linear elliptic operators. However, because (as discussed in the introduction) there is a difference in the ‘compactness’ behaviour of the dispersal operators, we need to prove the above assertion with care. The proof of the next theorem is given in Appendix B.1.

We consider $X, H : \mathcal{H} \rightarrow \mathcal{H}$ where

$$Hu(x) = h(x)u(x). \tag{3.1}$$

Let $\sigma(\cdot), \sigma_p(\cdot)$ denote the spectrum and point spectrum respectively.

Theorem 3.1. *Assume that (H1) holds and $h \in C$. Then H, X and M are self-adjoint, where $M = \rho X + H$. Suppose in addition that h is Lipschitz and $\beta(x, x) > 0$ ($x \in \Omega$). Then M has a PEV λ_0 given by*

$$\lambda_0 = \max_{\|u\|_2=1} (Mu, u), \tag{3.2}$$

the maximum being attained for a strictly positive, continuous eigenfunction ϕ , say, which is unique. Also $\sigma(M) \subset (-\infty, \lambda_0]$.

The eigenfunction problem for M is not, so far as the authors are aware, completely standard, and we therefore add the following remarks concerning the conditions on h and β in the theorem. As a glance at the proof reveals, there are a number of alternative conditions which would yield the result and our choice is simply for convenience in the present context. However, it is worth noting that if the Lipschitz condition on h is replaced by continuity then an example may be easily found for which the eigenvalue problem has no positive eigenfunction. A similar remark applies if the condition $\beta(x, x) > 0$ is weakened to $\beta(x, x) \geq 0$ and β non-trivial. It is inappropriate to pursue these questions in the present context.

It will be assumed in the rest of this section that (H1)–(H4) all hold. For $T > 0$, let $S_T = \Omega \times (0, T)$, $Q_T = \Omega \times \{T\}$ if $T < \infty$, otherwise $Q_T = \emptyset$.

Theorem 3.2. (Strong Maximum Principle). *Take $T > 0$. Assume that $\tilde{c} : \Omega \times [0, \infty)$ is continuous.*

Suppose that $u(x, 0) \geq 0$, u is continuous in x and C^1 in t and u satisfies the following equation:

$$\frac{\partial u}{\partial t} \geq Du + \tilde{c}(x, t)u.$$

Then the following alternative holds.

Either $u > 0$ in $Q_T \cup S_T$

or $\exists t^* \leq T$ such that $u(x, t) = 0$ ($x \in \Omega, 0 \leq t \leq t^*$) and $u(x, t) > 0$ ($x \in \Omega, t > t^*$).

Proof. See Appendix B.2. □

We consider the (scalar) equation

$$\frac{du}{dt} = Du + u(a - u), \tag{3.3}$$

with initial condition $u(x, t) = u(x, 0)$. By a solution (or subsolution or supersolution) we mean a function that is continuous in x and C^1 in t . Following [6], the notation $u \in C^1([0, \infty), C)$ will be used. We now develop the analogue of the sub/super solution method for the scalar parabolic case.

Definition 1. *A function \underline{u} with the above smoothness condition properties is said to be a subsolution in Q_T if*

$$\frac{d\underline{u}}{dt} \leq D\underline{u} + \underline{u}(a - \underline{u}). \tag{3.4}$$

A supersolution is defined similarly by reversing the inequality.

The following result is analogous to the classical theory (for example [9], Theorems 4.1 and 4.2 or [33], pp. 54, 55). Since essentially the same maximum principle holds (Theorem 3.2), the proofs of Theorems 3.3 and 3.4 are minor amendments and are omitted.

Theorem 3.3. (Comparison). *Let \underline{u}, \bar{u} be sub/supersolutions with $\underline{u}(x, 0) \leq \bar{u}(x, 0)$. Then either*

$$\underline{u} < \bar{u} \quad (\text{in } S_T \cup Q_T)$$

or, for some $t^ > 0$,*

$$\underline{u} = \bar{u} \quad (\text{in } S_{t^*} \cup Q_{t^*}).$$

Theorem 3.4. *Let $\underline{u}(x)$ be a stationary subsolution and u a solution with $u(x, 0) = \underline{u}(x)$. Then either \underline{u} is a solution or u is strictly increasing in t for each fixed x .*

Theorem 3.5. (Existence). *Suppose that $u(x, 0)$ is continuous and non-negative. Then equation (3.3) has a continuous, non-negative solution for all $t > 0$.*

Proof. Obviously a large constant, c , is a supersolution and zero is a subsolution, so $0 \leq u(x, t) \leq c$. We have an *a priori* bound and global existence in L^∞ follows from [6]. The continuity is just a consequence of the continuous dependence on initial conditions for an ordinary differential equation. \square

Theorem 3.6. *Suppose that $\underline{u}(x)$ is a non-trivial, non-negative subsolution. Then equation (3.3) has exactly one, strictly positive, C^1 , stationary solution. This is globally stable (for solutions with non-trivial initial conditions) in the sense of pointwise convergence.*

Proof. Since there is a constant supersolution, from Theorem 3.5 a continuous solution $u(x, t)$ with $u(x, 0) = \underline{u}(x)$ exists and from Theorem 3.4 it is non-decreasing. Now note that $\underline{u}(x)$ cannot be a solution if $\underline{u}(x_0) = 0$ for some x_0 . For if it were, putting $x = x_0$ in equation (3.3) gives

$$\int_{\Omega} \beta(x_0, y)\underline{u}(y) dy = 0,$$

which is impossible as $\beta > 0$ and \underline{u} is non-trivial. With this in view, we can see from Theorem 3.4 that $u(x, t) > 0$ ($x \in \Omega, t > 0$).

We thus have an increasing sequence $\{u(x, t_n)\}$, say of continuous functions which is bounded above and thus converges to an L^∞ function \tilde{u} (because continuous functions are measurable, and a convergent sequence of measurable functions is measurable and hence L^∞). From the monotone convergence theorem (proceeding to a sequence if necessary) we see that

$$\lim_{t \rightarrow \infty} \int_{\Omega} \beta(x, y)u(y, t) dy = \int_{\Omega} \beta(x, y)\tilde{u}(y) dy.$$

It follows that

$$\lim_{t \rightarrow \infty} Du(x, t) = (D\tilde{u})(x). \tag{3.5}$$

We claim that \tilde{u} is an equilibrium. To see this let

$$(Fu)(x, t) = \rho \int_{\Omega} \beta(x, y)u(y, t) dy + u(x, t)[a(x) - u(x, t) - \rho],$$

so that equation (3.3) is just $\partial u / \partial t = Fu$. We wish to show that $F\tilde{u} = 0$. Suppose the contrary. Then $\exists x_0 \in \Omega$ such that $(F\tilde{u})(x_0) > 0$. Thus from equation (3.5), $\exists \delta > 0$ and t_0 such that $F\tilde{u}(x, t) \geq \delta$ for $t \geq t_0$. Therefore

$$\frac{\partial u}{\partial t}(x_0, t) \geq \delta \quad (t \geq t_0),$$

and it follows that $u(x_0, t) \rightarrow \infty$ as $t \rightarrow \infty$. This contradicts the definition of \tilde{u} .

To show that \tilde{u} is C^1 , note first that since it is L^∞ and β is continuous, certainly $X\tilde{u}$ is continuous. Now \tilde{u} satisfies the equation

$$\tilde{u}^2 - (a + \rho)\tilde{u} - \rho X\tilde{u} = 0,$$

and the positive solution of this is

$$\tilde{u} = \frac{a + \rho + \sqrt{(a + \rho)^2 + 4\rho X\tilde{u}}}{2}.$$

The differentiability of \tilde{u} follows from that of a and $X\tilde{u}$.

The proof of uniqueness follows from a contradiction argument. Suppose that there are two positive solutions u_1 and u_2 . Set $w = u_1 - u_2$. Then

$$\begin{aligned} Du_1 + u_1(a - u_1) &= 0, \\ Dw + w(a - u_1 - u_2) &= 0. \end{aligned}$$

With $L = D + a - u_1$, these may be written as

$$\begin{aligned} Lu_1 &= 0, & (3.6) \\ Lw - u_2w &= 0. & (3.7) \end{aligned}$$

Equation (3.6) shows that zero is the principal eigenvalue of L . From equation (3.7) the principal eigenvalue of $L - u_2$ is also zero. However, this contradicts the variational principle 3.4 applied successively with $h = a - 1 - u_1$ and $h = a - 1 - u_1 - u_2$.

To prove the global stability, note first that if \underline{u} is a subsolution then so is $\epsilon\underline{u}$ for $0 \leq \epsilon \leq 1$. Clearly, if $u(x, 0)$ is non-trivial then $u(x, t) > 0$ for $t > 0$ (from Theorem 3.4). Hence one may choose t_0 and $\epsilon > 0$ such that $u(x, t) > \epsilon\underline{u}(x)$ for $t \geq t_0$. Changing the time origin to t_0 , it is clear that $u(x, t)$ lies between solutions u_1, u_2 with $u_1(x, 0) = \epsilon\underline{u}(x)$ and $u_2(x, 0) = C$, where C is a large constant, and u_1, u_2 are non-decreasing, non-increasing respectively. These tend to solutions as $t \rightarrow \infty$ and the global stability follows from the uniqueness. \square

Remark. Of course the condition that the PEV of the linearization about zero is positive, is sufficient for the existence of a subsolution.

Example 3.7 (The periodic case (including reflecting boundaries)). From equation (2.15), $\int_{\Omega} \beta(x, y) dy = 1$ for $x \in \Omega$. With $\underline{u} = \epsilon$, $D\underline{u} = 0$ and so

$$D\underline{u} + \underline{u}(a - \underline{u}) = \epsilon(a - \epsilon) > 0$$

for small enough ϵ . Therefore, from Theorem 3.6, there is a unique, positive solution \tilde{u} . Suppose now that $D = \tau E$ where $\tau > 0$ and E is fixed. Then it is easy to prove the following by the contraction mapping principle since E is a bounded operator. In the sense of C ,

$$\tilde{u} = a + O(\tau) \quad (\tau \ll 1). \tag{3.8}$$

Example 3.8 (Hostile Surroundings). Suppose that

$$Du = \rho_0 L^{-m}(X_L u - u).$$

With $m = 0$, \tilde{u} exists for small enough ρ_0 . To see this, note that

$$(Du)(x) \geq -\rho_0 u(x) \quad (x \in \Omega).$$

Hence, with $\underline{u} = \epsilon$,

$$\begin{aligned} D\underline{u} + \underline{u}(a - \underline{u}) &\geq -\rho_0\underline{u} + \underline{u}(a - \underline{u}) \\ &= \epsilon(a - \rho_0 - \epsilon) \\ &> 0 \end{aligned}$$

for $\rho_0 < \min_{x \in \Omega} a(x)$ and ϵ sufficiently small.

In contrast with the first case, if $m \neq 0$ it is not at all clear whether or not such a solution exists for all L . In fact there is numerical evidence (for the special case when the kernel is based upon $\frac{1}{2} \exp(-|x|)$ and $m = 1$) that the solution \tilde{u} may disappear and then reappear as L decreases! For this reason, we shall confine the investigation of the hostile surroundings case to the relatively easy case of $m = 0$.

4. Invasion by a mutant

We lay down here the basic structure for the analysis of dispersal. Consider a species whose dispersal depends smoothly on a dispersal parameter τ , and suppose that the species is at its non-zero equilibrium, \tilde{u} , say. Assume that a small number of a different phenotype with dispersal parameter τ' is introduced, possibly by mutation. We enquire whether or not the new phenotype will invade. This is done by carrying out a local stability analysis. Thus we consider the system

$$D(\tau)u + u(a - u - v) = \frac{\partial u}{\partial t}, \tag{4.1}$$

$$D(\tau')v + v(a - u - v) = \frac{\partial v}{\partial t}, \tag{4.2}$$

and enquire whether the equilibrium $(\tilde{u}, 0)$ is asymptotically stable.

Define linear operators L_1 and L_2 by setting

$$L_1(\tau)u = D(\tau)u + (a - \tilde{u})u, \tag{4.3}$$

$$L_2(\tau)u = L_1(\tau)u - \tilde{u}u. \tag{4.4}$$

Noting that both of these operators are of the form treated in Theorem 3.1, we let $\lambda(\tau)$ denote the PEV of $L_1(\tau)$. By definition, \tilde{u} satisfies the equation

$$D(\tau)\tilde{u} + (a - \tilde{u})\tilde{u} = 0, \tag{4.5}$$

which implies that \tilde{u} is the principal eigenfunction corresponding to the PEV zero. Therefore

$$\lambda(\tau) = 0. \tag{4.6}$$

A standard linearization procedure leads to the operator $L : \mathcal{H} \times \mathcal{H} \rightarrow \mathcal{H} \times \mathcal{H}$, where

$$L(u, v) = \begin{bmatrix} L_2(\tau)u & -\tilde{u}v \\ & L_1(\tau')v \end{bmatrix}. \tag{4.7}$$

This is of upper-triangular form and one may expect that the asymptotic stability of $(\tilde{u}, 0)$ is determined by $\lambda(\tau')$. This is confirmed by the following result which is proved in Appendix B.3.

Lemma 4.1. *The equilibrium $(\tilde{u}, 0)$ is asymptotically stable in C if $\lambda(\tau') < 0$ and unstable if $\lambda(\tau') > 0$.*

The biological consequence is as follows.

Note 4.1. By definition $\lambda(\tau')$ is the PEV of the equation

$$D(\tau')\phi + (a - \tilde{u})\phi = \lambda(\tau')\phi. \tag{4.8}$$

The phenotype with dispersal parameter τ' can invade from small numbers if $\lambda(\tau')$ is positive and will go to extinction if it is negative.

We shall see in the next section that in a certain class of cases, we can immediately deduce the sign of $\lambda(\tau)$. However, in general this will be analytically very difficult and we therefore now outline our approach for the case when $(\tau - \tau')$ is small. Biologically this is not unreasonable as it corresponds to a small phenotypic change. The analysis is formal (but could be justified rigorously) and is based on the broad assertion that an expansion in powers of ϵ , where $\epsilon = \tau' - \tau$, is valid. Recalling that $\lambda(\tau) = 0$, we write

$$\begin{aligned} \phi &= \tilde{u} + \epsilon\phi_1 + \dots, \\ \lambda(\tau + \epsilon) &= \epsilon\lambda_1(\tau) + \epsilon^2\lambda_2 + \dots, \\ D(\tau + \epsilon) &= D(\tau) + \epsilon D_1(\tau) + \dots \end{aligned}$$

Here $D_1(\tau)$ is a linear operator, the Frechét derivative of D with respect to τ evaluated at τ . Substituting into equation (4.8) and picking out the terms in ϵ^0, ϵ^1 we see that the first is satisfied because of equation (4.5) and we obtain the equation

$$D(\tau)\phi_1 + (a - \tilde{u})\phi_1 = \lambda_1(\tau)\tilde{u} - D_1(\tau)\tilde{u}. \tag{4.9}$$

Multiplying by \tilde{u} , integrating over Ω and using self-adjointness, we obtain the result

$$\lambda_1(\tau) = \frac{(D_1(\tau)\tilde{u}, \tilde{u})}{\|\tilde{u}\|^2}. \tag{4.10}$$

Therefore, for small ϵ , $\lambda(\tau + \epsilon)$ has the same sign as $(D_1(\tau)\tilde{u}, \tilde{u})$ and if this is known we can immediately determine whether a mutant with slightly different dispersal parameter will invade. In other words we can determine whether, in an evolutionary sense, a larger or smaller dispersal parameter is preferred.

It has to be emphasised that the analysis is local in two senses: the number of invaders is small and the phenotypic difference between the invaders and the residents is small. The global questions are a great deal harder than for the diffusion case treated in [4].

An important remark is that $\lambda_1(\tau_0) = 0$ when $(D_1(\tau_0)\tilde{u}, \tilde{u}) = 0$. We may then call τ_0 an *evolutionary equilibrium*, which will be an evolutionary attractor or repeller depending on whether $\lambda_2(\tau)$ is negative or positive.

The tactics that we have used to discuss the evolution of dispersal have similarities with evolutionarily stable strategies (ESS) and with adaptive dynamics. A criticism that can be levelled at the ESS approach is, to quote [26], that ‘only the

statics of adaptive evolution was considered'. Indeed, [26] suggest that evolutionarily unbeatable strategy is a better name for the fundamental entity involved. Although adaptive dynamics does, as its name suggests, attempt to overcome this shortcoming, it does not usually include a detailed model of population dynamics. Rather it is based upon the notion of fitness, see for example [16]. It is doubtful that such an approach is adequate in the context of dispersal.

5. The selection of dispersal rate

As outlined in the introduction, it is shown in [4] for the case of classical diffusion with dispersal $\rho \Delta u$, that small ρ is always selected. In the present case the first step must be to carry out an analogous investigation and consider the special case where the dispersal parameter ρ is simply a scaling factor multiplying a fixed non-local dispersal operator, which we designate by $E = X - I$, so that $D = \rho E$. It is convenient to abbreviate and use the term 'dispersal rate' for ρ . Remarkably, our analysis suggests exactly the same conclusion as [4]: the smaller rate is always selected. This result is surprisingly robust and requires only mild restrictions on the details of the dispersal.

Take then

$$(Eu)(x) = \int_{\Omega} \beta(x, y)u(y) dy - u(x) \quad (5.1)$$

$$Du = \rho Eu = \rho(Xu - u), \quad (5.2)$$

and consider the system

$$\frac{du}{dt} = \rho Eu + u(a - u - v), \quad (5.3)$$

$$\frac{dv}{dt} = \rho' Ev + v(a - u - v). \quad (5.4)$$

The following lemma, a generalisation of [4], Lemma 2.1, rapidly settles the issue. From Note 4.1, the asymptotic stability of $(\tilde{u}, 0)$ is determined by the sign of $\lambda(\rho')$, the PEV of the equation

$$\rho' E\phi + h\phi = \lambda(\rho')\phi, \quad (5.5)$$

where $h = a - \tilde{u}$ and \tilde{u} is the solution to

$$\rho E\tilde{u} + \tilde{u}(a - \tilde{u}) = 0. \quad (5.6)$$

Lemma 5.1. $(\rho' - \rho)\lambda(\rho') < 0$ for $\rho \neq \rho'$.

Proof. Suppose that $\rho' > \rho$. (The case $\rho' < \rho$ is similar.) From Theorem 3.1,

$$\begin{aligned} \lambda(\rho) &= \max_{\phi \in C} \frac{\rho(E\phi, \phi) + (h\phi, \phi)}{\|\phi\|^2} \\ &= \max_{\phi \in C} \frac{(\rho - \rho')(E\phi, \phi) + \rho'(E\phi, \phi) + (h\phi, \phi)}{\|\phi\|^2} \\ &\geq \frac{(\rho - \rho')[(E\phi(\rho'), \phi(\rho')) + \rho'(E\phi(\rho'), \phi(\rho')) + (h\phi(\rho'), \phi(\rho'))]}{\|\phi(\rho')\|^2} \\ &= \lambda(\rho') + (\rho - \rho')(E\phi(\rho'), \phi(\rho')), \end{aligned} \tag{5.7}$$

where $\phi(\rho')$ has been used as a trial function in the maximisation. Clearly the result will follow if can be shown that $(E\phi(\rho'), \phi(\rho')) < 0$.

From Lemma A.1, this inequality holds unless $\phi(\rho') = 1$ and $(E1)(x) = 0$ for $x \in \Omega$. But in this case there is equality and from equation (5.7) we have $0 = \lambda(\rho) \geq \lambda(\rho')$. If $\lambda(\rho') = 0$ then from equation (5.5) $h = 0$ and $\tilde{u} = a$. Equation (5.6) now implies that $Ea = 0$ and hence, from Lemma A.1, a is a constant which contradicts (H2). \square

Of course, from equation (5.6), $\lambda(\rho) = 0$, so the lemma shows that $\lambda(\rho') < 0$ if $\rho' > \rho$. We further remark that a slight extension of this result, taking $h = a - \tilde{u} - \tilde{v}$ instead, shows that there cannot be a stationary co-existence solution to equations (5.3) and (5.4). In the case of the analogous reaction-diffusion system, it was proved in [4] that, as a consequence, $(\tilde{u}, 0)$ is globally stable (for solutions with non-trivial initial values) and it is natural to enquire whether the same result also holds in the present case. The following result goes some way towards establishing this, but the basin of attraction is severely limited. It appears that a global result may be proved, but the details are complicated, and it is not appropriate to present them here.

The following convention is used for two functions $u_1, u_2 : u_1 \leq u_2$ means that $u_1(x) \leq u_2(x)$ ($x \in \Omega$), and $u_1 < u_2$ implies that, in addition, $u_1(x) \neq u_2(x)$ for some $x \in \Omega$. The proof of the following lemma is given in Appendix B.

Lemma 5.2. *Assume that $\rho' > \rho$. Suppose that $(u_0, v_0) \in C \times C$ satisfies the relations*

$$0 < u_0 < \tilde{u}, \quad 0 < v_0 < \tilde{v}.$$

Let (u, v) be the solutions of equations (5.3) and (5.4) with initial value (u_0, v_0) . Then $(u, v) \rightarrow (\tilde{u}, 0)$ pointwise.

We have thus proved that qualitatively exactly the same result holds for dispersal operators $D = \rho E$ as for the Laplacian. That is *smaller dispersal rate is always selected* (this refers to a small number of invaders, i.e. it is a local result). This holds for any biologically reasonable dispersal. The result is counter-intuitive from a biological point of view, although in some circumstances (birds on islands) lower dispersal appears to be selected very strongly, see [7]. In the next section, we proceed with the plan outlined in the introduction of considering a parameterized family of kernels, where the parameter determines the spread rather than the rate of dispersal, and examine the issue of selection.

6. The evolution of the dispersal spread – periodic environment

We showed in the last section that if only the rate of dispersal is considered then reduced rate is always favoured by evolution. We now consider variation in the dispersal width or spread, which was introduced in Section 2, while fixing the rate ρ_0 . We concentrate upon what is mathematically the easier case, that is the periodic environment, which includes the case of no migration across the boundary (and reflecting boundary). This has similarities with the case of zero Neumann conditions for the Laplacian. It is difficult to make any progress theoretically unless we assume that $(\rho - \rho')$ is small, and we do this here.

For definiteness, let us recall the notation from Section 2. We start with a dispersal mechanism on an infinite interval with kernel k_L . The dispersal operator D is then defined by

$$(Du)(x) = \rho_0 L^{-m} \left\{ \int_{\Omega} K_L(x - y)u(y) dy - u(x) \right\} \tag{6.1}$$

where ρ_0 and m are constants with $\rho_0 > 0$, $m \geq 0$ and K_L is given as in equation (2.13). The object in the rest of this section is to determine the sign of $\lambda_1(\rho)$ from equation (4.10) for the cases $L \ll 1$ and $L \gg 1$. The argument is heuristic.

6.1. $L \gg 1$

We first remark that if $m > 0$ then

$$\tilde{u}(x) \approx a(x) \quad (L \rightarrow \infty). \tag{6.2}$$

To see this, note that to a first approximation

$$\lim_{\epsilon \rightarrow \infty} \epsilon \sum_{n=-\infty}^{\infty} \alpha(n\epsilon) = \int_{-\infty}^{\infty} \alpha(x) dx = \rho.$$

Taking $\epsilon = l/L$, one finds that for large L

$$(Du)(x) \approx \rho_0 L^{-m} Eu(x)$$

where

$$Eu(x) = \frac{1}{l} \int_{\Omega} u(s) ds - u(x).$$

The approximation (6.2) follows from Example 3.7. When $m = 0$, an asymptotic estimate is somewhat more difficult to obtain, but is not needed in the following argument.

Let

$$\begin{aligned}
 u_{c,n} &= \frac{1}{l} \int_0^l \tilde{u}(x) \cos\left(\frac{2\pi nx}{l}\right) dx, \\
 u_{s,n} &= \frac{1}{l} \int_0^l \tilde{u}(x) \sin\left(\frac{2\pi nx}{l}\right) dx, \\
 \hat{k}(s) &= 4 \int_0^\infty k(x) \cos(sx) dx, \tag{6.3}
 \end{aligned}$$

$$\hat{k}_1(s) = 4s \int_0^\infty xk(x) \sin(sx) dx. \tag{6.4}$$

We impose rather strong assumptions on k , assuming that, in addition to (C1)–(C3), $xk(x) \in C^4[0, \infty)$ and $k^{(n)}$, the n th derivative of k , is $L^1(\mathbb{R}^+)$ for $0 \leq n \leq 4$. Integration by parts shows that for large s ,

$$\hat{k}(s) = O(s^{-2}), \tag{6.5}$$

$$\hat{k}_1(s) = -\frac{8}{s^2}k'(0) + O(s^{-3}). \tag{6.6}$$

We remark that an asymptotic approximation of the type (6.6) is only valid if one of the odd derivatives $k^{(2n+1)}(0) \neq 0$; otherwise it appears to be more difficult to obtain an approximation.

It is straightforward to show that λ has the same sign as

$$\begin{aligned}
 & -m \left[\left\{ \frac{1}{l} \int_\Omega \tilde{u}(x) dx \right\}^2 - \frac{1}{l} \int_\Omega \tilde{u}^2(x) dx \right] \\
 & - \sum_{n=1}^\infty \{u_{c,n}^2 + u_{s,n}^2\} \left\{ m\hat{k}\left(\frac{2\pi nL}{l}\right) + \hat{k}_1\left(\frac{2\pi nL}{l}\right) \right\}. \tag{6.7}
 \end{aligned}$$

Suppose first that $m > 0$ when, from the above remark, $\tilde{u} \rightarrow a$ as $L \rightarrow \infty$. Then it is easy to see from equations (6.5) and (6.6) that the summation term is small relative to the first term. We deduce from Schwarz’s inequality that this first term is positive and so λ is positive.

On the other hand, if $m = 0$ then, from equation (6.6),

$$\operatorname{sgn} \hat{k}_1(s) = -\operatorname{sgn} k'(0) \quad (s \gg 1). \tag{6.8}$$

If $k'(0) \neq 0$ then for large L every term in the summation in equation (6.7) has the same sign and it is clear that λ may take either sign. Simple examples are:

$$\begin{aligned}
 k(x) &= \frac{1}{2} \exp(-|x|) \quad \text{when } \lambda < 0, \\
 k(x) &= \frac{1}{2} |x| \exp(-|x|) \quad \text{when } \lambda > 0.
 \end{aligned}$$

6.2. The case $L \ll 1$

The argument is far from rigorous, but we later provide confirmatory numerical results. Unless x is near the boundary, one term is adequate in the summation for K_L and so

$$\begin{aligned} \frac{1}{L} \int_0^l K_L(x-y)u(y) dy &\approx \frac{1}{L} \int_0^l k\left(\frac{x-y}{L}\right)u(y) dy \\ &= \int_{-x/L}^{(l-x)/L} k(Y)u(x+LY) dY \\ &= \int_{-x/L}^{(l-x)/L} k(Y) \left[u(x) + LYu'(x) + \frac{L^2Y^2}{2}u''(x) + \dots \right] dY \end{aligned} \tag{6.9}$$

$$\approx u(x) + \frac{1}{2}L^2u''(x) \int_{-\infty}^{\infty} Y^2k(Y) dY + \dots \tag{6.10}$$

since k is even and is normalised by equation (2.8). Therefore, with c a positive constant, equation (6.1) becomes

$$(Du)(x) \approx c\rho_0L^{2-m}u''(x). \tag{6.11}$$

Also u satisfies the periodic boundary conditions

$$u(0) = u(l), \quad u'(0) = u'(l). \tag{6.12}$$

In the reflecting boundary case a similar argument gives

$$u'(0) = u'(l) = 0. \tag{6.13}$$

In both case we obtain the diffusion given by equation (6.11) together with a boundary condition leading to a self-adjoint operator. Therefore, when $m < 2$, we are back in the case considered by [4] (or a slight variant of it). We conclude that λ is negative when L is small and $m < 2$. That is *small dispersal spread is always favoured*.

6.3. Summary

A rather similar analysis applies for the case when the boundary is reflecting. The situation is summarised by Table 1.

For certain specific cases it may happen that $\hat{k}_1(s)$ is of one sign for all $s > 0$. For instance this occurs when $k(x) = \frac{1}{2} \exp(-|x|)$ and when $k(x) = \exp(-x^2)/\sqrt{\pi}$, where in each case $\hat{k}_1(s)$ is positive, and it follows that when $m = 0$, $\lambda < 0$ for all L . Usually $\hat{k}_1(s)$ will not be of constant sign and conclusions are harder to find.

For $m > 0$ progress may be made with Fourier series when $k(x) = \frac{1}{2} \exp(-|x|)$. When $a(x)$ is simply $2 + \cos(2\pi x)$ the curve in the (m, L) -plane upon which λ is zero is shown as the dashed curve in Fig. 1. However, by choosing $a(x)$ with Fourier components of higher order, it is possible to find examples in which λ changes sign more than once as L is increased. Such a case is illustrated in Fig. 1 by the solid curve. Clearly the fact that this curve can meander around the plane has important consequences for stability.

Table 1. Summary of behaviour for the periodic case.

L	m	$\text{sgn } \lambda$
$\gg 1$	> 0	$+1$
$\gg 1$	0	$\text{sgn } k'(0)$
$\ll 1$	$0 \leq m < 2$	-1

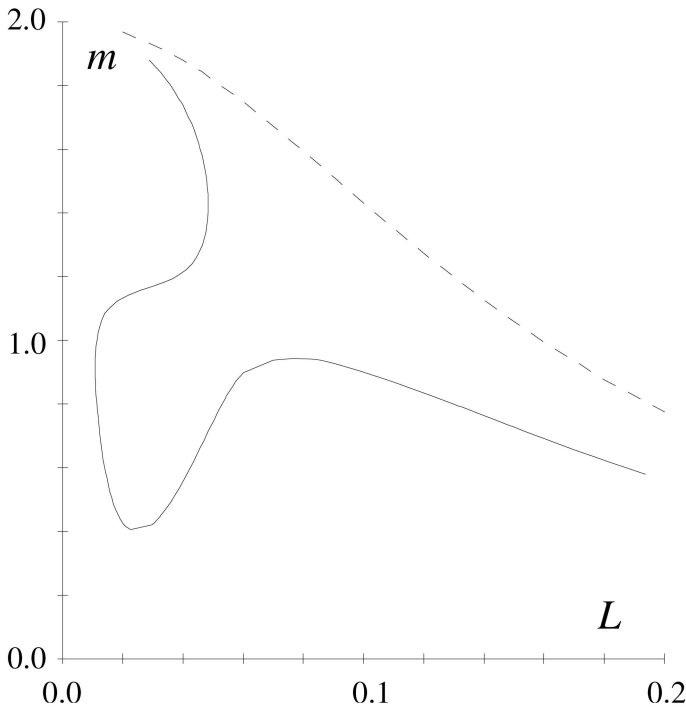


Fig. 1. Curves upon which λ is zero in the (L, m) -plane for two choices of a periodic environment.

7. Hostile surroundings

In this case animals/seeds will die if they pass across the boundary. We may thus expect that it is a disadvantage to have larger spread. Although this is generally the case, it is possible to construct examples where an *increase* in spread is selected. Indeed, there is an ‘evolutionarily unbeatable strategy’, see the discussion at the end of Section 4. We also note that in this case there may be co-existence for appropriate values of the parameters.

We remarked in Example 3.8 that if $m > 0$, that is there is a cost to dispersal (as discussed in Section 2), there may not even be a non-zero equilibrium solution. We

therefore restrict ourselves here to the case $m = 0$, and we assume, see Example 3.8, that $\rho_0 < \min_{x \in \Omega} a(x)$, so the semi-trivial solution \tilde{u} exists. The dispersal operator is then defined (using the notation of equations (2.10) and (2.8)) as follows;

$$(Du)(x) = \rho_0 \left\{ \frac{1}{L} \int_0^L k \left(\frac{x-y}{L} \right) u(y) dy - u(x) \right\}, \tag{7.1}$$

where, recall from Section 2, k is even, strictly positive, C^1 and

$$\int_{-\infty}^{\infty} k(z) dz = 1. \tag{7.2}$$

Suppose first that $L \gg 1$. Then from equation (4.10) the sign of λ is the same as

$$- \int_0^L \int_0^L \tilde{u}(x)\tilde{u}(y) \left\{ \left(\frac{x-y}{L} \right) k' \left(\frac{x-y}{L} \right) + k \left(\frac{x-y}{L} \right) \right\} dy dx. \tag{7.3}$$

Under the assumed smoothness condition on k , the bracketed term in (7.3) is clearly positive for large enough L , and we conclude that $\lambda < 0$ if L is large enough.

Consider next $L \ll 1$. From equation (6.10), for positive c , except for x close to 0 or l ,

$$(D\tilde{u})(x) \approx c\rho_0 L^2 \tilde{u}''(x). \tag{7.4}$$

Thus it is plausible that $\tilde{u} \rightarrow a$ in L^2 as $L \rightarrow 0$. In the notation of Section 6.1, λ always has the same sign as

$$- \int_0^\infty \hat{k}_1(yL) \{u_{c,y}^2 + u_{s,y}^2\} dy. \tag{7.5}$$

It is easy to show that, for small yL ,

$$\text{sgn } \hat{k}_1(yL) = \text{sgn} \left(\int_0^\infty x^2 k(x) dx \right) = 1$$

under the condition $y \int_y^\infty xk(x) dx \rightarrow 0$ as $y \rightarrow \infty$. If, as argued above, $\tilde{u} \rightarrow a$ then $u_{c,y}$ and $u_{s,y}$ will converge and be small for large y . Hence one may deduce that $\lambda < 0$ for $L \ll 1$. The argument yielding the sign of (7.5) could obviously be strengthened if the convergence of \tilde{u} could be proved. We shall not pursue this question here, but we remark that the analogous problem for the diffusion case is very well known.

We note that in the special case $k(x) = \frac{1}{2} \exp(-|x|)$, we can reduce the problem to an ODE and, with the help of Maple, prove that $\lambda < 0$ for all L . Numerical experiments on the kernel $\exp(-x^2)/\sqrt{\pi}$ confirm that $\lambda < 0$ for all L . However, this is not always the case. As an example, let

$$\begin{aligned} \tilde{u}(x) &= 1 + 3x^2 \quad (|x| \leq 1), \\ k(x) &= \frac{1}{2} [\delta(x - x_0) + \delta(x + x_0)]. \end{aligned}$$

It is now easy to use equation (7.5) to show that λ is positive for a finite interval of L . This example is artificial because \tilde{u} has been fixed rather than $a(x)$. However, since we know that λ is negative for large and small L , it does show that there is an $a(x)$ for which λ is positive somewhere, although the shape of the curve (as just calculated) will not be correct. The form of k above implies that all individuals disperse by exactly the same distance x_0 . It is clearly possible to approximate this artificial behaviour by a smooth, finite k and still retain the conclusion that λ changes sign. With the choices

$$\begin{aligned} a(x) &= 3.7 - 0.7 \cos(2\pi x) + 3 \cos(10\pi x) \quad (0 \leq x \leq 1) \\ k(x) &= |x| \exp(-x^2) \end{aligned}$$

the results of numerical computations are shown in Fig. 2 where λ evidently takes positive values. It is to be noticed that the bifurcation as λ passes through zero suggests that there can be stable co-existence.

We may broadly summarise the situation for the hostile environment case as follows. For both small and large dispersal spread L , smaller spread is selected. In all simple examples treated, this is true for all L . However, for some kernels k and environments $a(x)$ there will be a range of L in which larger spread is selected.

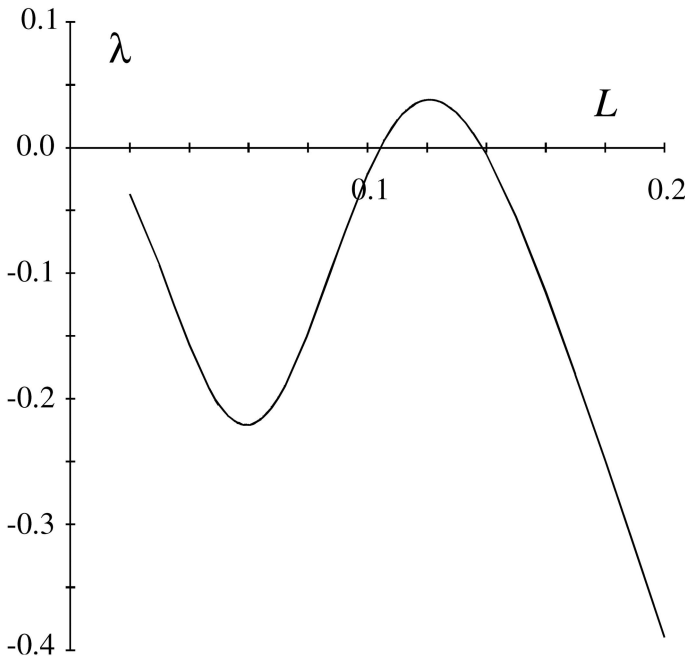


Fig. 2. λ (from 7.3) as a function of L when $k(x) = |x| \exp(-x^2)$ for a hostile environment.

8. Discussion

This paper has two principal aims. The first is to set up and study a continuous model of dispersal based upon an integral and the second is to consider the selection of dispersal in this framework. The mathematics is central in understanding the selection and the mathematical and biological points are inextricably intertwined. Nevertheless, it is convenient to comment on these issues separately.

8.1. Construction and properties of the model

Models (which are continuous in both space and time) for the dispersal of biological organisms have been discussed, for example, in [12] and [30]. While these derivations are of great interest, they are of a rather general nature and here we have derived in Section 2 a detailed model which takes into account the structure of dispersal and in particular defines the parameters L and ρ_0 for the spread and rate of dispersal. Also, in Section 2.4, we introduce the idea of a dispersal budget and the associated parameter m . Allowance is made for the effect of the environment by way of the function a and we are able to model situations in which it is periodic and also when there is a refuge surrounded by a hostile environment.

To put the problem in a broader mathematical context, we note that formally similar models have been studied in other physically motivated contexts, see [10], [11] and [38] and as a consequence there is a not inconsiderable amount of background theory available. Nevertheless, in order to study the problem under discussion, we need to develop the theory in a convenient form. The mathematics raises some interesting questions; for a valuable discussion see [6]. A big issue is the following. In contrast with the reaction-diffusion case, because the dispersal operator is bounded (and so has a less-strong effect on variations than the Laplacian) the semi-flow generated both in the scalar and system case do not have a smoothing effect and this makes for a theory which has important qualitative differences from the reaction-diffusion theory. If there is bistability, that is more than one attractor for the reaction system, the behaviour can be quite wild, see [6]. It is particularly interesting to speculate as to whether this represents a significant class of biological situations. In the context we consider here, there is exactly one attracting equilibrium. In this case much of the standard theory goes through. We note though that even the question of the existence of a principal eigenvalue is not as simple. Nonetheless there are some technical questions raised which at the present moment do not appear straightforward. For example, because of the lack of smoothing, asymptotic compactness does not hold, and even for a co-operative system it is not immediately possible to obtain the analogue of results along the lines of [35]. There appear to be several interesting, open lines of investigation.

8.2. Biological implications

The problem that we have concentrated upon is the selection of dispersal although we hope that our dispersal model will be used in other contexts. Of particular interest are situations where our model produces significantly different qualitative behaviour from the reaction-diffusion model.

As was indicated in the introduction we have avoided any comparison of the continuous models introduced with the patch models that are ubiquitous in the biological literature. We justified this on the grounds that a rigorous mathematical comparison is beyond the scope of this paper. This reasoning is, of course, not valid when we restrict our discussion to qualitative behaviour. It is clear that there are patch models which will reproduce the behaviour we have uncovered. This is especially true for the numerical results described below since the numerical method itself can be interpreted as a discrete patch model. On the other hand, one of the ‘disadvantages’ of using patch models is the extreme freedom that they provide. The integral model that we have presented allows for some of this freedom; there are very few constraints on the form of the integral kernel. Nevertheless, having fixed the kernel we have introduced two natural parameters, rate and spread, which we hope will provide useful constraints for the analysis of selection of dispersal.

We mentioned above the behaviour in the bistable case and it is unclear whether or not the pattern formation (via a Turing instability) will have a different flavour. However, we wish to close with some remarks and numerics on the selection of dispersal. We have shown in Section 5 that the selection of dispersal rate is almost certainly similar to the reaction-diffusion case treated in [4] and [13]. However, when it is the dispersal spread, L , that is being considered, the situation becomes much more complicated and interesting.

Consider first the easiest case, that is $m = 0$ (see Section 2.4). The biological idea of what we call ‘hostile surroundings’ is that of an island (or refuge or oasis) where organisms which leave the region die. If the dispersal spread is large, intuition suggests that this is deleterious and selection will favour a reduction in L ; this is established mathematically in Section 7. The case of L small is similar to that of small rate and again a reduction in L is selected. These ideas seem intuitively reasonable and one might hazard the guess that small L is always selected. However, this is *not* the case as we show by an example in Section 7. This suggests that, for an intuitive view to be correct, a much more sophisticated approach must be used and this is confirmed by our analysis of other cases. These remarks emphasize the importance of the mathematics.

The periodic case (Section 6) is important as it approximates a common situation where the region is very large and the effect of the boundary is smaller. With $m = 0$, we show that whilst still for $L \ll 1$, small L is selected as before, for $L \gg 1$ selection depends upon the kernel. In particular, with $k(x) = \frac{1}{2} \exp(-|x|)$ smaller L is selected but with $k(x) = \frac{1}{2}|x| \exp(-|x|)$ it is larger L which is selected. This seems hard to understand intuitively from the biology.

Let us now turn to cases where there is a dispersal ‘cost’, that is $m > 0$. We note that this raises the possibility of a ‘choice’ of whether it is better for a large numbers to move a small distance or small numbers a large one. The situation for small L is unchanged; smaller L is always selected. However, when L is large then *larger* L is selected. Note that this is in complete contrast to the selection of dispersal rate. Whatever the intuitive picture here, it is further complicated by the results of the numerical case mentioned in Section 6. The somewhat bizarre nature of Fig. 1 shows how complex the situation might be. Notice that, for a fixed value of m , the selection direction changes and there is an evolutionary attractor.

The main conclusion one can draw here is that selection of dispersal in the current model may depend crucially upon the details of the assumptions. Any simple conclusion is hard to find. Perhaps this reflects the real situation of dispersal.

Appendices

A. Lemma A.1

(i) For non-trivial $u \in C$,

$$(Eu, u) \leq 0 \quad (\text{A.1})$$

with strict equality if and only if both of the following hold:

$$u = \text{const} \quad (\text{A.2})$$

$$\int_{\Omega} \beta(x, y) dy = 1 \quad (x \in \Omega) \quad (\text{A.3})$$

- (ii) $(Eu)(x) = 0 \quad (\forall x \in \Omega)$ if and only if both equations (A.2) and (A.3) hold.
 (iii) $\int_{\Omega} (Eu)(x) dx \leq 0$ for non-trivial, non-negative u , with strict equality if equation (A.3) holds.

Proof. Since $\beta > 0$,

$$\int_{\Omega} \int_{\Omega} \beta(x, y) [u(x) - u(y)]^2 dx dy \geq 0, \quad (\text{A.4})$$

whence, from the symmetry of β ,

$$\begin{aligned} \int_{\Omega} \int_{\Omega} \beta(x, y) u(x) u(y) dx dy &\leq \int_{\Omega} u^2(x) dx \int_{\Omega} \beta(x, y) dy \\ &\leq \int_{\Omega} u^2(x) dx \end{aligned} \quad (\text{A.5})$$

from (H1). This proves that $(Eu, u) < 0$.

To deal with the equality case, first note that equations (A.2) and (A.3) imply that $(Eu, u) = 0$. To prove the converse, suppose that u is not a constant. Then the inequalities in the expression (A.4) and (A.5) become strict, from which it follows that $(Eu, u) < 0$. Therefore u is a constant. With $u(x) = 1$, $(Eu, u) = 0$ which implies that

$$\int_{\Omega} dx \int_{\Omega} \beta(x, y) dy = \int_{\Omega} dx,$$

which clearly cannot hold if $\int_{\Omega} \beta(x, y) dy < 1$ for some x . Equation (A.3) now follows. \square

B. Proofs

Throughout we take $\rho = 1$. It is clear that this may be done without loss of generality by re-scaling.

B.1. Proof of Theorem 3.1

Clearly H is self-adjoint and the first claim is standard. Define

$$\sigma_+ = \sup_{\|u\|_2=1} (Mu, u). \tag{B.1}$$

To complete the proof, we must first establish the following:

$$\max_{x \in \Omega} h(x) < \sigma_+. \tag{B.2}$$

Choose $x_0 \in \Omega$ such that $h(x_0) = \max_{x \in \Omega} h(x)$, and let $\Omega_\delta = \{x \in \Omega : |x - x_0| \leq \delta\}$. From the Lipschitz condition on h , there is a $C > 0$ such that

$$h(x_0) - h(x) \leq C|x - x_0| \quad (x \in \Omega). \tag{B.3}$$

From (H1) and the condition $\beta(x, x) > 0$ ($x \in \Omega$), there exist $\epsilon, \delta > 0$ such that

$$\beta(x, y) \geq \epsilon \quad (x, y \in \Omega_\delta). \tag{B.4}$$

To prove (B.2), it is enough to show that there is a $u \in L^2$ such that

$$h(x_0)\|u\|^2 - (Hu, u) < (Xu, u),$$

which is

$$\int_{\Omega} [h(x_0) - h(x)]u^2(x) dx < \int_{\Omega} \int_{\Omega} \beta(x, y)u(x)u(y) dx dy. \tag{B.5}$$

For some $\gamma > 0$ to be chosen later, define the trial function u as follows:

$$u(x) = \begin{cases} 1/(\gamma + h(x_0) - h(x)) & (x \in \Omega), \\ 0 & (x \notin \Omega). \end{cases}$$

From (B.3),

$$\begin{aligned} \int_{\Omega_\delta} \frac{dx}{\gamma + h(x_0) - h(x)} &\geq \int_{\Omega_\delta} \frac{dx}{\gamma + C|x - x_0|} \\ &\geq \frac{1}{C} \ln \left(\frac{C\delta + \gamma}{\gamma} \right) \\ &\geq \epsilon^{-1} \end{aligned} \tag{B.6}$$

for small enough γ . Then

$$\begin{aligned} \int_{\Omega} [h(x_0) - h(x)]u^2(x) dx &= \int_{\Omega_\delta} \frac{h(x_0) - h(x)}{[\gamma + h(x_0) - h(x)]^2} dx \\ &< \int_{\Omega_\delta} \frac{dx}{\gamma + h(x_0) - h(x)} \\ &\leq \epsilon \left\{ \int_{\Omega_\delta} \frac{dx}{\gamma + h(x_0) - h(x)} \right\}^2 \end{aligned} \tag{B.7}$$

from (B.6). Also, from (H1) and (B.4),

$$\epsilon \left\{ \int_{\Omega_\delta} \frac{dx}{\gamma + h(x_0) - h(x)} \right\}^2 \leq \int_{\Omega} \int_{\Omega} \beta(x, y)u(x)u(y) dx dy. \tag{B.8}$$

(B.7) and (B.8) yield (B.5) and hence (B.2).

Note next that $\sigma_+ \in \sigma(M)$ and $\sigma(M)$ lies to the left of σ_+ ([36] p.331, [18] p.167). It is standard that there is a sequence $\{u_n\}$ with $\|u_n\| = 1 (\forall n)$ such that

$$\lim_{n \rightarrow \infty} \|(M - \sigma_+ I)u_n\| = 0. \tag{B.9}$$

Define the operator $\tilde{H} : \mathcal{H} \rightarrow \mathcal{H}$ by setting

$$(\tilde{H}u)(x) = [\sigma_+ - h(x)]u(x).$$

From equation (B.2), \tilde{H} has a bounded inverse. Define next sequences $\{E_n\}, \{F_n\}$ in \mathcal{H} as follows:

$$\begin{aligned} E_n &= (M - \sigma_+ I)u_n \\ &= Xu_n - \tilde{H}u_n, \\ F_n &= \tilde{H}^{-1}Xu_n - u_n = \tilde{H}^{-1}E_n. \end{aligned}$$

It is an obvious consequence of equation (B.9) that $E_n \rightarrow 0$ and so $F_n \rightarrow 0$ as $n \rightarrow \infty$. Since X is compact, there is a subsequence, still denoted by $\{u_n\}$, such that $Xu_n \rightarrow v$, say; let $\phi = \tilde{H}^{-1}v$. Then $\lim_{n \rightarrow \infty} \tilde{H}^{-1}Xu_n = \tilde{H}^{-1}v = \phi$. It is easy to see that since $F_n \rightarrow 0, u_n \rightarrow \phi$. Hence

$$\tilde{H}^{-1}X\phi = \phi \implies X\phi = \tilde{H}\phi \implies M\phi = \sigma_+\phi. \tag{B.10}$$

Therefore $\sigma_+ \in \sigma_p(M)$, ϕ (assumed normalised) is an eigenfunction of M corresponding to the eigenvalue $\lambda_0 = \sigma_+$, and ϕ is continuous. Also $\lambda_0 = (M\phi, \phi)$. It is also clear from equation (3.2) that $\phi \geq 0$, for otherwise the choice of $|\phi|$ as test function would give a greater value for λ_0 . Strict positivity follows from the definition of X in equation (2.26) and the first inequality in (A.5). The uniqueness is then a simple consequence, for if ϕ_1, ϕ_2 were different eigenfunctions, $\phi_1 - \phi_2$ would be an eigenfunction. But this must change sign, contradicting the positivity. \square

B.2. Proof of Theorem 3.2

This proceeds in simple stages, the first being a routine observation.

- (i) It is enough to prove the result for the inequality

$$\frac{du}{dt} \geq Xu + c(x, t)u, \tag{B.11}$$

where c is strictly positive and bounded on $S_T \cup Q_T$. This follows from the substitution $u = v \exp(-\alpha t)$ and choice of α large enough.

The following notation will be useful:

$$c_M = \sup_{\substack{x \in \Omega \\ t \geq 0}} c(x, t),$$

$$U(t) = \min_{x \in \Omega} u(x, t),$$

$$U_-(t) = \min(U(t), 0).$$

Clearly

$$c(x, t)u(x, t) \geq c_M U_-(t) \quad (x \in \Omega, t \geq 0). \tag{B.12}$$

(ii) Claim: $u(x, t) \geq 0$ ($x \in \Omega, t \geq 0$). To prove this take $\delta = 1/[2(c_M + 1)]$. From (H2) and equations (B.11) and (B.12) it follows that

$$\frac{\partial u}{\partial t}(x, t) \geq (c_M + 1)U_-(t). \tag{B.13}$$

Define t_1 and then x_1 by taking

$$U_-(t_1) = \min_{0 \leq t \leq \delta} U_-(t),$$

$$u(x_1, t_1) = U_-(t_1).$$

Obviously $\int_0^{t_1} U_-(t) dt \geq t_1 U_-(t_1)$, and it follows on integrating equation (B.13) over $[0, t_1]$ and taking $x = x_1$ that

$$u(x_1, t_1) \geq u(x_1, 0) + (c_M + 1)t_1 U_-(t_1),$$

whence

$$U_-(t_1)[1 - (c_M + 1)t_1] \geq u(x_1, 0) \geq 0.$$

But $t_1 \leq \delta$, so this is impossible unless $U_-(t_1) = 0$. From the definition of t_1 , we must have $U_-(t) = 0$ for $0 \leq t \leq \delta$. A repetition of the argument proves the claim.

(iii) The result will obviously follow if we can establish the following claim: $u(x_0, t_0) = 0$ for some $x_0 \in \Omega, t_0 > 0$ implies that $u(x, t) = 0$ ($\forall x \in \Omega, (t \in [0, t_0])$). To prove this, it is enough to suppose that on the contrary $u(x_1, t_1) > 0$ for some $x_1 \in \Omega, t_1 \in [0, t_0)$, and use continuity for the case $t_1 = t_0$. Then from (H4), $(Xu)(x, t_1) > 0$ ($x \in \Omega$). From equation (B.11),

$$\frac{\partial u}{\partial t}(x, t_1) > 0 \quad (x \in \Omega).$$

Since this holds for every t_1 for which $u > 0$ for some $x \in \Omega$, we deduce that $u(x_0, t_0) > 0$, a contradiction. □

B.3. Proof of Lemma 4.1

First, consider $L_i(\tau)$ as operators $\mathcal{H} \rightarrow \mathcal{H}$. Let $\lambda_0(L_i(\tau))$ denote the PEV of $L_i(\tau)$ for $i = 1, 2$. We start by showing that $\sigma(L)$ is real and

$$\lambda_0(L_1(\tau')) < 0 \implies \sigma(L) \subset (-\infty, 0), \tag{B.14}$$

$$\lambda_0(L_1(\tau')) > 0 \implies \sigma(L) \cap (0, \infty) \neq \emptyset. \tag{B.15}$$

The fact that $\sigma(L)$ is real follows easily from L_1 and L_2 being self-adjoint. The proof is based on the simple observation, as in [4] Theorem 3.2, that the right hand end of $\sigma(L)$ is the larger, Λ say, of $\lambda_0(L_1(\tau'))$ and $\lambda_0(L_2(\tau))$. It is clear from equations (4.4) and (4.6) together with the variational characterization (3.2) that $\lambda_0(L_2(\tau)) < \lambda_0(L_1(\tau)) = 0$. Thus $\lambda_0(L_1(\tau')) < 0 \implies \Lambda < 0$. With $r(\cdot)$ denoting the resolvent set, we note that if $\lambda > \Lambda$ then $\lambda \in r(L_1(\tau')) \cap r(L_2(\tau))$. It is therefore clear that the following is the solution of $L(u, v) = (u_1, v_1)$ for any $(u_1, v_1) \in \mathcal{H} \times \mathcal{H}$:

$$\begin{aligned} v &= (L_1(\tau') - \lambda I)^{-1}v_1, \\ u &= (L_2(\tau) - \lambda I)^{-1}(u_1 + \tilde{u}v). \end{aligned}$$

This shows that $\lambda \in \rho(L)$ and proves (B.14).

On the other hand, if $\lambda = \lambda_0(L_1(\tau')) > 0$, and ϕ is the corresponding eigenfunction, then since $\lambda \in r(L_2(\tau))$ the pair

$$(\tilde{u}(L_2(\tau) - \lambda I)^{-1}\phi, \phi)$$

is an eigenfunction of L . Thus $\lambda \in \sigma(L)$, which establishes (B.15).

Consider next L as an operator on $C \times C$ into itself. We claim that if $\text{Re } \lambda > \Lambda$, then λ lies in the resolvent set of L . From the L^2 result, $(L - \lambda I)$ is injective, so it is enough to prove that it is surjective. Write $D(\tau) = \rho_\tau(X_\tau - I)$. Then from the definition of $\lambda_0(L_1(\tau))$, for the (positive) principal eigenfunction ψ ,

$$\rho_{\tau'}X_{\tau'}\psi + [a - \tilde{u} - \rho_{\tau'} - \lambda_0(L_1(\tau'))]\psi = 0.$$

Since $(X\psi)(x) > 0$ for all $x \in \Omega$, it follows that

$$a - \tilde{u} - \rho_{\tau'} - \lambda_0(L_1(\tau')) < 0 \quad (x \in \Omega).$$

By a similar argument

$$a - 2\tilde{u} - \rho_\tau - \lambda_0(L_2(\tau)) < 0 \quad (x \in \Omega).$$

We must show that for every $(w, z) \in C \times C$ the solution of the following equations is also in $C \times C$:

$$\rho_\tau X_\tau u + (a - 2\tilde{u} - \rho_\tau - \lambda)u - \tilde{u}v = w, \tag{B.16}$$

$$\rho_{\tau'} X_{\tau'} v + (a - \tilde{u} - \rho_{\tau'} - \lambda)v = z. \tag{B.17}$$

But $X_{\tau'}v \in C$ and from what was proved above,

$$a - \tilde{u} - \rho_{\tau'} - \text{Re } \lambda < 0.$$

Therefore, from equation (B.17), $v \in C$. From equation (B.16), a similar argument shows that u also belongs to C and this proves the claim.

To prove the asymptotic stability, note that by [14] Theorem 1.3.4, for some $c, \beta > 0$,

$$\|e^{Lt}\| \leq ce^{-\beta t} \quad (t \geq 0).$$

The asymptotic stability follows from a standard argument for ordinary differential equations on a Banach space, see [14] Theorem 5.1.1, for example. \square

B.4. Proof of Lemma 5.2 (Outline only)

A complete and detailed proof of this result is lengthy and it is not appropriate to include it here. In outline, however, it follows a PDE-type argument and is straightforward. First, a maximum principle holds for the scalar equation (see Section 3) and this extends in the usual way to a competitive system, see for example [23] chapter 1 or [15]. As proved in Section 5, there is no coexistence solution. Also, from the scalar comparison theorem, the region $0 < u < \tilde{u}$ and $0 < v < \tilde{v}$ is positively invariant.

Let $z = (z_1, z_2)$ and $w = (w_1, w_2)$. Write $w \leq z$ if $w_1 \geq z_1$ and $w_2 \leq z_2$. For the system (5.3), (5.4) expressed in the form $dw/dt = F(x, w)$, suppose that for some stationary \underline{w} in the invariant set above, $F(x, \underline{w}) \leq 0$. If \underline{w} is not a solution then the solution of the system with $w(x, 0) = \underline{w}(x)$ is such that $w_1(x, \cdot)$ and $w_2(x, \cdot)$ are strictly increasing and decreasing respectively. Since they are bounded, they must each tend to a limit, which must be a solution. Since there is no coexistence solution, this has to be $(\tilde{u}, 0)$. It is thus enough to find such a \underline{w} .

From the linearization about $(0, \tilde{u})$, the eigenvalue equations are

$$\begin{aligned} \mu D\phi + (a - \tilde{v})\phi &= \lambda\phi, \\ \nu D\psi + (a - 2\tilde{v})\psi - \tilde{v}\phi &= \lambda\psi, \end{aligned}$$

with $\lambda > 0, \phi > 0$. Then

$$\nu D\psi + (a - 2\tilde{v} - \lambda)\psi = \tilde{v}\phi.$$

However, the PEV of the operator is negative. Further, an analogue of the positivity result of [15] Theorem 16.6(i) clearly holds, and we conclude that $\psi < 0$.

Now let $\underline{w} = (\delta\phi, \tilde{v} + \delta\psi)$ for some small $\delta > 0$. We need only to check that $F(x, \underline{w}) \leq 0$. For the first component, we have

$$\mu D\phi + (a - \tilde{v} - \delta\phi - \delta\psi)\phi = \lambda\phi - \delta(\phi + \psi)\phi > 0$$

for small δ . For the second component, we have

$$\begin{aligned} \nu D(\tilde{v} + \delta\psi) + (a - \delta\phi - \tilde{v} - \delta\psi)(\tilde{v} + \delta\psi) \\ = \delta[\mu D\psi + (a - 2\tilde{v})\psi - \tilde{v}\psi] - \delta^2\psi(\phi + \psi) \\ = \delta\lambda\psi - \delta^2\psi(\phi + \psi) \\ < 0 \end{aligned}$$

since $\lambda > 0$ and $\psi < 0$. This completes the proof. \square

References

1. Chmaj, A., Ren, X.: Homoclinic solutions of an integral equation: existence and stability. *J. Diff. Eqns.* **155**, 17–43 (1999)
2. Dieckmann, U., Law, R., Metz, J. *The Geometry of Ecological Interactions: Simplifying Spatial Complexity*. Cambridge University Press, Cambridge, 2000
3. Dieckmann, U., O'Hara, B., Weisser, W.: The evolutionary ecology of dispersal. *Trends Ecol. Evol.* **14**, 88–90 (1999)
4. Dockery, J., Hutson, V., Mischaikow, K., Pernarowski, M.: The evolution of slow dispersal rates. *J. Math. Biol.* **37**, 61–83 (1998)
5. Doebli, M., Ruxton, G.: Evolution of dispersal rates in metapopulation models: branching and cyclic dynamics in phenotype space. *Evolution* **51**, 1730–1741 (1997)
6. Duncan, D., Grinfeld, M., Stoleriu, I.: Coarsening in an integro-differential model of phase transitions. *Euro. J. Appl. Math.* **11**, 511–572 (2000)
7. Feduccia, A.: *The Origin and Evolution of Birds*. Yale University Press, New Haven, 1996
8. Ferriere, R., Belthoff, J., Olivieri, I., Krackow, S.: Evolving dispersal: where to go next? *Trends Ecol. Evol.* **15**, 5–7 (2000)
9. Fife, P.: *Mathematical Aspects of Reaction-Diffusion Equations*, Lecture Notes in Mathematics. **28**, Springer-Verlag, New York, 1979
10. Fife, P.: An integro-differential analog of semilinear parabolic PDE's. In: *Partial Differential Equations and Applications*, eds. P. Marcellini, G. Talenti and E. Vesenti, Lecture Notes in Pure and Applied Mathematics 1996, **177**, Marcel Dekker.
11. Fife, P.: Clines and material interfaces with non-local interaction. In: *Nonlinear Problems in Applied Mathematics*, eds. T.S. Angell, L. Pamela Cook, R.E. Kleinman and W.E. Olmstead 1996, 134–149, SIAM Publications.
12. Hadeler, K.P.: Reaction-transport systems in biological modelling. In: *Mathematics Inspired by Biology*, eds. V. Capasso and O. Diekmann, Lecture Notes Mathematics 1999, **1714**, Springer-Verlag, New York
13. Hastings, A.: Can spatial variation alone lead to selection for dispersal? *Theor. Pop. Biol.* **24**, 244–251 (1983)
14. Henry, D.: *Geometric Theory of Semilinear Parabolic Equations*. Springer-Verlag, Berlin, 1981
15. Hess, P.: *Periodic-Parabolic Boundary Value Problems and Positivity*. Pitman Research Notes in Mathematics, **247**. Longman, New York, 1991
16. Hofbauer, J., Sigmund, K.: Adaptive dynamics and evolutionary stability. *Appl. Math. Lett.* **3**, 75–79 (1990)
17. Holt, R., McPeck, M.: Chaotic population dynamics favours the evolution of dispersal. *Amer. Nat.* **148**, 709–718 (1996)
18. Hutson, V., Pym, J.S.: *Applications of functional analysis and operator theory*. Academic Press, London, 1980
19. Hutson, V., Mischaikow, K., Poláčik, P.: The evolution of dispersal rates in a heterogeneous, time-periodic environment. *J. Math. Biol.* **42**, 501–533 (2001)
20. Johnson, M., Gaines, M.: The evolution of dispersal: theoretical models and empirical tests using birds and mammals. *Ann. Rev. Ecol. Syst.* **21**, 449–480 (1990)
21. Kot, M., Schaffer, W.: Discrete-time growth-dispersal models. *Math. Biosc.* **80**, 109–136 (1986)
22. Kot, M., Lewis, M., van den Driessche, P.: Dispersal data and the spread of invading organisms. *Ecology* **77**, 2027–2042 (1996)
23. Leung, A.W.: *Systems of partial differential equations*. Kluwer, Dordrecht, 1989

24. McPeck, M., Holt, R.: The evolution of dispersal in spatial and temporally varying environments. *Amer. Nat.* **140**, 1010–1027 (1992)
25. Méndez, V., Pujol, T., Fort, J.: Dispersal probability distributions and the wave-front speed problem. *Phys. Rev. E* **65**, 1–6 (2002)
26. Metz, J.A.J., Geritz, S.A.H., Meszéna, G., Jacobs, F.J.A., van Heerwaarden, J.S.: Adaptive dynamics: A geometrical study of the consequences of nearly faithful reproduction. In: S.J. van Strien, S.M. Verduyn Lunel (eds), *Stochastic and Spatial Structures of Dynamical Systems*, KNAW Verhandelingen afd. Natuurkunde Eerste Reeks deel 45. North Holland, Amsterdam, Netherlands 1996, pp. 183–231
27. Moody, M.: The evolution of migration in subdivided populations I. Haploids. *J. Theor. Biol.* **131**, 1–13 (1988)
28. Moody, M., Ueda, M.: A diffusion model for the evolution of migration. Preprint.
29. Murray, J.D.: *Mathematical Biology*, Springer-Verlag, New York, 1989
30. Nagylaki, T.: The diffusion model for migration and selection. In: *Some Mathematical Questions in Biology: Models in Population Biology* ed. A. Hastings, *Lectures on Mathematics in the Life Sciences*, **20**, 1989
31. Neuhauser, N.: Mathematical challenges in spatial ecology. *Notices of the AMS* 1304–1314 (2001)
32. Othmer, H.G., Dunbar, S.R., Alt, W.: Models of dispersal in biological systems. *J. Math. Biol.* **26**, 263–298 (1988)
33. Pao, C.V.: *Nonlinear Parabolic and Elliptic Equations*. Plenum Press, New York, 1992
34. Roughgarden, J.: *Theory of Population Genetics and Evolutionary Ecology: An Introduction*. Macmillan, New York, 1979
35. Smith, H.L.: *Monotone dynamical systems*. Amer. Math. Soc., Providence, R.I. 1995
36. Taylor, A.E.: *Introduction to functional analysis*. Wiley, London, 1958
37. Travis, J.M.J., Dytham, C.: Habitat persistence, habitat availability and the evolution of dispersal. *Proc. Roy. Soc. Lon. B* **266**, 723–728 (1999)
38. Wang, X.: Metastability and patterns in a convolution model for phase transitions. *J. Diff. Eqns.* **183**, 434–461 (2002)
39. Weinberger, H.: Asymptotic behaviour of a model in population genetics. In: Chadan, J. (ed.) *Nonlinear PDE's and Applications*, Springer Lecture Notes **648**, 47–96 (1970)