New Journal of Physics
The open–access journal for physics

Time-delayed fronts from biased random walks

Joaquim Fort[1] **and Toni Pujol**[2]

[1] Departament de Física, Campus de Montilivi, Universitat de Girona, 17071 Girona, Catalonia, Spain
[2] Departament de Mecànica, Campus de Montilivi, Universitat de Girona, 17071 Girona, Catalonia, Spain
E-mail: joaquim.fort@udg.es

New Journal of Physics **9** (2007) 234
Received 19 April 2007
Published 17 July 2007
Online at http://www.njp.org/
doi:10.1088/1367-2630/9/7/234

Abstract. We generalize a previous model of time-delayed reaction–diffusion fronts (Fort and Méndez 1999 *Phys. Rev. Lett.* **82** 867) to allow for a bias in the microscopic random walk of particles or individuals. We also present a second model which takes the time order of events (diffusion and reproduction) into account. As an example, we apply them to the human invasion front across the USA in the 19th century. The corrections relative to the previous model are substantial. Our results are relevant to physical and biological systems with anisotropic fronts, including particle diffusion in disordered lattices, population invasions, the spread of epidemics, etc.

Contents

1. Introduction

Fronts are solutions to reaction–diffusion equations [1, 2]. They describe propagating profiles for the particle concentration, individual number density, temperature, etc. They are used to describe combustion flames [3]–[6], population invasions [7]–[9], virus infections [10], and many other interesting phenomena in physical and biophysical systems [1, 2].

Anisotropic reaction–diffusion fronts have been analysed by several authors. Previous theoretical results include a Hamilton–Jacobi derivation of the front position [11], a propagation failure condition for random walks biased in the opposite direction to that of the front propagation [12], velocity-curvature relations [13], nucleation of spiral waves [14], etc. An interesting application is the recent explanation (via computer simulations) of the nonhomogeneous speed of Neolithic fronts, based on anisotropic diffusion due to enhanced transport along major rivers [15].

In this paper, we will follow a microscopic approach to deal with systems in which particles (or individuals) move with a direction-dependent probability, i.e. following a biased (or anisotropic) random walk. Such a microscopic behaviour has been considered for many systems, such as particle diffusion in disordered lattices (see section 10 in [16]) diffusion-limited aggregation [17], experimental populations of micro-organisms [18], human populations invading a geographical region [15], etc.

In this paper, we will first develop two models leading to anisotropic reaction–diffusion, and derive their front speeds (sections 2 and 3). Then we shall present an application of these results (section 4), and finally include some concluding remarks (section 5).

Correlation between the directions of successive jumps will not be included in our models. The reason is that there is no reason to think that such a correlation is relevant in the application we will tackle (section 4) and many similar situations. Therefore, we will deal with biased, *uncorrelated* random walks. This is a fundamental difference between our models and those

on correlated (or persistent) random walks. In our models, we will allow the probability of jump to depend on the angle relative to a fixed direction. In contrast, in correlated models the probability of jump depends on the angle relative to the direction of motion before performing the jump [19]. Models incorporating both effects have also been proposed [20]. Our framework is similar to other phenomena (not considered in the present paper) such as bacterial chemotaxis, where the probability of jump also depends on the angle relative to an external direction (e.g. that of the chemoattractant gradient). However, the way in which detection of an external gradient or stimulus by a cell leads to directed movement requires modelling of intracellular mechanisms [21] that are beyond population-dynamics models such as that in the present paper. Another fundamental difference between the present paper and those references is that, we will include population reproduction in addition to diffusion.

Again, in order for our models to be useful in the application we want to discuss, we will deal with a two-dimensional (2D) space throughout this work.

2. First model: non-sequential fronts from biased random walks

2.1. Microscopic derivation

In this section, we generalize the framework in [9] to the case of biased random walks. Let $p(x, y, t)$ stand for the population (or particle) number per unit area at position (x, y) and time t.

In the present paper, we define the dispersal kernel $\phi(\triangle_x, \triangle_y)$ as the probability per unit area that an individual (or particle) who was at $(x-\triangle_x, y-\triangle_y, t)$ jumps to $(x, y, t+T)$.[3]

Let T stand for the mean time interval between two subsequent jumps (in biophysical applications, usually $T = 1$ generation [9, 10]). Let $R[p(x, y, t)]$ stand for the number of new individuals (or particles) due to the reproduction process (or chemical reactions), produced during the time interval T per unit area centred at (x, y). From these definitions, the evolution equation is typically written down as follows [9]

$$p(x, y, t+T) - p(x, y, t) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x-\triangle_x, y-\triangle_y, t)\,\phi(\triangle_x, \triangle_y)\,\mathrm{d}\triangle_x\,\mathrm{d}\triangle_y - p(x, y, t) + R[p(x, y, t)], \tag{1}$$

where the first and second terms in the right-hand side correspond to population dispersal, and the last one $R[p(x, y, t)]$ is a source term due to net reproduction (or chemical reactions). In general, a sink term can also be added to the former equation but here it is not included explicitly because, in the next subsection, we shall consider biological populations and then, $R[p(x, y, t)]$ will describe the net effect of both sources and sinks (i.e. of the reproduction and death of individuals, respectively).

Usually second-order Taylor expansions in space ('diffusion') and time are performed. Then equation (1) becomes

$$\frac{\partial p}{\partial t} + \frac{T}{2}\frac{\partial^2 p}{\partial t^2} = -U_x\frac{\partial p}{\partial x} - U_y\frac{\partial p}{\partial y} - U_{xy}\frac{\partial^2 p}{\partial x\,\partial y} + D_x\frac{\partial^2 p}{\partial x^2} + D_y\frac{\partial^2 p}{\partial y^2} + F + \frac{T}{2}\frac{\partial F}{\partial t}. \tag{2}$$

[3] Note that in [9], we defined $\phi(\triangle_x, \triangle_y)$ as the probability of a jump $(x+\triangle_x, y+\triangle_y, t) \to (x, y, t+T)$. Here we use negative signs instead, because otherwise the kernel $\Phi(\theta)$ in section 4 would correspond to a jump with angle $-\theta$, which would be rather confusing.

Here F is the time derivative of $p(x, y, t)$ due to reproduction, i.e. [9]

$$R[p(x, y, t)] = TF + \frac{T^2}{2!}\frac{\partial F}{\partial t} + \frac{T^3}{3!}\frac{\partial^2 F}{\partial t^2} + \cdots, \tag{3}$$

D_x and D_y are direction-dependent diffusion coefficients

$$D_x(x, y) = \frac{\langle \triangle_x^2 \rangle}{2T}, \tag{4}$$

$$D_y(x, y) = \frac{\langle \triangle_y^2 \rangle}{2T}, \tag{5}$$

and we have defined

$$U_x(x, y) = \frac{\langle \triangle_x \rangle}{T}, \tag{6}$$

$$U_y(x, y) = \frac{\langle \triangle_y \rangle}{T}, \tag{7}$$

$$U_{xy}(x, y) = \frac{\langle \triangle_x \, \triangle_y \rangle}{T}, \tag{8}$$

where the mean value of an arbitrary function $\zeta(\triangle_x, \triangle_y)$ is defined as

$$\langle \zeta(\triangle_x, \triangle_y) \rangle \equiv \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \zeta(\triangle_x, \triangle_y)\, \phi(\triangle_x, \triangle_y, x, y)\, \mathrm{d}\triangle_x\, \mathrm{d}\triangle_y.$$

In general, the dispersion kernel $\phi(\triangle_x, \triangle_y, x, y)$ can depend on position (x, y) in addition to the jump vector components (Δ_x, Δ_y). Then, the macroscopic parameters above (D_x, D_y, etc) also depend on position. However, for our purposes here it will be sufficient to consider mainly the homogeneous case. Then, we can simply write $\phi(\triangle_x, \triangle_y)$ instead of $\phi(\triangle_x, \triangle_y, x, y)$.

The first three terms in the right-hand side of equation (2), which correspond to equations (6)–(8), did not appear in [9] because there we assumed that the kernel $\phi(\triangle_x, \triangle_y)$ was isotropic (thus $D_x = D_y \equiv D$).

An interesting 1D anisotropic equation similar to (2) was considered by Fedotov [11], but the diffusion terms had a different form. Also, in [12] a general continuous-time anisotropic reaction–dispersion equation was analysed in 1D. Here, we consider the 2D reaction–diffusion equation (2) because its microscopic derivation above will make it possible to tackle a specific application in section 4 by estimating the values of the parameters by means of the 2D equations (4)–(8).

A macroscopic derivation of this model is included in the appendix. Such a macroscopic derivation may be suitable for other applications, but not for the one we will discuss in the present paper (section 4) because there the macroscopic parameters (D_x, U_x, etc) cannot be directly measured. The macroscopic derivation (appendix) does not relate them to the dispersion kernel $\phi(\triangle_x, \triangle_y)$. The latter can be measured, so we will need the microscopic derivation above in order to obtain numerical estimations of the macroscopic parameter values.

2.2. The speed of front solutions

As in [9], let us consider a reproduction function that has been widely applied in biophysics problems, namely the logistic function

$$F = ap(x, y, t)\left(1 - \frac{p(x, y, t)}{p_{\max}}\right), \tag{9}$$

where a is called the initial growth rate and $p_{\max}$ the saturation density. This function includes the net effect of reproduction and death of individuals [22].

In contrast to some previous work on non-delayed non-isotropic fronts [13, 14, 23], we will not tackle the problem of the dependence of front speed on the direction. Instead, in order to derive results relevant to the application we are interested in (section 4), we need to focus our attention into the speed of time-delayed fronts along the x-direction.

We recall that the front speed can be found most easily by assuming that for $t \to \infty$ the front curvature is negligible at scales much larger than that of individual dispersal events. This approach that has been successful in previous work [1, 24]. In other words, we consider a region centred about the x-axis which is sufficiently small so that the y-dependence of $p(x, y, t)$ can be neglected. Let c stand for the front speed. We look for constant-shape solutions with the form

$$p = p_0 \exp[-\lambda(x - ct)] \tag{10}$$

as $x - ct \to \infty$, with $c > 0$ and $\lambda > 0$. In this way, from equations (2) and (9) we obtain the characteristic equation

$$\lambda^2\left(D_x - \frac{Tc^2}{2}\right) - \lambda\left(c - U_x - \frac{aTc}{2}\right) + a = 0. \tag{11}$$

Solving this equation for λ and requiring for it to be real, we obtain the condition

$$f(c) \equiv c^2\left(1 + \frac{aT}{2}\right) - 2cU_x\left(1 - \frac{aT}{2}\right) - 4aD_x - U_x^2 \geqslant 0. \tag{12}$$

It is easily seen that $f(c)$ is convex from below, and that the equation $f(c) = 0$ has one negative and one positive root for c, say c_- and c_+. Therefore, the minimum possible value for $c > 0$ corresponds to c_+. Let us now assume, as usual, that this minimum possible speed c_+ is that selected by the front (this is usually called linear or marginal stability analysis [1, 2]). In this way, we finally obtain

$$c = \frac{U_x(1 - (aT/2)) + 2\sqrt{aD_x(1 + (aT/2))^2 - (aT/2)U_x^2}}{(1 + (aT/2))^2}. \tag{13}$$

We note that

$$\lim_{U_x \to 0} c = \frac{2\sqrt{aD}}{1 + aT/2}, \tag{14}$$

where $D \equiv D_x$, so that we recover, as a special case, the result derived in [9] for non-biased random walks.

For a biased random walk with negligible delay time ($T \ll \frac{1}{a}$), equation (13) becomes $c = U_x + 2\sqrt{aD_x}$. If we consider the additional limit of a non-biased random walk (equation (14)), we obtain $c \to 2\sqrt{aD}$, which is Fisher's well-known speed [25].

For a macroscopic derivation and a discussion about this model, see the appendix.

3. Second model: sequential fronts from biased random walks

3.1. Derivation of the second model

Equation (1) has been widely applied [1, 9, 10, 24], and it is a reasonable starting point. However, it assumes simultaneous dispersal (first and second terms in the right-hand side) and reproduction (last term). But for application to biological populations (section 4), a model with non-simultaneous reproduction and dispersal seems more realistic. A simple way to see this is that, according to the last term in equation (1), the population at (x, y, t) reproduces causing a population number increase at the same space point (x, y) one generation later $(t + T)$. But, in fact, that parent population will no longer be at (x, y) at time $t + T$ because of dispersal [first and second terms in the rhs equation (1)]. So we can say that in the first model 'parents leave their children behind'. This is not realistic for most biological species, e.g. humans, because children need to spend some time depending on their parents (until they become adults or independent). Therefore, it is more realistic to replace equation (1) by

$$p(x, y, t + T) = R\left[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x - \triangle_x, y - \triangle_y, t)\phi(\triangle_x, \triangle_y)\,\mathrm{d}\triangle_x\,\mathrm{d}\triangle_y\right], \tag{15}$$

so that after the population moves (from $(x+\triangle_x, y+\triangle_y)$ into (x, y)), it reproduces at the arrival location (x, y) according to a reproduction function $R[p(x, y, t)]$. In this sense, equation (15) is a sequential (or time-ordered) evolution equation, whereas equation (1) is not.

Concerning the reproduction function $R[p(x, y, t)]$, we cannot use a logistic form for it (i.e. we cannot set $R[p] = ap(1 - p/p_{\max})$). The reason is that it is known from non-spatial models [25] that it yields negative values for the particle (or population) number density $p(x, y, t)$ for discrete-time equations such as (15), which makes no physical sense. Therefore, we will simply assume net reproduction to be proportional to the population density but bounded by a maximum value, $p_{\max}$,

$$R[p(x, y, t)] = \begin{cases} R_0 p(x, y, t) & \text{if } p < p_{\max}, \\ p_{\max} & \text{if } p \geqslant p_{\max}, \end{cases} \tag{16}$$

where R_0 is called the net reproductive rate (or fecundity) per generation.

In section 3.3, we will show that equation (16) and the logistic equation (9) differ only at high values of $p(x, y, t)$ (and thus, they would give the same front speed, if both were applied to this second model).

Equations similar to (15) have been previously considered and applied to predict front speeds, but mainly in 1D space and only for non-biased kernels $\phi(\triangle_x, \triangle_y)$ [26, 27]. In contrast, here we are dealing with the 2D case, and also consider biased kernels (both points are necessary for the applications we shall present in section 4).

In sequential (i.e. time-ordered) models, it is always assumed that $R_0 > 1$ [26]. Otherwise, the number density of particles (or individuals) $p(x, y, t)$ would decrease according to equation (16), i.e. $R_0 < 1$ would correspond to a negative net population reproduction, which cannot lead to invasion fronts.

As for the first model, we shall apply the linearization method, i.e. consider the leading edge of the invasion front

$$z \equiv x - ct \to \infty.$$

Therefore, we can assume low values for the population density $p(x, y, t)$ and, using equation (16), we can write equation (15) as

$$p(x, y, t+T) = R_0 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} p(x+\triangle_x, y+\triangle_y, t)\phi(\triangle_x, \triangle_y)\, \mathrm{d}\triangle_x\, \mathrm{d}\triangle_y. \tag{17}$$

Performing Taylor expansions up to second order in space and time, this equation becomes

$$\tfrac{1-R_0}{T} p + p_t + \tfrac{T}{2} p_{tt} = R_0 \left(-U_x p_x - U_y p_y - U_{xy} P_{xy} + D_x p_{xx} + D_y p_{yy}\right), \tag{18}$$

where U_x, U_y, D_x, D and U_{xy} are given by equations (4)–(8).

Equation (18) is the sequenced (or time-ordered) analogue to equation (2). As explained at the beginning of this section, for biophysical applications the second model (equation (18)) is more realistic than the first one (equation (2)).

Note that equation (17) is time-sequential in spite of the fact of having used the linear net reproduction rate (16) (instead of the logistic (9), which is nonlinear). Indeed, mathematically the non-sequential character of equation (1) does not correspond to a nonlinearity in the reproduction rate. It corresponds to the sum of two terms in equation (1), corresponding to independent processes (dispersion and reproduction). In equation (17), instead of such a sum we have a composite function: reproduction is applied to the result of dispersion, so the model is time-sequenced (albeit linear).

3.2. The speed of front solutions

As in section 2.2, we look for solutions with the form $p = p_0 \exp[-\lambda(x - ct)]$ as $x - ct \to \infty$, with $c > 0$ and $\lambda > 0$. Then equation (18) yields the characteristic equation

$$\lambda^2 \left(D_x R_0 - \frac{Tc^2}{2}\right) + \lambda\left(-c + U_x R_0\right) + \frac{R_0 - 1}{T} = 0. \tag{19}$$

Solving this equation for λ and requiring for it to be real, we obtain the condition

$$g(c) \equiv c^2\left(1 + 2\left(R_0 - 1\right)\right) - 2cU_x R_0 + U_x^2 R_0^2 - 4R_0 \frac{R_0 - 1}{T} D_x \geqslant 0. \tag{20}$$

Again, it is easily seen that $g(c)$ is convex from below, and that the equation $g(c) = 0$ has one negative and one positive root for c, say c_- and c_+. Therefore, the minimum possible value for $c > 0$ corresponds to c_+, and we finally obtain the speed

$$c = \frac{R_0 U_x + \sqrt{R_0\left(R_0 - 1\right)\left[(4/T)\left(2R_0 - 1\right) D_x - 2R_0 U_x\right]}}{\left(2R_0 - 1\right)}. \tag{21}$$

For the special case of a non-biased random walk, this becomes

$$\lim_{U_x \to 0} c = \sqrt{\frac{4R_0 D}{T} \frac{R_0 - 1}{2R_0 - 1}}, \tag{22}$$

where we have introduced $D \equiv D_x$.

3.3. Comparison to the first model

In order to compare the results of both models in section 4, it will be necessary to establish the connection between the low-density population growth parameters a (appearing in the first model, e.g. equation (9)) and R_0 (which appears in the second model, e.g. equation (16)). This relationship can be obtained most easily as follows. In the absence of dispersal, both equations (1) and (15) become

$$p(x, y, t+T) - p(x, y, t) = R[p(x, y, t)] \tag{23}$$

which, combined with equation (3), implies that

$$\frac{\partial p}{\partial t} = F. \tag{24}$$

Thus, for low values of $p(x, y, t)$, the logistic form (9) yields

$$p(x, y, t+T) = p(x, y, t)\exp[aT], \tag{25}$$

whereas equation (16) yields

$$p(x, y, t+T) = R_0 p(x, y, t), \tag{26}$$

so that the reproduction function (16) and the logistic(9) are consistent with each other at low values of the population density, provided that

$$a = \frac{1}{T}\ln R_0. \tag{27}$$

On the other hand, the reproduction function (16) and the logistic (9) will give different results for high values of the population density $p(x, y, t)$. However, the high-density behaviour is not accurately known for biological populations outside the laboratory, because there are no experimentally well-established trends in the population numbers versus time (except at low population densities) [28]. Moreover, it has been observed in non-biased models that the high-p behaviour of reproduction does not affect the speed of fronts for the two cases considered the present paper, i.e. for (i) the logistic growth function in non-sequential evolution equations (first model) [9] and (ii) integro-difference equations with a linear source term (second model) [26]. However, the high-p behaviour may be important in other cases, which are out of the scope of the present paper (e.g. the Arrhenius source function, widely used in combustion science [5]).

Although a comparison to the first model does not seem possible for an arbitrary bias U_x, it is possible in the non-biased limit ($U_x = 0$). For this purpose, using equation (27) into (22) it is easily seen that the speed from the second model (22) will be higher than that from the first model, equation (14), provided that

$$\exp[\tilde{T}](\exp[\tilde{T}] - 1)(1 + \tilde{T}/2)^2 - \tilde{T}(2\exp[\tilde{T}] - 1) > 0,$$

where $\tilde{T} = Ta > 0$. Plotting the left-hand side for $\tilde{T} > 0$, it is easily seen that this condition is always fulfilled. The physical interpretation is that the first model corresponds to simultaneous dispersal and reproduction, which means that 'parents leave their children behind' (see the beginning of this section). Intuitively, this should lead to slower invasion fronts. This is the physical interpretation of the fact that the second model leads to faster front speeds.

In the previous paragraph, we have considered arbitrary values of $\tilde{T} = Ta > 0$. But we may note that, in fact, it is sufficient to consider low values of $\tilde{T}$. The reason is that the validity

of the second-order Taylor expansion (2) implies that $\frac{aT}{2} < 1$, as it is clear from the fact that equation (14) must then involve a small correction to Fisher's speed, $2\sqrt{aD}$. This condition $\frac{aT}{2} < 1$ implies that the characteristic dispersion time T is sufficiently small (as compared to $\frac{2}{a}$, which may be considered a characteristic reproduction time) and is satisfied for many biological populations, e.g. humans [9].

4. Application

4.1. Connection between microscopic and macroscopic parameters

In order to apply our models, we need to assume some function for the kernel $\phi(\triangle_x, \triangle_y)$ appearing in equations (4)–(8). Many forms for the kernel have been considered in the literature. Unless there is some sort of evidence for a correlation between the lengths and the directions of the jumps, it is assumed that they are independent from each other. For this reason, it is usual to assume that

$$\phi(\triangle_x, \triangle_y) = \Psi(\triangle)\Phi(\theta), \tag{28}$$

where $\triangle = \sqrt{\triangle_x^2 + \triangle_y^2}$ and $\theta = \tan^{-1}\frac{\triangle_y}{\triangle_x}$. Correspondingly, we write the normalization condition of the kernel, namely

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \phi(\triangle_x, \triangle_y)\,\mathrm{d}\triangle_x\,\mathrm{d}\triangle_y = 1, \tag{29}$$

as a normalization condition for the length jump probability distribution,

$$\int_0^{\infty} \Psi(\triangle)\,\triangle\,\mathrm{d}\triangle = 1, \tag{30}$$

and another one for the probability distribution of the jump direction,

$$\int_0^{2\pi} \Phi(\theta)\,\mathrm{d}\theta = 1. \tag{31}$$

Several functions $\Phi(\theta)$ have been used in the literature on biased random walks [18, 29]. In order to illustrate the use of our models, it will be enough to consider the simple form

$$\Phi(\theta) = a \pm b\cos\theta, \tag{32}$$

where $b \geqslant 0$ and $a = \frac{1}{2\pi}$ from the normalization condition (31). Therefore

$$\Phi(\theta) = \frac{1}{2\pi} \pm b\cos\theta. \tag{33}$$

The diffusion coefficient D_x and the macroscopic bias parameter U_x appearing in the front speed (equation (13) in the first model; equation (21) in the second) can now be related to the microscopic bias parameter b, by performing the integrations in equations (4) and (6). This yields

$$D_x = \frac{\langle\triangle_x^2\rangle}{2T} = \frac{\langle\triangle^2\rangle}{4T} = \frac{1}{4T}\int_0^{\infty} \Psi(\triangle)\,\triangle^3\,\mathrm{d}\triangle, \tag{34}$$

$$U_x = \frac{\langle\triangle_x\rangle}{2T} = \pm\pi b\frac{\langle\triangle\rangle}{T} = \pm\frac{\pi b}{T}\int_0^{\infty} \Psi(\triangle)\,\triangle^2\,\mathrm{d}\triangle. \tag{35}$$

The following two cases can be considered.

(i) The positive sign in equations (33) and (35) corresponds to the case in which the random walk is biased towards the front propagation direction (recall that in sections 2 and 3, we have computed macroscopic front speeds along the x-direction ($\theta = 0$)). Then, the jump probability along the front direction ($\theta = 0$) is $\Phi = a + b$. It decreases with increasing values of $|\theta|$, down to the minimum $\Phi = a - b$ (which is attained for $\theta = \pi$).

(ii) The negative sign in equations (33) and (35) corresponds to the case in which the minimum probability is attained along the front direction, namely $\Phi(\theta = 0) = a - b$. It increases for increasing values of $|\theta|$, up to the maximum possible value $\Phi(\theta = \pi) = a + b$. Note that the kernel (33) is a probability distribution, so it must be positive for all the values of θ. Thus, in case (ii) we have the condition

$$0 \leqslant b \leqslant \frac{1}{2\pi}. \tag{36}$$

In both cases (i) and (ii), the dimensionless parameter

$$\beta \equiv \frac{b}{a} = 2\pi b \geqslant 0, \tag{37}$$

may be called the bias of the random walk. In case (ii), we see from equation (36) that

$$0 \leqslant \beta \leqslant 1. \tag{38}$$

Note that we may have case (i) at one point of space and case (ii) in another point because U_x, as defined by equation (6), is space-dependent. This may be interesting to describe systems with nonhomogeneous rates of front spread. For example, in biological invasions individuals may have a preference to jump in the front direction at some areas (case (i), $U_x > 0$), e.g. because they are attracted by more favourable habitats. But if other regions are difficult to colonize, the random walk of individuals may be strongly biased against the front invasion direction (case (ii), $U_x < 0$) and the front speed will become slower. An application of case (ii) ($U_x < 0$) is presented below.

4.2. Application: human invasion of America during the 19th century

The speed of the human population front colonizing North America in the period 1790–1910 can be easily determined, either from detailed population maps [30] or from the centre-of-mass population trajectory [31]. Both approaches yield essentially the same range for the observed speed, namely $(13.5 \pm 0.8)\,\mathrm{km\,yr^{-1}}$ (95% confidence-level interval) [32]. On the other hand, mean migration data of individuals are strongly biased in the direction opposite to that of the front propagation [33]. Therefore, we are dealing with case (ii) in the former subsection.

The parameter values for this application can be estimated as follows. Lotka fitted a logistic growth function (9) to the population of the USA and obtained for the initial growth rate $a = 0.031\,\mathrm{yr^{-1}}$ [22]. It is worth noting that this estimation agrees almost exactly with independent estimations for human populations in other places and time intervals [9]. Diffusion parameters are more difficult to estimate. Sometimes a relatively small sample of migration distances from genealogies are combined with persistence data from other sources [32], but demographers have pointed out that genealogy data are not representative of the whole population [33]. Ferrie has analysed migration distances for the USA in the 19th century [33]. Using his data for regions with more than 500 observations (i.e. a total of 3804 individuals) yields $D_x = 6075\,\mathrm{km^2\,yr^{-1}}$ using equation (34). Finally, we can estimate the macroscopic bias

parameter U_x using equation (35) and values for $\frac{\langle\Delta\rangle}{T} = 24.42\,\mathrm{km\,yr^{-1}}$ and b (or, equivalently, β),

$$U_x = -\frac{\beta}{2}\frac{\langle\Delta\rangle}{T}. \tag{39}$$

The values of β and b can also be estimated from Ferrie's data cited above[4]. However, we prefer to use the anisotropy parameter β as a free parameter (horizontal axis in figure 1) because Ferrie's data contain only a few directions, so it does not seem possible to obtain a precise value for β.[5]

In figure 1 we present the speeds predicted by the first model (equation (13), lower curve) and the second model (equation (21), upper curve) as a function of the random walk bias β, see equation (38). The hyperbolic reaction–diffusion (HRD) speed (14), which was derived in [9] and corresponds to the non-biased limit ($\beta = 0$) of first model, is also shown. It is seen that the difference of the first model relative to the HRD speed can be substantial, as large as 30%. The differences of the second model relative to the HRD speed are still more important (up to 55%).

In figure 1, the first model seems compatible with the observed speed for high enough values of β, whereas the second model seems not. In principle, we expected the second model to be superior to the first model for this application (because it involves a biological population). But from figure 1, it appears that it is not. However, this may be too strong a conclusion in view of the uncertainty of the values of the parameters. We think that dispersion data in many directions should be analysed in order to estimate the mobility (D_x) and bias (U_x) parameters for this human population accurately, as well as their error ranges and their dependence on position. This would yield a nonhomogeneous framework which, in contrast to that in [32], would be free of some relatively strong assumptions (e.g. the fractal nature of pathways, the use of adjustable parameters, etc). Assuming that sufficient data could be found for this purpose, such a project would certainly require very tedious work and discussions, which we feel more appropriate for a specialized demography publication. Here, our aim is not to present an in-depth analysis of the demographic data. Rather, the main point in the present paper is to show that physical models (arising from biased random walks) can be useful to describe such kind of biophysical processes. Indeed, they yield quite different speeds than non-biased models (curves versus horizontal lines in figure 1, respectively). So our new speed formulae can be useful in several biophysical applications. An especially augurious field is that of microorganisms, because there the experiments can be replicated and the parameter values are much more certain [10]. Although we are not yet aware of experimental front speeds of microorganisms arising from biased random walks, our work provides a theoretical basis that can be useful when they become available.

[4] For the migration data in [33] only the adult subpopulation is considered, well-known 2D diffusion theory can be applied (without reproduction terms) and we can estimate Dx and U_x from $\langle\Delta^2\rangle/(4T) = \langle r^2\rangle/(4t) = D_x$ (see equation (34)) and $\langle\Delta\rangle/T = \langle r\rangle/T = \sqrt{\pi D/T}$ (see also [25], equation (9.10)).

[5] One can try to estimate b from equation (33) with the minus sign and the migration data in [33] into the ENC region (which has more migration directions than the other regions). Those data are clearly biased, with more migrations from the West than from the East. However, the results are quite different if we estimate b using the horizontal directions, than if diagonal directions are used (e.g. the latter approach gives $b = 0.113$, thus $\beta = 0.7$). This shows the need for more detailed data, i.e. in many directions, so that a fit can be made to equation (33) in order to try to estimate b accurately. We are not aware of such detailed data, so it is more reasonable to analyse the front speed as a function of β (figure 1) at this stage.

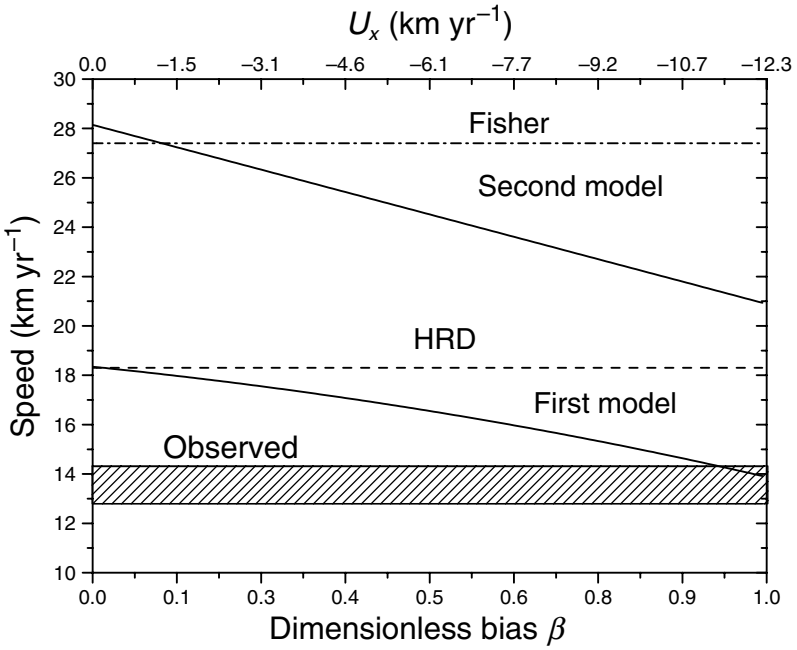

Figure 1. Predicted speeds for the human invasion of the USA in the 19th century, as a function of the random walk bias in the migration of individuals. The speeds shown are that according to the first model, equation (13), and to the second model, equation (21). The HRD speed (14), which was derived in [9], and its limit $T \to 0$ (Fisher) are also shown for comparison. The observed speed range is shown as a hatched rectangle.

At this stage, we think that the most important point is to realize that a bias in the random walk can have an important effect on the front speed. Note that both in the first model and in the second one, there is an advection term and a diffusion term. Their relative importance as a function of distance can be estimated by means of the Peclet number,

$$Pe = \frac{U_x L}{D_x}, \tag{40}$$

which for the human invasion application in this section becomes of order 1 for distances L of the order of 500 km. This is a scale similar to that in which the front speed is measured [30, 31], which supports our proposal that both advection and diffusion can be important in this illustrative application.

5. Concluding remarks

We have presented microscopic derivations of reaction–diffusion equations arising from biased random walks, and formulae for the speeds of their front solutions. Two models have been presented. Firstly, we have generalized a widely-used approach [1, 9, 10, 24] (the first or non-sequenced model, section 2). Then, we have presented a more realistic model which includes the effect of the time-order of events (the second or sequenced model, section 3). As an illustration, in section 4 we have applied both models to the human population colonization of North-America in the 19th century, and concluded that the effect of the bias in the random

walk can substantially change the value of predicted front speed. Although more detailed data are required for an in-depth analysis of this specific application, it shows that the biased front models we have presented can be useful for a variety of physical and biophysical applications dealing with biased fronts, such as particle diffusion in disordered lattices [16], nucleation of spiral waves [14], human and nonhuman population invasions [7], the spread of epidemics [34], cultural fronts, etc. For purely physical applications [14, 16] (not involving biological reproduction), the first model introduced here (section 2) is more appropriate. For biophysical ones [7, 9, 10, 34], the second model (section 3) seems more reasonable.

Acknowledgments

This work was funded by the European Commission (grant NEST-28192-FEPRE), the MEC-FEDER (grant FIS-2006-12296-C02-02) and the Generalitat de Catalunya (grant SGR-2005-00087).

Appendix. Macroscopic derivation of the first model

In section 2.2, we have considered a region centred about the x-axis which is sufficiently small so that the y-dependence of $p(x, y, t)$ can be neglected. Then equation (2) becomes simply

$$\frac{\partial p}{\partial t} + \frac{T}{2}\frac{\partial^2 p}{\partial t^2} = -U_x \frac{\partial p}{\partial x} + D_x \frac{\partial^2 p}{\partial x^2} + F + \frac{T}{2}\frac{\partial F}{\partial t}. \tag{A.1}$$

It is easy to see that this equation can also be derived by combining the following set of phenomenological equations

$$\begin{cases} \dfrac{\partial p}{\partial t} + \dfrac{\partial J}{\partial x} = F, \\ J + \tau \dfrac{\partial J}{\partial t} = U_x p - D_x \dfrac{\partial^2 p}{\partial x^2}, \end{cases} \tag{A.2}$$

where J is the diffusion flux and $\tau \equiv \frac{T}{2}$ is the relaxation time. The first equation of this set is just a mass balance equation, whereas the latter one a first-order Taylor expansion for a time-delayed flux,

$$J(x, t + \tau) = U_x p - D_x \frac{\partial^2 p}{\partial x^2}. \tag{A.3}$$

From this equation, we can say that the macroscopic effect arising from a direction-dependent microscopic motion of the particles is to introduce an additional flux $U_x p$ to the usual diffusion flux $-D_x \frac{\partial^2 p}{\partial x^2}$. In contrast, the effect of a finite jump time ($\tau \neq 0$) is to introduce a delay in the whole flux J. Equation (A.3) for the non-delayed limit $\tau = 0$ is well-known to arise from biased random walks [35].

This simple macroscopic derivation of equation (A.1) from the set (A.2) is not enough to apply the model to experimental data (section 4). The reason is that the set (A.2) is written in terms of the relaxation time τ and the macroscopic 'speed' U_x. But the identification of the relaxation time τ as half the time interval between successive jumps (i.e. $\tau = T/2$) and of the macroscopic 'speed' as $U_x = \langle \triangle_x \rangle / (4T)$ are absolutely necessary to predict the front speed (13) for each specific system (section 4). These two key results can be derived only by the microscopic derivation in section 2.1—but not from the macroscopic set (A.2).

A macroscopic derivation of the second model does not seem possible. The reason is that such physical macroscopic equations do not take the time order of biological reproduction and dispersal into account. Indeed, this is the main feature of the second model.

References

[1] Fort J and Méndez V 2002 *Rep. Prog. Phys.* **65** 895
[2] van Saarloos W 2003 *Phys. Rep.* **386** 29
[3] Merikoski J, Maunuksela J, Myllys M and Timonen J 2003 *Phys. Rev. Lett.* **90** 024501
[4] Fort J, Pujol T and Cukrowski A S 2000 *J. Phys. A: Math. Gen.* **33** 6953–73
[5] Fort J, Campos D, González J R and Velayos J 2004 *J. Phys. A: Math. Gen.* **37** 7185–98
[6] Campos D, Llebot J E and Fort J 2004 *J. Phys. A: Math. Gen.* **37** 6609–21
[7] Shigesada N and Kawasaki K 1997 *Biological Invasions: Theory and Practice* (Oxford: Oxford University Press)
[8] Ammerman A J and Cavalli-Sforza L L 1984 *The Neolithic Transition and the Genetics of Population in Europe* (Princeton, NJ: Princeton University Press)
[9] Fort J and Méndez V 1999 *Phys. Rev. Lett.* **82** 867
[10] Fort J and Méndez V 2002 *Phys. Rev. Lett.* **89** 178101
[11] Fedotov S 2000 *J. Phys. A: Math. Gen.* **33** 7033
[12] Méndez V, Fedotov S, Campos D and Horsthemke W 2007 *Phys. Rev.* E **75** 011118
[13] Bär M, Hagbert A, Meron E and Thiele U 1999 *Phys. Rev. Lett.* **83** 2664
Bär M, Hagbert A, Meron E and Thiele U 2000 *Phys. Rev.* E **62** 366
[14] Wei H, Lilienkamp G, Davidsen J, Bär M and Imbihl R 2006 *Phys. Rev.* E **73** 016210
[15] Davison K, Dolukhanov P, Sarson G R and Shukurov A 2006 *J. Arch. Sci.* **33** 641
[16] Haus J E and Kehr K W 1987 *Phys. Rep.* **150** 263
[17] Meakin P 1983 *Phys. Rev.* B **28** 5221
[18] Hill N A 1997 *J. Theor. Biol.* **186** 503
[19] Nossal R and Weiss G H 1974 *J. Theor. Biol.* **47** 103
[20] Tranquillo R T and Lauffenburger D A 1987 *J. Math. Biol.* **25** 229
[21] Dawes A T and Edelstein-Keshet L 2007 *Biophys. J.* **92** 744
[22] Lotka A J 1956 *Elements of Mathematical Biology* (New York: Dover) pp 64–9
[23] van den Bosch F, Metz J A J and Diekmann O 1990 *J. Math. Biol.* **28** 529
[24] Fort J and Méndez V 1999 *Phys. Rev.* E **60** 5894
[25] Murray J D 2002 *Mathematical Biology* 2 vols, 3rd edn (Berlin: Springer)
[26] Weinberger H F 1978 *Nonlinear Partial Differential Equations and Applications* ed J Chadam (Berlin: Springer)
[27] Clark J S 1998 *Am. Nat.* **152** 204
[28] Hall C A S 1988 *Ecol. Modell.* **43** 5
[29] Colding E A and Hill N A 2005 *J. Theor. Biol.* **233** 573
[30] Paulin C O 1932 *Atlas of the Historical Geography of the United States* (Westport: Greenwood Press) pp 76–9
[31] Flanders S A 1998 *Atlas of American Migration* (New York: Facts of File)
[32] Campos D, Fort J and Méndez V 2006 *Theor. Popul. Biol.* **69** 88
[33] Ferrie J P 2006 *Historical statistics of the US, millenial edition* 5 vols ed S B Carter, S Gartner, M R Haines, A L Olmstead, R Sutch and G Wright (Cambridge: Cambridge University Press) chapter Ac, table 2
Ferrie J P 1996 *Historical Methods* **29** 141
[34] Ferguson N M *et al* 2003 *Nature* **425** 681
[35] Berg H C 1983 *Random Walks in Biology* (Princeton, NJ: Princeton University Press)