Propagating Fronts for Competing Species Equations with Diffusion

MIN MING TANG & PAUL C. FIFE

Communicated by J. B. MCLEOD

1. The notion of a traveling wave front in the context of population dynamics is a natural one and has undoubtedly been in existence a long time. This concept, however, was apparently first modeled by a differential equation by FISHER in 1937 [1]. Fisher had in mind a community in which the processes of natural selection and random spatial migration were evident. The front then represented a wave of spatial advance of a favorable gene. In the final analysis, FISHER's equation was simply the usual logistic growth equation, supplemented by an extra term u_{xx} representing the effect of spatial diffusion:

$$u_t = u_{xx} + ru(1-u).$$

He found, and Kolmogorov et al. [2] established rigorously, that for every number $c \geqq 2\sqrt{r}$, there is a wave front solution

$$u(x,t) = U(x-ct)$$

with velocity c. The front joins the stable state $u=1$ with the unstable state, $u=0$, in the sense that

$$U(-\infty)=1, \qquad U(\infty)=0.$$

Furthermore, for each such c the profile U is unique (modulo translations in x).

Besides the natural selection situation FISHER had in mind, his equation also applies to single species dynamic models incorporating logistic growth and spatial diffusion. The logistic equation for a single species is, of course, easily generalized to a pair of equations for the dynamics of two fairly arbitrarily interacting species:

$$\dot{u} = u f(u,w) \tag{1a}$$

$$\dot{w} = w g(u,w). \tag{1b}$$

The study of such systems goes back to the work of LOTKA and VOLTERRA in the early part of the century, and that work is by now classical. If $r_1 = f(0,0)$ and $r_2 = g(0,0)$ are positive, the origin is unstable for (1). Supposing there is, in addition, a stable state $(\tilde{u}, \tilde{w})$ for (1), one may ask a question analogous to FISHER's: when the effect of spatial migration is introduced into the model by inserting diffusion terms, does the resulting system support wave fronts joining the unstable state (0, 0) to the stable state $(\tilde{u}, \tilde{w})$?

The purpose of our paper is to answer this question affirmatively, in the case when the equilibrium $(\tilde{u}, \tilde{w})$ comes about through competition and crowding effects, i.e., when f and g are decreasing functions of u and w in a neighborhood of $(\tilde{u}, \tilde{w})$. The only additional major assumption we make is that $f(0,0)>f(u,w)>0$ and $g(0,0)>g(u,w)>0$ for $u\in(0,\tilde{u})$, $w\in(0,\tilde{w})$.

As with FISHER's original equation, fronts turn out to exist for all c greater than or equal to a certain minimal speed c^*. In FISHER's case, $c^*=2\sqrt{r}$, and in our case, if both species' diffusivities are equal to one, then $c^*=2\,\mathrm{Max}\,[\sqrt{r_1}, \sqrt{r_2}]$. But *unlike* the scalar case, we find that the fronts are not unique, even for a given c. In fact, for each $c\geqq c^*$, there is a one-parameter family of fronts. This multiplicity generally does not affect the rates of approach of the profiles to their limits at $\pm\infty$, as this is determined by c alone; it is rather reflected in the properties of these profiles for finite values of their argument.

Wave fronts of the type we study represent a progressive replacement of one equilibrium state (ahead of the front) by another (behind the front). It is important to realize that simple wave fronts do not provide the only mechanism for such a replacement. For example, this might be accomplished in two stages, an intermediate equilibrium being involved, and each of the two stages perhaps with its own wave front. An easy specific example would be when $f=r_1(1-u)$, $g=r_2(1-w)$, so that the two equations are uncoupled:

$$u_t-u_{xx}=r_1u(1-u)$$
$$w_t-w_{xx}=r_2w(1-w).$$

Then the dynamics of u and w are independent. The transition from the state (0, 0) to (1, 1) might then happen by means of a u-front traveling with minimal velocity $2\sqrt{r_1}$ (first stage), combined with a w-front traveling with minimal velocity $2\sqrt{r_2}$ (second stage). If $r_1>r_2$, then eventually the second stage, representing a transition from (1, 0) to (1, 1), will lag behind the first, a transition from (0, 0) to (1, 0). This type of combination of fronts is not covered in the present paper.

Other work has been done on wave fronts for systems of the form (2), though their existence has not been proved; see, for example, [3].

2. The model we examine, then, consists of (1) with spatial migration in a one-dimensional habitat accounted for by diffusion terms:

$$u_t-k_1u_{xx}=uf(u,w), \tag{2a}$$
$$w_t-k_2w_{xx}=wg(u,w), \tag{2b}$$
$$(k_1>0,\ k_2>0).$$

Such systems are of considerable current interest. The functions $u(x,t)$ and $w(x,t)$ represent the densities of the two species, as functions of space and time.

We assume the existence of an equilibrium point $(\tilde{u}, \tilde{w})$ of (1) with $\tilde{u}>0$, $\tilde{w}>0$. We further assume it is stable, so the eigenvalues of the linearization of the right-hand side about $(\tilde{u}, \tilde{w})$ both have negative real parts. This is equivalent to the pair of conditions

$$D \equiv \det \begin{bmatrix} f_1 & f_2 \\ g_1 & g_2 \end{bmatrix} > 0; \qquad \tilde{u} f_1 + \tilde{w} g_2 < 0, \tag{3}$$

where $f_1 \equiv \frac{\partial f}{\partial u}(\tilde{u}, \tilde{w})$, etc.

In accordance with our interpretation of (1) as competing species equations, we also assume that

$$f_1 < 0, \quad g_2 < 0, \quad f_2 \leqq 0, \quad g_1 \leqq 0 \tag{4}$$

and

$$0 < f(u, w) < r_1, \quad 0 < g(u, w) < r_2 \tag{5}$$

for

$$u \in (0, \tilde{u}), \quad w \in (0, \tilde{w}),$$

where we have set $r_1 = f(0,0)$, $r_2 = g(0,0)$. Our basic result is the following.

Theorem. *For each* $c \geqq [\mathrm{Max}(4 r_1 k_1, 4 r_2 k_2)]^{1/2} \equiv c^*$, *there exists a one parameter family of wave front solutions of* (2):

$$u = U(x - ct), \quad w = W(x - ct),$$

joining $(0,0)$ *with* $(\tilde{u}, \tilde{w})$, *in the sense that*

$$U(-\infty) = \tilde{u}, \quad W(-\infty) = \tilde{w}, \quad U(\infty) = W(\infty) = 0. \tag{6}$$

The functions U and W are positive for finite values of their argument. Moreover, there exist no positive wave fronts for $c < c^*$.

3. The functions U and W will necessarily satisfy the pair of equations

$$k_1 U'' + c U' + U f(U, W) = 0, \tag{7a}$$

$$k_2 W'' + c W' + W g(U, W) = 0, \tag{7b}$$

primes denoting differentiation with respect to $z = x - ct$, and the proof of the theorem involves a qualitative analysis of the trajectories of (7). These equations are more conveniently written as a system of first order equations, identifying $q_1 = U$, $q_2 = U'$, $q_3 = W$, $q_4 = W'$:

$$q_1' = q_2 \tag{8a}$$

$$q_2' = -\frac{c}{k_1} q_2 - \frac{q_1}{k_1} f(q_1, q_3) \tag{8b}$$

$$q_3' = q_4 \tag{8c}$$

$$q_4' = -\frac{c}{k_2} q_4 - \frac{q_3}{k_2} g(q_1, q_3). \tag{8d}$$

Linearization of this system about the critical point (0, 0, 0, 0) reveals that no trajectory can approach the origin as $z\to\infty$ with q_1 and q_3 remaining positive, unless $c\geqq c^*$. This establishes the last part of the theorem.

4. All the trajectories we construct will lie within a certain region $\Delta\subset\mathbb{R}^4$, which we now define. Let

$$\boldsymbol{q}=(q_1,\dots,q_4),$$
$$m_i=-\frac{1}{2k_i}\left(c-\sqrt{c^2-4r_ik_i}\right)<0, \qquad i=1,2,$$
$$\Delta=\{\boldsymbol{q}: m_1q_1<q_2<0,\ m_2q_3<q_4<0,\ 0<q_1<\tilde{u},\ 0<q_3<\tilde{w}\}.$$

Lemma 1. *Let $c\geqq c^*$. Let $\boldsymbol{q}(z)$ be a trajectory of (8) satisfying $\boldsymbol{q}(0)\in\Delta$. Then* $\lim\limits_{z\to\infty}\boldsymbol{q}(z)=0$.

Proof. Since $q_1'=q_2<0$ and $q_3'=q_4<0$ in Δ, we see that q_1 and q_3 decrease with z as long as $\boldsymbol{q}(z)\in\Delta$. If $\boldsymbol{q}$ leaves Δ for some $z>0$, let z_0 be the first value at which $\boldsymbol{q}(z)\notin\Delta$. Then at $z=z_0$, one of the four inequalities

$$m_1q_1<q_2<0, \qquad m_2q_3<q_4<0$$

must be replaced by an equality. We consider the various possibilities in turn. If $q_2(z_0)=0$ with $q_1>0$, then necessarily $q_2'(z_0)\geqq 0$, so that from (8b)

$$0\leqq k_1q_2'(z_0)=-cq_2-q_1f(q_1,q_3)=-q_1f(q_1,q_3).$$

This inequality contradicts (5). A similar contradiction shows that $q_4(z_0)\neq 0$ if $q_3(z_0)>0$.

Now suppose that $m_1q_1(z_0)=q_2(z_0)$. Then since the point $(q_1(z_0),q_2(z_0))$ must be approached from within Δ, it follows that

$$\frac{q_2'(z_0)}{q_1'(z_0)}\geqq m_1.$$

Substituting (8) into this inequality and using the fact that $q_2=m_1q_1<0$, we find this implies

$$f(q_1,q_3)\geqq -m_1^2k_1-cm_1.$$

By the definition of m_1, the right-hand side is equal to r_1. This again contradicts (5), unless of course $q_1=q_3=0$ at $z=z_0$. But this is impossible because the uniqueness property of differential equations would then imply $q_1\equiv q_3\equiv 0$.

In a similar manner we see that it is impossible for the trajectory to first arrive at the boundary of Δ at a point where $q_4=m_2q_3$.

Therefore the trajectory stays within Δ for all z, with q_1 and q_3 decreasing as z increases. Hence $\lim\limits_{z\to\infty}q_i=\overline{q_i}$, $i=1$ and 3. It also follows from the boundedness of $\boldsymbol{q}(z)$ and (8) that $\boldsymbol{q}'$ is bounded. In particular, the boundedness of $q_2'=q_1''$, together with the limit relation $q_1\to\overline{q_1}$, implies $q_1'(z)=q_2(z)\to 0$. Thus a point on the line $q_2=0$ is approached as $z \to \infty$. The analogous argument above can be used again to show $\overline{q_1} = 0$. Similarly, $q_3(\infty) = \overline{q_3} = 0$ and $q_4(\infty) = 0$. This completes the proof.

5. Our purpose, of course, is to show the existence of trajectories $\boldsymbol{q}(z)$ with $\boldsymbol{q}(-\infty) = (\tilde{u}, 0, \tilde{w}, 0) \equiv \tilde{\boldsymbol{q}}$ and $\boldsymbol{q}(\infty) = 0$. In view of Lemma 1, it suffices to show the existence of trajectories with $\boldsymbol{q}(-\infty) = \tilde{\boldsymbol{q}}$ and $\boldsymbol{q}(z) \in \Delta$ for some z. Such trajectories lie on the unstable manifold M of $\tilde{\boldsymbol{q}}$. Locally at $\tilde{\boldsymbol{q}}$, M is tangent to the unstable manifold Q of the linearization of (8) about $\tilde{\boldsymbol{q}}$, and a small neighborhood of $\tilde{\boldsymbol{q}}$ on one is the homeomorphic image of a small neighborhood of $\tilde{\boldsymbol{q}}$ on the other. The existence of such trajectories will therefore follow, if it can be shown that

$$Q \cap \Delta \neq \emptyset. \tag{9}$$

We set $\boldsymbol{p} = \boldsymbol{q} - \tilde{\boldsymbol{q}}$ and write the linearization of (8) about $\tilde{\boldsymbol{q}}$ in the form

$$\boldsymbol{p}' = A\,\boldsymbol{p},$$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\dfrac{\tilde{u}}{k_1} f_1 & -\dfrac{c}{k_1} & -\dfrac{\tilde{u}}{k_1} f_2 & 0 \\ 0 & 0 & 0 & 1 \\ -\dfrac{\tilde{w}}{k_2} g_1 & 0 & -\dfrac{\tilde{w}}{k_2} g_2 & -\dfrac{c}{k_2} \end{bmatrix}.$$

The manifold Q is spanned by the eigenvectors of A corresponding to eigenvalues with positive real part.

Lemma 2. *$\tilde{\boldsymbol{q}}$ is a hyperbolic critical point, and Q is two-dimensional.*

For the proof, we introduce the matrix $A(\rho)$ defined, for $\rho \in [0, 1]$, by simply replacing the quantity f_2 in the second row, third column of A by ρf_2. We show that $A(\rho)$ has two positive and two negative eigenvalues. This fact for $\rho = 1$ is the assertion of the lemma. The characteristic equation of $A(\rho)$ is

$$F_\rho(\lambda) \equiv \lambda^4 + a\lambda^3 + b\lambda^2 + c\lambda + d(\rho) = u,$$

where

$$d(\rho) = \frac{\tilde{u}\tilde{w}}{k_1 k_2} D(\rho), \qquad D(\rho) = \begin{vmatrix} f_1 & \rho f_2 \\ g_1 & g_2 \end{vmatrix},$$

and the coefficients a, b, c are independent of ρ. By (3) and (4),

$$D(\rho) = D + (1 - \rho) g_1 f_2 \geqq D > 0,$$

whence

$$d(\rho) = F_\rho(0) > 0. \tag{10}$$

First, consider the case $\rho = 0$. Then the set of eigenvalues of $A(0)$ is the union of the eigenvalues of its two principal 2×2 blocks

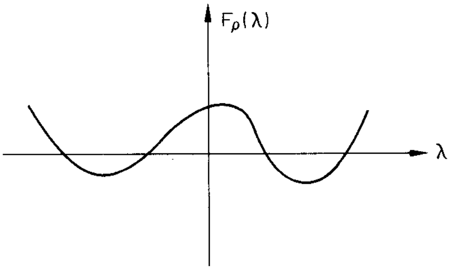
$$\begin{bmatrix} 0 & 1 \\ -\dfrac{\tilde{u}}{k_1} f_1 & -\dfrac{c}{k_1} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 1 \\ -\dfrac{\tilde{w}}{k_1} g_2 & -\dfrac{c}{k_2} \end{bmatrix}.$$

These are exactly calculated and found to be

$$-\frac{1}{2k_1}\left(c\pm\sqrt{c^2-4k_1 f_1}\right) \quad \text{and} \quad -\frac{1}{2k_2}\left(c\pm\sqrt{c^2-4k_2 g_2}\right). \tag{11}$$

We see from (4) that these are real; two are positive, and two are negative.

Therefore for $\rho=0$, $F_\rho(\lambda)$ has the properties indicated in the figure: $F_0(0)>0$, $F_0(\lambda)>0$ for large $|\lambda|$, and $F_0(\lambda)<0$ for some positive and negative values of λ. The assertion of the lemma is true in this case.



When ρ is now increased from 0 to 1, the effect on the graph of the function F_ρ is to shift it upwards or downwards. Furthermore since $g_1 f_2 \geqq 0$, the coefficient d depends monotonically on ρ, so when ρ is increased, d does not increase, and the graph of F_ρ is not shifted upward. This means it still assumes negative values for some positive and negative values of λ. By virtue of (10), however, $F_1(0)=d(1)$ remains positive, so that F_1 continues to have two negative and two positive roots. This proves the lemma.

6. For the following, we introduce the notation

$Q(\rho)=$ the unstable manifold of $\boldsymbol{p}'=A(\rho)\,\boldsymbol{p}$;

$N=$ the negative cone in $\mathbb{R}^4$:

$$N=\{\boldsymbol{p}\in\mathbb{R}^4\colon p_i<0,\ i=1,\dots,4\};$$

$S=$ unit sphere in $\mathbb{R}^4$;

$R(\rho)=Q(\rho)\cap N\cap S$.

Lemma 3. $Q\cap N\neq\emptyset$.

Proof. The manifold $Q(\rho)$ is the span of the eigenvectors of $A(\rho)$ with positive eigenvalues. By Lemma 2, it is two-dimensional. Clearly $Q=Q(1)$. Furthermore, $A(\rho)$ and $Q(\rho)$ depend continuously on ρ and we can identify normalized eigenvectors $\phi^1(\rho)$ and $\phi^2(\rho)$ on $Q\cap S$ which are linearly independent and depend continuously on ρ. Let λ_1 and λ_2 be the corresponding positive eigenvalues.

When $\rho=0$, we may calculate ϕ^i explicitly:

$$\phi^1(0)=(1+\lambda_1^2)^{-1/2}(0,0,1,\lambda_1),$$

where from (11),

$$\lambda_1=(2k_2)^{-1}\left(-c+(c^2-4k_2 g_2)^{1/2}\right);$$

also

$$\phi^2(0)=K(1,\lambda_2,\beta,\beta\lambda_2)$$

for some constants K and β. It is clear that a linear combination of $\phi^1(0)$ and $\phi^2(0)$ exists which lies in N. Therefore $Q(0)\cap N\neq\emptyset$.

Now as ρ increases from 0 to 1, either $Q(\rho)\cap N$ remains non-empty or there is a first value $\rho=\rho_0$ at which $Q(\rho)\cap N=\emptyset$. We show that the second alternative cannot occur.

If $f_2=0$, nothing changes with ρ and we are through. A similar argument establishes the result if $g_1=0$. Therefore below we assume $f_2<0$, $g_1<0$.

Since $Q(\rho)$ is a two-dimensional plane, $Q(\rho)\cap S$ is a curve in $\mathbb{R}^4$, and (for $\rho<\rho_0$) $R(\rho)$ an arc of that curve, with nonzero endpoints on the boundary of N. This boundary consists of portions of the coordinate hyperplanes $T_i=\{\boldsymbol{p}: p_i=0\}$, $i=1,\dots,4$. Let $\psi_1(\rho)$ and $\psi_2(\rho)$ be the endpoints of $R(\rho)$. When $R(\rho)\neq\emptyset$, the points ψ_1 and ψ_2 lie on different T_i's. This is clear because $\psi_1(\rho)$ and $\psi_2(\rho)$ generate $Q(\rho)$ and if both were on the same T_i, then $Q(\rho)$ would lie on T_i, contradicting the fact that $Q(\rho)\cap N\neq\emptyset$.

Suppose the second alternative were true. Since the points $\psi_i(\rho)$, $0\leqq\rho<\rho_0$, lie in a bounded set, there exists a sequence $\rho_n\uparrow\rho_0$ with $\psi_i(\rho_n)\to\bar{\psi}_i$, so that $\bar{\psi}_i\in Q(\rho_0)$ and each $\bar{\psi}_i$ is on some T_i. Two cases arise:

(1) $\bar{\psi}_1\neq\bar{\psi}_2$. Then $\bar{\psi}_1$ and $\bar{\psi}_2$ generate $Q(\rho_0)$. Moreover, they must lie on a common hyperplane T_l; otherwise a linear combination of them would lie in N, which is contrary to the definition of ρ_0. It follows that the entire manifold $Q(\rho_0)$ must lie on T_l, so that for some $l=1, 2, 3$, or 4, the l^{th} component of all vectors in $Q(\rho_0)$, in particular of ϕ^1 and ϕ^2, vanish.

The equation $A(\rho)\phi=\lambda\phi$ in component form is

$$\lambda\phi_1=\phi_2, \tag{11a}$$

$$\lambda\phi_2=-\frac{c}{k_1}\phi_2-\frac{\tilde{u}}{k_1}(f_1\phi_1+\rho f_2\phi_3), \tag{11b}$$

$$\lambda\phi_3=\phi_4, \tag{11c}$$

$$\lambda\phi_4=-\frac{c}{k_2}\phi_4-\frac{\tilde{w}}{k_2}(g_1\phi_1+g_2\phi_3). \tag{11d}$$

We return to the statement that $Q(\rho_0)\in T_l$. If $l=1$ or 2, then by (11a), $\phi_1^i=\phi_2^i=0$; and if $l=3$ or 4, $\phi_3^i=\phi_4^i=0$, so in fact *two* components vanish for all vectors in $Q(\rho_0)$.

(2) $\bar{\psi}_1=\bar{\psi}_2\equiv\bar{\psi}$. Then since $\psi_i(\rho)$ lie on distinct T_i for $\rho<\rho_0$, it must happen that $\bar{\psi}$ lies on the intersection of two T's, so that two components vanish.

Thus in *either* case, some nontrivial linear combination of ϕ^1 and ϕ^2 has two components vanishing:

$$C_1\phi_l^1+C_2\phi_l^2=0, \quad l=l_1 \;\; \textit{and} \;\; l_2, \; C_i \text{ not both zero.}$$

First, suppose $(l_1,l_2)=(1,2)$. Then by (11a),

$$C_1\phi_1^1+C_2\phi_1^2=0 \tag{12}$$

and

$$C_1 \lambda_1 \phi_1^1 + C_2 \lambda_2 \phi_1^2 = 0. \tag{13}$$

But since $\lambda_1 \neq \lambda_2$, this implies $\phi_1^1 = \phi_1^2 = 0$. From (11a), this further yields $\phi_2^1 = 0$; from (11b), $\phi_3^1 = 0$; and from (11c), $\phi_4^1 = 0$. This is impossible, because $\phi^1 \in S$. Similarly, the case $(l_1, l_2) = (3, 4)$ is excluded.

Next, suppose $(l_1, l_2) = (1, 3)$. Then (13) is replaced by

$$C_1 \phi_3^1 + C_2 \phi_3^2 = 0. \tag{14}$$

Equations (11c) and (11d) yield the following relation between ϕ_3^i and ϕ_1^i:

$$P(\lambda_i)\phi_3^i = -\frac{\tilde{w}}{k_2} g_1 \phi_1^i,$$

where $P(\lambda) = \lambda^2 + \frac{c}{k_2}\lambda + \frac{\tilde{w}}{k_2} g_2$. Here we may exclude the possibility that $P(\lambda) = 0$, since this would imply $\phi_1^i = \phi_2^i = 0$, and we would be in the case considered above. Since $P(\lambda_i) \neq 0$, (14) becomes

$$\frac{C_1}{P(\lambda_1)} \phi_1^1 + \frac{C_2}{P(\lambda_2)} \phi_1^2 = 0.$$

As before, this equation and (14) have a nontrivial solution only if $P(\lambda_1) = P(\lambda_2)$, which implies

$$\lambda_1 + \lambda_2 + \frac{c}{k_2} = 0.$$

But this is impossible, since $\lambda_i > 0$ and $c > 0$.

The only other case is $(l_1, l_2) = (2, 4)$. By the same argument as before, we have

$$\frac{P(\lambda_1)}{\lambda_1} = \frac{P(\lambda_2)}{\lambda_2},$$

which implies $0 < \lambda_1 \lambda_2 = \frac{\tilde{w}}{k_2} g_2 \leqq 0$. Again, this is impossible.

All possibilities being exhausted, we conclude that at no value $\rho \in [0, 1]$ does it happen that $Q(\rho) \cap N = \emptyset$. In particular, for $\rho = 1$ we obtain the desired conclusion.

7. From Lemma 3, it follows that the manifold M enters N. And since M is two-dimensional, we have in fact a one-parameter family of trajectories on $M \cap N$. Each of these trajectories goes to the critical point $\tilde{\boldsymbol{q}}$ as $z \to -\infty$ (by the definition of M), and to the origin as $z \to \infty$ (by Lemma 1). This proves the theorem.

8. For each $c \geqq c^*$, we have shown the existence of a one-parameter family of wave fronts for which each component u, w is monotone decreasing. This monotonicity property is clear from our arguments because $\boldsymbol{q}(z) \in \Delta$, this implying that $q_2 = U' < 0$ and $q_4 = W' < 0$. At the same time, it is clear that monotonicity is not a necessary condition, but only a convenience for our proof. Very likely nonmonotone fronts exist as well.

Note. This research was sponsored in part by the National Science Foundation under Grant MPS-74-06835-01.

References

1. FISHER, R. A., The wave of advance of advantageous genes. Ann. of Eugenics **7**, 355–369 (1937)
2. KOLMOGOROV, A. N., PETROVSKIĬ, I. G., and PISKUNOV, N. S., A study of the equation of diffusion with increase in the quantity of matter, and its application to a biological problem. Bjul. Moskovskovo Gos. Univ. **17**, 1–72 (1937)
3. MURRAY, J. D., On travelling wave solutions in a model for the Belousov-Zhabotinskii reaction. J. Theor. Biol. **56**, 329–343 (1976)

Department of Mathematics
University of Missouri
Rolla, Missouri
and
Department of Mathematics and
Program in Applied Mathematics
University of Arizona
Tucson, Arizona

(Received September 15, 1978)