ELECTROPHORETIC TRAVELING WAVES

P. C. FIFE, O. A. PALUSINSKI AND Y. SU

ABSTRACT. An existence-uniqueness-approximability theory is given for a prototypical mathematical model for the separation of ions in solution by an imposed electric field. The separation is accomplished during the formation of a traveling wave, and the mathematical problem consists in finding a traveling wave solution of a set of diffusion-advection equations coupled to a Poisson equation. A basic small parameter ε appears in an apparently singular manner, in that when $\varepsilon = 0$ (which amounts to assuming the solution is everywhere electrically neutral), the last (Poisson) equation loses its derivative, and becomes an algebraic relation among the concentrations. Since this relation does not involve the function whose derivative is lost, the type of "singular" perturbation represented here is nonstandard. Nevertheless, the traveling wave solution depends in a regular manner on ε, even at $\varepsilon = 0$; and one of the principal aims of the paper is to show this regular dependence.

1. **Introduction.** Electrophoresis is a term used to describe the movement of charged particles in solution exposed to an electric field. Electrophoretic techniques are extensively used in the separation of proteins and other biological materials [De]. The compounds of interest in solution are generally capable of reacting either as an acid or as a base (amphoretic compounds); in the ionization process the species may acquire either positive or negative charges. The charge of a given species in a solvent is a complex function of the sum total of the compound's dissociable groups and the inherent properties of the solvent. When a solution containing amphoretic compounds is exposed to an electric field, the movement of ions and uncharged species occurs simultaneously with dissociation-recombination reactions. The resulting transport is highly specific, so that a mixture of different compounds can be separated on the basis of different transport properties.

The dissociation-recombination reactions occur on much shorter time scales than that of transport by diffusion or electromigration, and it is usually assumed that those reactions are always in equilibrium. Under this assumption, the model for electrophoresis simplifies, taking the form of nonlinear advection-diffusion equations coupled with nonlinear algebraic relations characterizing the chemical equilibria. Moreover, the experimental setup is often such that variation occurs in only one spatial dimension (along a column). A mathematical model for electrophoresis with these simplifications (which we assume here) was given in [SP].

There are several different electrophoretic techniques which are described by the same equations, but which differ in the boundary and/or initial conditions. For example if the ends of the column are impermeable to the amphoretic compounds, the process will reach steady state and the various species will focus in neighborhoods of their respective isoelectric positions. This separation technique is widely used and is called isoelectric focusing. The initial concentrations of compounds in the isoelectric focusing process are usually uniform.

In this paper we present a study of another extensively used separation technique, called isotachophoresis (ITP). The usual setup for ITP consists of a reaction column connected at both ends to reservoirs of buffer electrolytes. The reservoirs are large enough that the processes in the column have a negligible effect on the composition of the buffers. This assures a fixed concentration of components at the column ends; Dirichlet boundary conditions are therefore appropriate. The column is filled with specially prepared solutions (Figure 1), called common, leading, and terminating electrolytes, and a sample containing a mixture of species is then inserted as the interface between the leading and terminating solutions. Such nonuniform initial conditions are characteristic for the ITP process. After the sample insertion, an electric field is applied. Following a transient period, a set of contiguous zones is formed (Figure 2) with all zones traveling along the column with the same constant velocity. This particular feature of the process gives rise to its name.

LEADING ELECTROLYTE

TERMINATING ELECTROLYTE

COMMON ELECTROLYTE (COUNTERION)
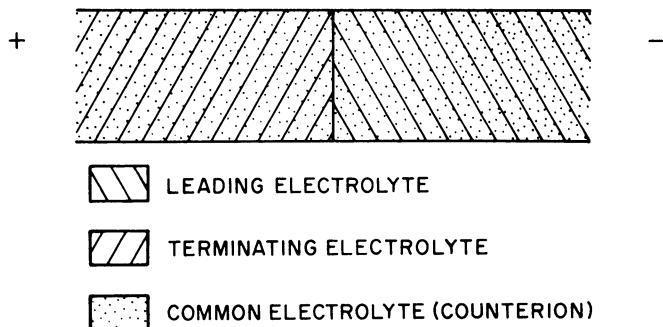
INITIAL SETUP OF ELECTROLYTES IN ITP SEPARATION

FIGURE 1

A quantitative description of the traveling waves uses the concept of mobility. An ion will move with velocity proportional to the product of its charge and the electric field. The proportionality constant is defined to be its mobility Ω. The mobilities $(m^2/v \cdot \text{sec})$ of the positive ions A_1^+, A_2^+, and A_3^+ in the three zones depicted in Figure 2 are Ω_1, Ω_2, and Ω_3, and they satisfy the inequalities

$$\Omega_1 < \Omega_2 < \Omega_3.$$

Within a given zone, the composition is approximately uniform, changes being mainly concentrated in transitional regions between the zones. The composition increments from one zone to the next are governed by well-known regulating functions, first derived by Kohlrausch [K] for strong (fully ionized) electrolytes. The

ZONE 3	ZONE 2	ZONE 1	
$[C-]\Omega_-$	$[C-]\Omega_-$	$[C-]\Omega_-$	
$[A_1^+]\Omega_1$	$[A_2^+]\Omega_2$	$[A_3^+]\Omega_3$	

(to the left of the table: $+$; to the right: $-$)

TRAVELING ZONES IN ITP STEADY STATE

FIGURE 2

regulating relation between zones 1 and 2 is

$$\frac{\hat{n}_1}{\hat{n}_2} = \frac{\Omega_1(\Omega_2 + \Omega_-)}{\Omega_2(\Omega_1 + \Omega_-)},$$

where $\hat{n}_1$ and $\hat{n}_2$ are the positive species concentrations in zones 1 and 2 respectively, and Ω_- is the mobility of the common, negatively charged, species $C-$ called counterion. This regulating relation is readily derived from the governing equations, assuming experimentally observable constant velocity of movement and uniform composition away from the transition regions.

There are several variations on the idea of regulating functions in the literature; they determine the concentrations of species inside the various zones and the traveling velocity. They were derived for various systems like multicomponent configurations involving zones traveling in both directions [**Do**], simple weak acids and bases [**A**] and simple ampholytes [**SP**].

The analysis of the transition regions is much more complicated and is usually performed with the aid of a digital computer used to generate numerical approximations to the solutions of governing equations.

In this paper we consider a prototypical transition region separating two positive ions with different mobilities; a common negative counterion is also present. The types of solutions we obtain are expected to describe typical transition regions in more complex isotachophoretic waves with a greater number of ions; but we idealize by assuming that only the ions mentioned are present.

The mathematical problem (formulated in §2) becomes that of finding a connection between two critical points of a fourth-order system of differential equations (the four dependent variables represent the concentrations of the three kinds of ions and the electric field). The last of these equations involves a small parameter ε multiplying its derivative term. This singularly perturbed form is nonstandard because when $\varepsilon = 0$, the resulting algebraic equation does not involve the function (the electric field) whose derivative has been lost.

The existence of a solution is proved by means of a compactness argument in §3. A crucial part of the method is to convert the singularly perturbed equation mentioned above to a second order equation subject to analysis by a maximum principle. This conversion is useful for the numerical analysis as well, as will be seen in a later paper. When $\varepsilon = 0$, an explicit unique solution can be obtained; this is shown in §2. §4 is devoted to showing that the solution proved to exist in §3 for $\varepsilon > 0$ lies uniformly $O(\varepsilon)$ close to the solution for $\varepsilon = 0$; in this sense, the perturbation turns out to be regular rather than singular. Finally, the uniqueness of the traveling wave connection is proved in §5. The most crucial step in the argument in these last two sections involves showing that the exponentially decaying solutions

of a certain linear system satisfy a certain inequality (4.8); this result in turn is based on some very delicate structural properties of a Jacobian matrix.

2. The mathematical model for ITP; **electrically neutral solution.** The basic object of study is a mixture of charged species in solution, subject to an imposed electric field. Shortly we will specialize to the case when the number of species is only 3, but for now denote it by m. The concentrations of the m species (in moles/volume) are denoted by the vector $\hat{n} = (\hat{n}_1, \ldots, \hat{n}_m)$. The basic assumption is that each species has a characteristic mobility constant Ω_i such that the velocity of its ions, when acted upon by a potential gradient, is given by

$$(2.1) \qquad v_i = -\Omega_i z_i \frac{\partial \phi}{\partial \hat{x}},$$

where z_i is its charge, in units of the absolute value of the electron charge, and ϕ is the potential. The species also diffuse, with diffusivities D_i. The flux f_i of the ith species is then given by

$$f_i = v_i \hat{n}_i - D_i \frac{\partial \hat{n}_i}{\partial \hat{x}}.$$

From this and (2.1), assuming strong electrolytes, we may write the species balance equations for a one-dimensional configuration in the form

$$(2.2) \qquad \frac{\partial \hat{n}_i}{\partial t} = \frac{\partial}{\partial \hat{x}} \left[z_i \Omega_i \hat{n}_i \frac{\partial \phi}{\partial \hat{x}} + D_i \frac{\partial \hat{n}_i}{\partial \hat{x}} \right], \qquad i = 1, \ldots, m.$$

There is an additional equation, since ϕ must satisfy Poisson's equation

$$(2.3) \qquad \hat{\varepsilon} \frac{\partial^2 \phi}{\partial \hat{x}^2} = -e \sum z_k \hat{n}_k,$$

where e is the molar charge and $\hat{\varepsilon}$ the permittivity of the solvent.

We assume the Einstein relation $D_i = (RT/e)\Omega_i$. For the ITP problem, it is appropriate to look for traveling wave solutions, and so we introduce the traveling wave coordinate $\bar{x} = \hat{x} - Ut$, U being the velocity. Let $\hat{E} = \partial \phi / \partial \hat{x}$ be the electric field. Written in the traveling wave coordinate, (2.2) becomes an ordinary differential equation which can be integrated once to yield

$$(2.4) \qquad \frac{RT}{e} \Omega_i \frac{d\hat{n}_i}{d\bar{x}} = \Omega_i z_i \hat{n}_i \hat{E} - U\hat{n}_i + C_i,$$

the C_i being integration constants. Also (2.3) is transformed in the obvious manner. These equations are to be solved under appropriate boundary conditions at $\pm\infty$. In general there may also be integral conditions, but not in the special case which we shall now consider.

Suppose that $m = 3$, there being two positive ions ($\hat{n}_1$ and $\hat{n}_2$) with different mobilities $\Omega_1 < \Omega_2$, and one negative ion ($\hat{n}_3$). Thus $z = (1, 1, -1)$. We look for traveling wave solutions which serve to separate the two positive ions, the one with higher mobility ($\hat{n}_2$) being ahead of the other. In other words,

$$(2.5) \qquad \hat{n}_1(+\infty) = \hat{n}_2(-\infty) = 0.$$

The remaining boundary conditions, the velocity U, and the constants C_i are parameters in the problem which should either be prescribed or deduced from the

equations. Still another important parameter is the (constant) electric current through the medium, which is given in terms of the other parameters by

$$I = e \sum z_i (f_i + U\hat{n}_i).$$

In view of (2.3) (the right side vanishes at $\pm\infty$) and the fact that the diffusion term in f_i also vanishes there, the expression for the current simplifies when evaluated at $\pm\infty$:

(2.6) $$I = e \sum (z_i)^2 \Omega_i \hat{n}_i E_\pm,$$

where $E_\pm \equiv \hat{E}(\pm\infty)$.

It turns out to be physically reasonable to specify only the concentration ν of the leading positive ion at $+\infty$ ($\hat{n}_2(\infty) \equiv \mu$) and the current I; the velocity U and all other boundary values, as we shall see, can then be obtained a priori in a unique manner. First of all, if we set $\nu = \hat{n}_1(-\infty)$, the vanishing of the right side of (2.3) at $\pm\infty$, together with (2.5) implies

(2.7) $$\hat{n}_1 = \hat{n}_3 = \nu \text{ at } -\infty; \qquad \hat{n}_2 = \hat{n}_3 = \mu \text{ at } +\infty.$$

We use the following dimensionless variables and parameters:

$$x = \frac{eE_-}{RT}\hat{x}, \quad n_i = \frac{\hat{n}_i}{\mu}, \quad E = \frac{\hat{E}}{E_-}, \quad b_i = \frac{\Omega_1}{\Omega_i}, \quad i = 2, 3, \quad \theta = \frac{\nu}{\mu}, \quad \varepsilon = \frac{\hat{\varepsilon}E_-^2}{RT\mu}.$$

The inequality involving mobilities translates into

(2.8) $$b_2 < 1.$$

Introducing this notation into (2.4), (2.3) and evaluating the C_i from the conditions at both $+\infty$ and $-\infty$ yields the following system, dots denoting differentiation with respect to x:

(2.9a–d)
$$\dot{n}_1 = n_1 E - n_1, \quad \dot{n}_2 = n_2 E - b_2 n_2,$$
$$\dot{n}_3 = -n_3 E - b_3 n_3 + \theta(1 + b_3), \quad \varepsilon \dot{E} = n_1 + n_2 - n_3.$$

As boundary conditions, we have

(2.10a)   At $x = -\infty$: $n = (\theta, 0, \theta), \quad E = 1$;

(2.10b)   At $x = +\infty$: $n = (0, 1, 1), \quad E = b_2$.

Finally, all the remaining parameters $E_\pm$, ν, U, θ are obtained in terms of the Ω_i, μ, and I by the relations

$$I = e(\Omega_1 + \Omega_3)\nu E_- = e(\Omega_2 + \Omega_3)\mu E_+,$$
$$E_+ = b_2 E_-, \quad U = \Omega_1 E_-, \quad \theta = \nu/\mu.$$

These relations among the parameters are necessary for there to exist a traveling wave solution. The first of them implies

(2.11) $$\theta(1 + b_3) = (b_2 + b_3).$$

In typical cases, ε is small, of the order 10^{-3} or less.

The existence, uniqueness, and continuous dependence on ε (at $\varepsilon = 0$) of a solution of (2.9), (2.10) will be proved in later sections. In this, the positive parameters

b_i, ε, and θ are completely arbitrary except for the necessary restrictions (2.8) and
(2.11), and the restriction that ε be small enough. For $\varepsilon = 0$ (the electroneutrality
assumption), the problem has been solved explicitly, see e.g. [**SP**]. In this case,
(2.9d) becomes an algebraic relation among the n_i.

We may eliminate E from (2.9a) and (2.9b) to get

$$\frac{1}{n_1}\frac{dn_1}{dx} - \frac{1}{n_2}\frac{dn_2}{dx} = b_2 - 1.$$

Choosing the origin to be at the point where $n_1 = n_2$, we obtain the relation

(2.12) $$n_2 = n_1 e^{(1-b_2)x},$$

(2.13) $$E = 1 + \frac{1}{n_1}\frac{dn_1}{dx}.$$

From (2.9d), (2.12),

(2.14) $$n_3 = n_1(1 + e^{(1-b_2)x}).$$

Substitution of (2.13) and (2.14) into (2.9c) easily yields a linear differential
equation for n_1 alone, of the form

$$\dot{n}_1 = -A(x)(n_1 - B(x)),$$

where $A(x) = \kappa + \frac{1}{2}\alpha e^{\alpha x}(1 + e^{\alpha x})^{-1} > 0$, $B(x) = 2\theta\kappa/(2\kappa + (2\kappa + \alpha)e^{\alpha x})$, $\kappa = \frac{1}{2}(1 + b_3)$, $\alpha = 1 - b_2$. The only bounded solution of this equation is

(2.15) $$n_1(x) = \kappa\theta e^{-\kappa x}(1 + e^{\alpha x})^{-1/2}\int_{-\infty}^{x} e^{\kappa t}(1 + e^{\alpha t})^{-1/2}\,dt.$$

Similar explicit formulas for n_2, n_3, and E can be obtained from this and (2.12–14).

It can be checked that n_1, given by (2.15), satisfies the correct boundary condi-
tions $n_1(-\infty) = \theta$, $n_1(+\infty) = 0$. Moreover, $n_1(x)$ is a strictly decreasing function.
To see this, note that $B(x)$ is decreasing and satisfies the same boundary conditions
as n_1. If, for some finite $x = x_0$ (which we take to be the leftmost such point) it
were true that $n_1(x_0) = B(x_0)$, then $\dot{n}_1 = 0$ there by the differential equation,
which means that n_1 would change from being less than B (hence $\dot{n} > 0$) to being
greater than B ($\dot{n} < 0$) at that point. To the left, n_1 would remain strictly increas-
ing, and could not attain its limit at $-\infty$. This contradiction shows that n_1 must
always remain either greater or less than B. Hence n_1 is strictly monotone, and
hence decreasing because of the boundary conditions.

Similar differential equations can be derived for n_2 and n_3, showing that they
are monotone increasing. All functions n_i and E satisfy the correct boundary
conditions. This behavior is shown in Figure 3.

In short, for $\varepsilon = 0$, we have, by explicit construction, the existence and unique-
ness (modulo shifts in x) of a solution, and its monotonicity.

3. Existence and qualitative properties for $\varepsilon > 0$. For positive constants
b_i and θ, and nonnegative ε, we consider the system (2.9), (2.10).

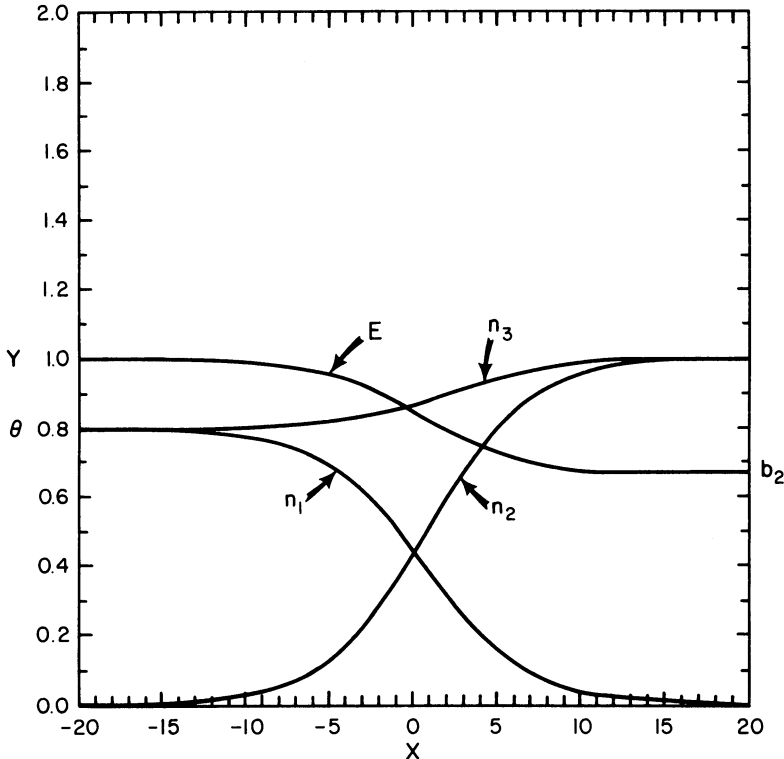From (2.8), (2.11), the parameters are subject to the restrictions

(3.1) $$b_2 < 1;$$

TRAVELING WAVES IN ITP. $\varepsilon = 0$

FIGURE 3

(3.2) $$\theta(1 + b_3) = (b_2 + b_3).$$

The problem may be reformulated by defining the vectors $z = (1, 1, -1)$, $b = (1, b_2, -b_3)$, and the quantity $N \equiv n_1 + n_2 + n_3$.

Then (2.9d) may be written as $\varepsilon \dot{E} = z \cdot n$, the latter symbol denoting scalar product. Taking the scalar product of (2.9a–c) with z now yields

$$\varepsilon \ddot{E} = NE - b \cdot n - \theta(1 + b_3).$$

Or defining

(3.3) $$F(n) \equiv \frac{b \cdot n + \theta(1 + b_3)}{N},$$

we have

(3.4) $$\varepsilon \ddot{E} - N(E - F(n)) = 0.$$

Equation (3.4) will be taken as replacement for (2.9d). Our basis problem now is

P: to find a solution (n, E) of (2.9a–c), (3.4) (subject to (3.1), (3.2)), satisfying (2.10).

THEOREM 3.1. *For every small enough $\varepsilon \geq 0$,[1] there exists a solution of* **P** *satisfying the monotonicity properties* **M**: n_1 *and* E *are strictly decreasing in* x; n_2 *and* n_3 *are strictly increasing.*

In preparation for the proof, we first examine an analogous problem $\mathbf{P}_K$ on a finite interval, defined as follows.

$\mathbf{P}_K$: to find a solution of (2.9a–c), (3.4) defined on the interval $(-K, K) \equiv I_K$, satisfying the boundary conditions

(3.5a) $$n_1(-K) = \theta,$$

(3.5b) $$n_2(+K) = 1,$$

(3.5c) $$n_3(-K) = \theta,$$

(3.6) $$\dot{E}(+K) = \dot{E}(-K) = 0.$$

This problem, as shown below, can be interpreted as a fixed point problem on the set

$$U_K \equiv \{E \in C^1(\bar{I}_K): \ \|E\|_1 \leq r; \ E \text{ is monotone nonincreasing}$$
$$\text{on } I_K; \text{ and } E(x) \in [b_2, 1] \text{ for } x \in I_K\}.$$

Here r is a positive number which will be selected later, and $\|E\|_1$ denotes the norm in C^1. (Later we shall use $\|E\|_m$ and $\|n\|_m$ to denote norms in C^m.) For future reference, we define the related set $U_K^0 \subset U_K$ which differs from the latter only in that the inequality and the monotonicity are both supposed to be strict, and $E(x) \in (b_2, 1)$ for $x \in I_K$.

A transformation T_K on U_K is defined as follows.

Given $E \in U_K$, first define the function $n(x) = (n_1(x), n_2(x), n_3(x))$ as the unique solution of (2.9a–c) under boundary conditions (3.5). Then let E' be defined as the solution of (3.4), (3.6) (see Lemma 3.2 below). In the case $\varepsilon = 0$, the boundary conditions (3.6) are omitted. We define $E' \equiv T_K E$.

LEMMA 3.2. *Given* $n \in C^2(\bar{I}_K)$ *with* $N > 0$, *for* ε *positive there exists a unique solution* E *of* (3.4), (3.6). *Moreover,*

(3.7) $$|E(x) - F(n(x))| \leq C_1 \varepsilon$$

for x bounded away from $\pm K$, and for a constant C_1 depending only on that bound and on n.

PROOF. The function

$$\bar{E} \equiv f(x) + M\varepsilon - \varepsilon^{1/2}\eta_1 \exp[-a\varepsilon^{-1/2}(x + K)] + \varepsilon^{1/2}\eta_2 \exp[a\varepsilon^{-1/2}(x - K)],$$

where $f(x) \equiv F(n(x))$, will be an upper solution, provided that the constants M, a, η_1, and η_2 are chosen so that $M \geq \sup(|f(x)|/N(x))$, $0 < a < N(x)$, $\eta_1 = \dot{f}(-K)a^{-1}$, and similarly for η_2 (this choice of η_i will guarantee that $\bar{E}$ satisfies the right boundary conditions). There is a similar lower solution of the form $f - M\varepsilon +$ the same kind of boundary corrections. Therefore there exists

---
[1]We have learned that Professors Li Zhen-yuan and Ye Qi-xiao have succeeded in removing the restriction on ε.

an exact solution between the two. This proves existence, together with (3.7). Uniqueness follows from the maximum principle.

It is clear that fixed points of T_K will be solutions of $\mathbf{P}_K$. For convenience, the subscript K will be dropped from T_K.

PROPOSITION 3.3. *For $E \in U_K$, the functions $n_i(x)$ satisfying (2.9a–c), (3.5) have the following properties for $x \in I_K$:*
1. $n(x) > 0$.
2. $(-\dot{n}_1, \dot{n}_2, \dot{n}_3) \geq 0$.
3. $n_1 \in (0, \theta]$, $n_2 \in (0, 1]$, $n_3 \in [\theta, 1]$.

PROOF. Part 1 for the first two components n_1 and n_2 follows directly from equations (2.9a,b) and the positivity of the initial data. These same equations and the fact that $E(x) \in [b_2, 1]$ also imply Part 2 for those components. Now consider n_3. It is easily seen from differentiating (2.9c) and applying the initial condition that as long as $n_3 \geq 0$,

$$\ddot{n}_3 = -\dot{n}_3(E + b_3) - n_3\dot{E} \geq -\dot{n}_3(E + b_3).$$

This and the fact that $\dot{n}_3(-K) = -\theta(E + b_3) + \theta(1 + b_3) \geq 0$ now imply that $\dot{n}_3 \geq 0$ for $x \in I_K$ such that $n_3(x) \geq 0$. But in order for n_3 to vanish, it must be true that $\dot{n}_3 < 0$ at a previous value of x. This is impossible, and the proof of Part 2 is finished.

In Part 3, the statements regarding n_1 and n_2 follow from Parts 1 and 2, as does the fact that $n_3 \geq \theta$. Now if n_3 ever surpassed the value 1 in I_K, it would follow from (2.9c) and the fact that $E \geq b_2$ that $\dot{n}_3 \leq (-n_3 + 1)(b_2 + b_3) < 0$, which contradicts Part 2. This completes the proof.

PROPOSITION 3.4. *Let n be as in Proposition 3.3, and $F(n)$ be defined by (3.3). Then the function $f(x) \equiv F(n(x))$ satisfies*
1. *$\dot{f} \leq 0$ when $F \in [b_2, 1]$.*
2. *$f(x) \in (b_2, 1)$ for all $x \in I_K$.*
3. *For $F \in (b_2, 1)$, $F(n)$ is increasing in n_1 and decreasing in n_2 and n_3.*
4. *There is a number $F_0 \in (b_2, 1)$, independent of K, such that $f(x)$ assumes the value F_0 for some $x_K \in I_K$.*

PROOF. Taking differentials in (3.3) shows that

(3.8) $N dF = (1 - F) dn_1 - (F - b_2) dn_2 - (F + b_3) dn_3$.

In particular, we have that

(3.9) $N\dot{f} = (1 - F)\dot{n}_1 - (F - b_2)\dot{n}_2 - (F + b_3)\dot{n}_3$.

Part 1 now follows from (3.9) and Proposition 3.3.2.

Next, we calculate $f(-K) = F(\theta, n_2, \theta) = (2\theta + b_2 n_2)/(2\theta + n_2)$, and it is easily checked that this latter number lies in the interval $(b_2, 1)$. This, together with Part 1, shows that $f(x) < 1$ for all x.

Moreover, it cannot attain the value b_2; for if it did so at some point $x_0 \in I_K$ (which we take to be the first such point), then necessarily $f(x_0) = b_2$; $\dot{f}(x_0) \leq 0$. But these two inequalities together with (3.9) and Proposition 3.3.2 imply $\dot{n}_1(x_0) = 0$. From (2.9a), this could only happen if $E(x_0) = 1$, which would imply, since E is

nonincreasing, that $E(x) = 1$ for all $x \in (-K, x_0)$. This and (2.9a) would further imply that n_1 is constant, therefore equal to θ, on $(-K, x_0)$. Similarly, (2.9c) with Proposition 3.3.2 and 3.3.3 imply that n_3 is also equal to θ on that interval. But then $f(x_0) = F(\theta, n_2, \theta)$, which we have already seen to be larger than b_2. This is a contradiction. This establishes Part 2.

Part 3 follows from (3.8). For Part 4, consider that Proposition 3.3.2 implies that $n_1 \leq \theta$, $n_2 \leq 1$, $n_3 \geq \theta$. From this, the boundary conditions, and Part 3, we have

$$f(-K) = F(\theta, n_2(-K), \theta) \geq F(\theta, 1, \theta) \geq F(n_1(K), 1, n_3(K)) = f(K).$$

Now let $F_0 \equiv F(\theta, 1, \theta) \equiv (2\theta + b_2)/(2\theta + 1)$. It is clearly in $(b_2, 1)$, and since the boundary values of f lie on either side of it, it is attained by f at some point x_0. This completes the proof.

PROPOSITION 3.5. *Let f be as in the previous proposition. Then $\dot{f} < 0$ for all $x \in I_K$.*

PROOF. From (3.9) and Propositions 3.4.2, 3.3.2, it is clear that $\dot{f} = 0$ at any point implies $\dot{n} = 0$, which from (2.9a,b) would imply that $E = 1$ and $E = b$ at that point, which is impossible. The conclusion now follows from Proposition 3.3.1.

PROPOSITION 3.6. *For all $E \in U_K$, $TE = E'$ lies in a ball in $C^1(I_K)$ of radius r_0 independent of r, $\varepsilon \geq 0$, and K. Moreover, the second derivative of E' is bounded independently of ε and K (but may depend on r).*

PROOF. It is easily seen directly from the differential equations (2.9a–c) and the uniform boundedness of n (Proposition 3.3.3) that $|\dot{n}|$, hence $|d/dx\, F(n(x))| = |\dot{f}|$, is bounded from above independently of E, r, ε, and K, as long as $E \in U_K$. In the case $\varepsilon = 0$, $TE = F(n(x))$ and the first statement follows. In the case $\varepsilon > 0$, E' satisfies

$$\varepsilon \overset{3}{\dot{E}}{}' = \dot{N}(E' - f) + N(\dot{E}' - \dot{f}),$$

where the notation on the left side denotes the third derivative. At a point where E' has a minimum on I_K (recall it vanishes at the endpoints), we have $\overset{3}{\dot{E}}{}' \geq 0$; from this and the fact that $\dot{f} \leq 0$, we obtain a uniform lower bound on $\dot{E}'$ at that minimum, hence everywhere in I_K. The same sort of argument establishes such an upper bound (but see below, where it is shown that 0 is an upper bound). The first statement of the proposition follows.

To establish the second, let $v(x) = E'(x) - f(x)$. It satisfies

(3.10) $$\varepsilon \ddot{v} - Nv = -\varepsilon \ddot{f}.$$

Now by differentiating (2.9a–c), we see that $\ddot{n}$ and hence $\ddot{f}$ are bounded in terms of $\|E\|_1 \leq r$. If $\varepsilon = 0$, we are done because $E' = f$. Suppose $\varepsilon > 0$. The right side of (3.10) is bounded in absolute value by εM for some M independent of ε and K. The function $v(x)$ cannot have a maximum on the boundary of I_K unless $\dot{v} = 0$ there; this is simply because at the left boundary (say), $\dot{E}' = 0$, so $\dot{v} = -\dot{f} \geq 0$. And since $\dot{v} \leq 0$ at such a boundary maximum, we must have that $\dot{v} = 0$. It follows that at the point where v attains a maximum, whether it be in the interior or at the boundary, necessarily $\ddot{v} \leq 0$, hence from (3.10), $v \leq \varepsilon M_1$. A similar

argument at the minimum of v establishes that in fact $|v| \leq \varepsilon M_2$ for all x and for some constant M_2, independent of the usual things. This and (3.10) again in turn implies that $\ddot{v}$, hence $\ddot{E}'$, is subject to a uniform bound, which completes the proof of the proposition.

At this point we specify the radius r appearing in the definition of U_K to be some number larger than r_0.

LEMMA 3.7. *T maps U_K into U_K^0.*

PROOF. The function $E' \equiv TE$ satisfies the boundary value problem

$$(3.11) \qquad \varepsilon \ddot{E}' - N(E' - f(x)) = 0,$$

$$(3.12) \qquad \dot{E}'(\pm K) = 0,$$

where $f(x)$ is the function of Propositions 3.4 and 3.5, and N is defined in terms of n following (3.2). Note that $N > \theta$ by Proposition 3.3.3. At a local maximum (boundary points are not excluded—see (3.12)), it must be true that $\ddot{E}' \leq 0$, hence from (3.11) that $E' \leq f$. Similarly at a local minimum, $E' \geq f$. Suppose E' had two successive local extrema, at x_1 and x_2, the first a minimum and the second a maximum. We would have $f(x_1) \leq E'(x_1) \leq E'(x_2) \leq f(x_2)$. This contradicts the fact that f is decreasing (Proposition 3.5). It follows that there is no interval on which E' is strictly increasing.

On the other hand if $\dot{E}' = 0$ at some point x_0 in the interior, then the nonincreasing property implies that $\ddot{E}' \geq 0$ at points to the left of and arbitrarily close to x_0, and the opposite immediately to the right. From (3.11), this implies that $E' \geq f$ at those points to the left, which together with Proposition 3.5 contradicts the hypothesis that $\dot{E}'(x_0) = 0$. Therefore E' is strictly decreasing on I_K.

Since E' has a local maximum at $-K$ and a local minimum at K, it follows that $E'(-K) \leq f(-K) < 1$ with a similar inequality at K: $E'(K) \geq f(K) > b_2$ by Proposition 3.4.2.

We have established that TE has the proper range and monotonicity. The fact that it lies in the ball of radius r in C^1 follows from Proposition 3.6 and the requirement that $r > r_0$. This completes the proof.

LEMMA 3.8. *Problem $\mathbf{P}_K$ has a solution (n, E) with E in U_K^0 for every positive K and nonnegative ε.*

PROOF. The set U_K is closed and convex in the space $C^1(I_K)$. The above lemma shows that T maps it into itself. The last statement of Proposition 3.6 shows that T is compact, and the existence of a fixed point E of T follows from Schauder's theorem. It is in U_K^0 by Lemma 3.7. The function $E(x)$, together with the associated function $n(x)$, is the desired solution of $\mathbf{P}_K$.

PROPOSITION 3.9. *The solution of $\mathbf{P}_K$ satisfies*

$$(3.13) \qquad n_1 + n_2 - n_3 = \varepsilon \dot{E} + n_2(-K).$$

PROOF. It follows from (2.9a–c) that

$$\dot{n}_1 + \dot{n}_2 - \dot{n}_3 = N(E - F(n)).$$

The proposition follows by integrating (3.4) with this representation made, and applying the known boundary values at $x = -K$.

We are now ready to complete the proof of Theorem 3.1 by passing to the limit as $K \to \infty$. The first step is to shift the independent variable in the solutions of $\mathbf{P}_K$, proved to exist in Lemma 3.8, so that always $F(n(0)) = F_0$ (this is possible by Proposition 3.4.4). The interval on which the solution is defined will no longer be $(-K, K)$, but will remain of length $2K$ and contain the origin. Nevertheless, we still denote it by I_K. Let K assume only integer values, and let it approach ∞. The uniform bounds on TE given in Proposition 3.6 hold also for the solutions, which we now call E_K, and by Arzela's theorem the following result is obtained.

LEMMA 3.10. *There is an infinite interval I^* and a subsequence $\{K_j\}$ along which the E_K approach a limit E^*, defined on I^*. The limit $E^*(x)$, together with its associated vector function $n^*(x)$, satisfies (2.9a–c), (3.4).*

PROOF. A subsequence can be chosen so that the set $I^* \equiv \{x \colon x \in I_{K_j}$ for all $j > j_0$ for some $j_0(x)\}$ is infinite and contains the origin. This follows merely from the facts that the length of the intervals approaches ∞, and that they all contain the origin. As indicated above, by Arzela's theorem a further subsequence exists along which the E_K approach some E^*. The approach is uniform together with first derivatives on any finite subinterval of I^*, so the limit functions E^*, n^* satisfy the same differential equations. This completes the proof.

Since I^* is infinite and contains the origin, it either
(a) contains the entire semiaxis $(-\infty, 0]$, or
(b) contains the entire semiaxis $[0, \infty)$.
We shall show that both occur.

If case (a) holds but not (b), then the right-hand endpoints of the intervals I_{K_j} of the subsequence are bounded and, possibly by taking a further subsequence, we may assume that they converge to a limit X_+ which will be the right-hand endpoint of I^*. By continuity, we shall have

$$(3.14) \qquad\qquad n_2^*(X_+) = 1.$$

PROPOSITION 3.11. *If case (a) holds and ε is small enough, then (n^*, E^*) satisfies boundary condition (2.10a) and case (b) must also hold.*

PROOF. The limit function n^* continues to satisfy the conclusions of Proposition 3.3, possibly with strict inequalities becoming nonstrict and intervals being closed. The monotonicity and boundedness of these functions imply they have limits $n_i^*(-\infty) \equiv n_i^-$. The same holds for E^*. Set $F_1 \equiv F_0 - C_1\varepsilon$, the latter being the constant in (3.7). Then $E^* > F_1$ on $(-\infty, 0)$ and $F_1 > b_2$ for small ε. This together with the fact that $n_2^* \le 1$ indicates that $n_2^*(x) \le \exp[(F_1 - b_2)x]$ for $x \le 0$. Therefore

$$(3.15) \qquad\qquad n_2^- = 0.$$

Passing to the limit in (3.13) yields

$$(3.16) \qquad\qquad n_1^* + n_2^* - n_3^* - \varepsilon E^* = 0,$$

hence

$$(3.17) \qquad\qquad n_1^- + n_2^- - n_3^- = 0.$$

Combining (3.17), (3.15) with the facts that $n_1^- \leq \theta$ and $n_3^- \geq \theta$ (Proposition 3.3.3) tells us that

$$(3.18) \qquad\qquad n^*(-\infty) = (\theta, 0, \theta).$$

Now suppose that case (b) does not hold. By the above, I^* has a finite endpoint X_+. Furthermore, (3.14) together with (3.16) and the facts that $n_1^*(X_+) \geq 0$, $n_3^*(X_+) \leq 1$ and $\dot{E}^*(X_+) = 0$ yields $n_1^*(X_+) = 0$ and $n_3^*(X_+) = 1$. The former would imply that $n_1^* \equiv 0$, which we know to be incorrect because $n_1^*(-\infty) = \theta$. Therefore the assumption that case (b) does not hold is wrong, and case (b) is established. This completes the proof.

PROPOSITION 3.12. *If case* (b) *holds, then* (n^*, E^*) *satisfies boundary condition* (2.10b), *and case* (a) *holds.*

PROOF. Assume case (b). Then in place of (3.15), we have the analogous result

$$(3.19) \qquad\qquad n_1^+ = 0.$$

Also

$$(3.20a) \qquad\qquad 0 = \dot{n}_2^*(\infty) = n_2^+(E^+ - b_2),$$

$$(3.20b) \qquad\qquad 0 = \dot{n}_3^*(\infty) = -n_3^+(E^+ + b_3) + (b_2 + b_3).$$

Suppose that $n_2^+ \neq 0$. Then by (3.20a), $E^+ = b_2$, and by (3.20b), $n_3^+ = 1$. This in fact must be the case, because if $n_2^+ = 0$, the fact (analogous to (3.17)) that $n_1^+ + n_2^+ - n_3^+ = 0$ and (3.19) would imply $n_3^+ = 0$, which contradicts the fact that $n_3^* \geq \theta$. This establishes (2.10b).

Now suppose that case (a) did not hold. Then I^* would have a finite (negative) left endpoint X_-, where $n_1 = n_3 = \theta$ and $\varepsilon \dot{E} = 0$. As before, (3.16) would imply that $n_2^*(X_-) = 0$, which by (2.9b) further implies $n_2^* \equiv 0$, which cannot happen. Therefore cases (a) and (b) hold.

PROOF OF THEOREM 3.1. The above propositions show that for small enough ε, the limits E^*, n^* satisfy (2.9a–c), (3.4), and the boundary conditions (2.10). Moreover, we have shown (3.16) that (2.9d) holds. The only thing left to show is the monotonicity property M, i.e. that the inequality in Proposition 3.3.2 for n^* is strict, and that E^* is strictly decreasing. This is left to the reader.

In preparation for the next section, we shall now show that despite the apparently singular type of appearance of the small parameter ε in (2.9d) and (3.4), all derivatives of the solution are bounded independently of ε. The norm $||E||_m$ is defined as the supremum of the absolute value of all derivatives of E up to order m, as functions of x, for all real values of x.

LEMMA 3.13. *For each positive integer* m *there is a constant* C_m *independent of* ε *such that*

$$(3.21) \qquad\qquad ||E||_m + ||n||_m \leq C_m.$$

Moreover, $E(x) = F(n(x)) + \varepsilon G(x)$, *where*

$$(3.22) \qquad\qquad ||G||_m < C_m.$$

PROOF. First of all, (3.22) follows from (3.21) and (3.4); in fact, $G(x) = E(x)/N(x)$. The estimate (3.21) is clearly true for $m = 0$. We first derive it for $m = 2$. Let

$$A \equiv ||E - F(n)||_0; \qquad B \equiv ||E||_1.$$

In the following the same symbol C will denote various constants independent of ε. From the proof of Lemma 3.2 and by differentiating (2.9), it is clear that $A \leq C\varepsilon||n||_2$ and $||n||_2 \leq C(1 + B)$, hence

$$(3.23) \qquad\qquad A \leq C\varepsilon(1 + B).$$

Also from (3.4) and an interpolation inequality,

$$(3.24) \qquad\qquad ||\ddot{E}||_0 \leq C\varepsilon^{-1}A$$

and

$$(3.25) \qquad\qquad B \leq C||\ddot{E}||_0^{1/2} \leq C\varepsilon^{-1/2}A^{1/2}.$$

The solution of inequalities (3.23), (3.25) is

$$A \leq C\varepsilon, \qquad B \leq C.$$

Applying this result to (3.24) and the above estimate for $||n||_2$ gives the conclusion of the lemma in the case $m = 2$. The estimates for higher m can be obtained along similar veins by differentiating (2.9), (3.4). This completes the proof.

LEMMA 3.14. *The solution (n, E) established in Theorem 3.1 approaches its limits as $x \to \pm\infty$ exponentially, at a rate which is bounded below independently of ε. More specifically,*

$$(3.26) \qquad |n(x) - n(\infty)| + |E(x) - E(\infty)| \leq Me^{-ax},$$

with the analogous inequality holding for $-\infty$, where M and a are positive constants independent of ε.

PROOF. The assertion in the case $\varepsilon = 0$ follows from (2.15), so we assume $\varepsilon > 0$. Linearization of (2.9a–d) at the boundary values at $x = \pm\infty$ yields matrices with all real distinct eigenvalues bounded away from 0 independently of ε. To establish (3.26), consider for example the first component n_1, as $x \to +\infty$. It turns out that there are two negative eigenvalues $-\lambda_1$ and $-\lambda_2, \lambda_2 > \lambda_1$, bounded away from zero (but not bounded) independently of ε. The solution lies on the stable manifold of the critical point at $+\infty$, so

$$n_1(x) = p_1 e^{-\lambda_1 x}(1 + o(1)) + p_2 e^{-\lambda_2 x}(1 + o(1))$$

as $x \to \infty$. Here the p_i may depend on ε. In view of (3.21), this function and its derivative

$$\dot{n}_1(x) = -\lambda_1 p_1 e^{-\lambda_1 x}(1 + o(1)) - \lambda_2 p_2 e^{-\lambda_2 x}(1 + o(1))$$

are both bounded independently of x and ε. Since the λ's are distinct, it follows that each individual term $p_i e^{-\lambda_i x}(1 + o(1))$ is so bounded. Let x_0 be a number large enough so that the terms $o(1)$ here are less than $\frac{1}{2}$. Then for some constant M independent of ε, we have $|p_i e^{-\lambda_i x_0}| < M$, so that

$$|p_i e^{-\lambda_i x}(1 + o(1))| \leq 2Me^{-\lambda_1(x-x_0)},$$

and the inequality (3.26) for n_1 alone follows. The other components, and the approach to $-\infty$, are handled in the same way. This completes the proof.

4. Approximation of the solution for small ε. Having proved the existence of solutions of (2.9), (2.10) in §3 for all small $\varepsilon \geq 0$, we now wish to show their uniqueness and to show that the solution for $\varepsilon > 0$ may be approximated, within an error $O(\varepsilon)$, by the one for $\varepsilon = 0$. The latter will be done first.

For purposes of Theorem 4.1, (n, E) will be a solution with some fixed $\varepsilon > 0$, and (n^0, E^0) will denote a solution with $\varepsilon = 0$. In both cases, the origin will be chosen at the cross-over point for the graphs of n_1 and n_2,

$$(4.1) \qquad n_1(0) - n_2(0) = n_1^0(0) - n_2^0(0) = 0.$$

(In view of the construction in §2, (n^0, E^0) is uniquely determined.)

Rewrite (2.9a–c) in the form

$$(4.2) \qquad \dot{n} = \tilde{g}(n, E).$$

Because of (3.4), $E^0 = F(n^0)$ and

$$(4.3) \qquad \dot{n}^0 = \tilde{g}(n^0, F(n^0)) \equiv g(n^0).$$

Moreover, from Lemma 3.13 we have $E(x) = F(n(x)) + \varepsilon G(x)$, where G is smooth, with all derivatives bounded independently of ε. Hence

$$(4.4) \qquad \dot{n} = \tilde{g}(n, F(n) + \varepsilon G) \equiv g(n(x)) + \varepsilon h(x),$$

where h also has all derivatives bounded independently of ε. Finally, Lemma 3.14 yields that h decays as $x \to \pm\infty$ and obeys an inequality like (3.26).

Let $u(x) \equiv n(x) - n^0(x)$. Subtracting (4.3) from (4.4), we obtain

$$(4.5) \qquad \dot{u} = g'(n^0 + \tau u)u + \varepsilon h, \qquad 0 < \tau < 1,$$

where g' denotes the Jacobian matrix. We define $A_1(x) \equiv g'(n^0(x) + \tau(x)u(x))$. Then (4.5) becomes

$$(4.6) \qquad \dot{u} = A_1(x)u + \varepsilon h(x).$$

THEOREM 4.1. *There is a constant C, independent of ε, such that $\|u\|_0 \leq C\varepsilon$.*

The proof relies on several lemmas.

LEMMA 4.2. *The matrix $A_1(x)$ approaches the limit A_- (which is independent of ε) exponentially and uniformly in ε, as $x \to -\infty$. The matrix A_- has one negative, one positive, and one zero eigenvalue. There is a set of three linearly independent solutions of*

$$(4.7) \qquad \dot{w} = A_1(x)w$$

with the property that only one of the three (call it $w^+(x)$) decays to zero exponentially as $x \to -\infty$. Moreover

$$(4.8) \qquad w_1^+(0) \neq w_2^+(0).$$

PROOF. Lemma 3.14 shows that $n(x)$ and $n^0(x)$ approach their limits exponentially, so A_1 does as well. The three eigenvalues of A_- are found by direct computation, use being made of (3.9). (For future reference, the positive one turns out to be $1 - b_2$.) The behavior of solutions of (4.7) as $x \to -\infty$ is the same as that of the solutions of $\dot{w} = A_- w$ (see [CL, Theorem 3.8.1]), which is the behavior described.

It only remains to prove (4.8). An explicit expression for A_1 will be needed. We write $\hat{n}(x) = n^0(x) + \tau(x)u(x)$, $\hat{F}(x) = F(\hat{n}(x))$, and $\mu_i(x) = \hat{n}_i/(\hat{n}_1 + \hat{n}_2 + \hat{n}_3)$. Then by definition, use being made of (3.8),

$$A_1(x)u = g'(\hat{n}(x))u = \tilde{g}_n(\hat{n}, \hat{F})u + \tilde{g}_E(\hat{n}, \hat{F})F'(\hat{n})u$$

(4.9)
$$= \begin{bmatrix} -(1 - \hat{F})(1 - \mu_1) & -\mu_1(\hat{F} - b_2) & -\mu_1(\hat{F} + b_3) \\ \mu_2(1 - \hat{F}) & (1 - \mu_2)(\hat{F} - b_2) & -\mu_2(\hat{F} + b_3) \\ -\mu_3(1 - \hat{F}) & \mu_3(\hat{F} - b_2) & -(1 - \mu_3)(\hat{F} + b_3) \end{bmatrix}.$$

The relation (2.12) is derived from (2.9a,b) alone, and holds whether or not $\varepsilon = 0$. Therefore again setting $\alpha = 1 - b_2$, we have

$$n_2 = n_1 e^{\alpha x}, \quad n_2^0 = n_1^0 e^{\alpha x},$$

and so

(4.10)
$$u_2 = u_1 e^{\alpha x}, \quad \mu_2 = \mu_1 e^{\alpha x}, \quad \hat{n}_2 = \hat{n}_1 e^{\alpha x}.$$

From (4.9), (4.10), we may write the first two equations of the system $\dot{w} = A_1 w$ in the form

(4.11)
$$\dot{w}_1 = -(1 - \hat{F})(1 - \mu_1)w_1 - \mu_1(\hat{F} - b_2)w_2 - \mu_1(\hat{F} + b_3)w_3,$$

(4.12)
$$\dot{w}_2 = \mu_1(1 - \hat{F})w_1 e^{\alpha x} + (1 - \mu_1 e^{\alpha x})(\hat{F} - b_2)w_2 - \mu_1(\hat{F} + b_3)w_3 e^{\alpha x}.$$

We now define the function $v(x)$ by

$$w_2(x) = w_1(x)e^{\alpha x} + v(x).$$

Make this substitution for w_2 in (4.11), (4.12), multiply (4.11) by $e^{\alpha x}$, and subtract the two equations. There results a simple equation for v:

(4.13)
$$\dot{v} = (\hat{F} - b_2)v.$$

Hence either $v(x) \equiv 0$ or v never vanishes. We shall show that for the particular solution $w = w^+$, the second is the case.

By [**CL**, Theorem 3.8.1], the solution $w^+(x)$ has the following behavior for large negative x (we omit superscripts $^+$ for simplicity):

$$\lim_{x \to -\infty} w(x)e^{-\alpha x} = p,$$

where p is an eigenvector of A_- associated with the positive eigenvalue $\lambda = \alpha$. Again, direct calculation shows that $p = (p_1, p_2, p_3)$ where $p_1 < 0$, $p_2 > 0$, and $p_3 > 0$ (or the reverse).

It follows that for large negative x, $w_1(x)$ and $w_2(x)$ have opposite signs, and so $v = w_2 - w_1 e^{\alpha x} \neq 0$. We conclude that v never vanishes. In particular, $v(0) = w_2(0) - w_1(0) \neq 0$. This completes the proof.

It will be shown later that w^+ is in fact bounded for all x. Moreover, it should be noted that all the relevant properties of the matrix A_1 used in deriving this result as well as the above are completely independent of ε, and so we may choose w^+ itself to be bounded independently of ε, and simultaneously to have $v(0)$ bounded away from 0 independently of ε.

LEMMA 4.3. *Equation (4.6) has a unique solution decaying exponentially as* $x \to -\infty$ *and satisfying* $u_1(0) = u_2(0)$.

PROOF. By construction, the function $u = n - n^0$ satisfies (4.6) and those conditions, so existence is trivial. Now every other solution decaying at $-\infty$ must be of the form $u^* = u + cw^+$, where c is an arbitrary constant and w^+ is the function from Lemma 4.2. But we are requiring

$$0 = u_1^*(0) - u_2^*(0) = u_1(0) - u_2(0) + c(w_1^+(0) - w_2^+(0)).$$

Now (4.8) implies that $c = 0$, completing the uniqueness proof.

PROOF OF THEOREM 4.1. One solution $\tilde{u}(x)$ of (4.6), decaying as $x \to -\infty$, may be constructed as the solution of an integral equation, as in [**CL**, Theorem 3.8.1]. For that purpose we write $A_1(x) = A_- + R(x)$, with R, like h, decaying exponentially as $x \to -\infty$. Specifically,

$$|R(x)| + |h(x)| \le Me^{\alpha x},$$

$\alpha = 1 - b_2$ being the positive eigenvalue of A_-, and M being independent of ε. With no loss of generality, it will be assumed that A_- is diagonal: $A_- = \mathrm{diag}(\alpha, 0, -\beta)$, these being the three eigenvalues. The solution will be constructed in the form of a solution, defined for $x \le -X < 0$, X being a suitably large constant, of an integral equation

(4.14) $$\tilde{u}(x) = \varepsilon J h(x) + J R(x)\tilde{u}(x).$$

Here the operator J is defined componentwise as follows:

$$(Jh(x))_1 \equiv \int_{-X}^{x} e^{\alpha(x-s)} h_1(s)\, ds,$$

$$(Jh(x))_2 \equiv \int_{-\infty}^{x} h_2(s)\, ds,$$

$$(Jh(x))_3 \equiv \int_{-\infty}^{x} e^{-\beta(x-s)} h_3(s)\, ds.$$

It is easily verified that a solution of (4.14) will satisfy (4.6) and decay exponentially as $x \to -\infty$.

As in [**CL**], the solution is constructed by iteration: $\tilde{u}^1 = \varepsilon J h$, $\tilde{u}^{j+1} = \varepsilon J h + J R \tilde{u}^j$. The iteration will converge if X is sufficiently large, and in fact direct estimates show that for any fixed δ, $0 < \delta < \alpha$, $\{\tilde{u}^j\}$ is a Cauchy sequence with respect to the norm

$$\|\tilde{u}^\wedge\| \equiv \|\tilde{u}e^{-\delta x}\|_0^{(-\infty, X)}$$

(this last superscript means that the function is to be restricted to the interval $(-\infty, X)$). The operator J is bounded in this norm, and therefore the solution will satisfy

(4.15) $$\|\tilde{u}^\wedge\| \le C\varepsilon\|h^\wedge\|,$$

C of course being independent of ε.

This defines the desired solution $\tilde{u}$ on the interval $(-\infty, -X)$. It can be extended uniquely to the entire real line. It is now appropriate to consider the possible

behavior of solutions of (4.6) for large positive x. The matrix $A_1(+\infty) = A_+$ has two negative ($\lambda_1 = b_2 - 1; \lambda_2 = -(b_2 + b_3)/2$) and one zero eigenvalue ($\lambda_3 = 0$).

The solution $\tilde{u}$ may be represented in the form

$$\tilde{u}(x) = \varepsilon J^+ h(x) + J^+ R^+(x)\tilde{u}(x) + U_0(x)$$

for $x > X > 0$, X sufficiently large, where

$$R^+(x) \equiv A(x) - A_+,$$

which decays exponentially as $x \to \infty$,

$$J^+ h(x) \equiv \int_X^x e^{A_+(x-s)} h(s) \, ds,$$

and

$$U_0(x) = e^{A_+(x-X)}\tilde{u}(X).$$

With the same small enough $\delta > 0$ as before, we let

$$||h^\wedge||_+ \equiv \sup\{|h(x)e^{\delta x}| : x > X\}.$$

It can be verified that $||J^+ R(x)\tilde{u}||_0^{(X,\infty)} < \frac{1}{2}||\tilde{u}||_0^{(X,\infty)}$ if X is a suitably large constant. Therefore since

$$||\tilde{u}(x)||_0^{(X,\infty)} \le \varepsilon||J^+ h(x)||_0^{(X,\infty)} + ||J^+ R(x)\tilde{u}||_0^{(X,\infty)} + ||U_0(x)||_0^{(X,\infty)}$$

we have

$$||\tilde{u}(x)||_0^{(X,\infty)} \le 2\varepsilon||J^+ h(x)||_0^{(X,\infty)} + 2|\tilde{u}(X)|.$$

It can be shown that

$$|J^+ h(x)| \le C||h^\wedge||_+,$$

with C independent of ε. Therefore

(4.16) $$||\tilde{u}(x)||_0^{(X,\infty)} \le 2C\varepsilon||h^\wedge||_+ + 2|\tilde{u}(X)|.$$

In view of the continuity of solutions of initial value problems with respect to initial data, we have

$$|\tilde{u}(x)| \le C|\tilde{u}(-X)| + C\varepsilon||h||_0, \qquad x \in [-X, X],$$

with C independent of ε but possibly depending on X. Since (from (4.15))

$$||\tilde{u}(x)||_0^{(-\infty,X)} \le C\varepsilon||h^\wedge||,$$

it follows that

$$||\tilde{u}||_0^{(-X,X)} \le C_1\varepsilon(||h^\wedge|| + ||h||_0).$$

Hence by (4.16)

(4.17) $$||\tilde{u}(x)||_0^{(X,\infty)} \le C_1\varepsilon(||h^\wedge|| + ||h||_0 + ||h^\wedge||_+) \le C_2\varepsilon.$$

By the uniqueness result in Lemma 4.3, we have

$$u = \tilde{u} + cw^+,$$

where w^+ is the function from Lemma 4.2, and where

$$c = (\tilde{u}_1(0) - \tilde{u}_2(0))/(w_1^+(0) - w_2^+(0)).$$

The same type of argument as in the paragraph above and in fact mentioned already following Lemma 4.2, may be used to show that the solution w^+ may be chosen so that $||w^+||_0$ and $||w^{+\wedge}||$ are bounded independently of ε, and also $w_1^+(0) - w_2^+(0)$ is bounded away from 0 independently of ε. By (4.17) and Lemma 4.2, we obtain that $c = O(\varepsilon)$, and the conclusion of the theorem follows.

5. Uniqueness. Much of the same apparatus developed in §4 will be used here.

THEOREM 5.1. *There is an $\varepsilon_0 > 0$ such that for fixed positive $\varepsilon \leq \varepsilon_0$, there can exist no more than one solution of* (2.9), (2.10) *satisfying* $n_1(0) = n_2(0)$.

Suppose, in fact, that there were two such solutions, (n, E) *and* $(\bar{n}, \bar{E})$. *Let* $u \equiv n - \bar{n}$, $\gamma \equiv E - \bar{E}$, $V \equiv E - F(n)$, $\bar{V} \equiv \bar{E} - F(\bar{n})$, *and finally* $w \equiv V - \bar{V}$. *Then with D denoting differentiation,*

$$(5.1) \qquad \varepsilon \ddot{V} = NV - \varepsilon D_x^2 F(n(x)).$$

and the same equation holds with bars over V, N, and n. Subtracting the two, we obtain

$$(5.2) \qquad \varepsilon \ddot{w} = Nw + q,$$

where $q = (N - \bar{N})\bar{V} - \varepsilon D_x^2(F(n) - F(\bar{n}))$.

In the following, a special norm will be used:

$$||u^\sim|| \equiv ||u^\wedge|| + ||u||_0 + ||u^\wedge||_+,$$

the norms on the right being defined as in §4 with the positive numbers X and δ used there.

LEMMA 5.2. *Let $||h^\sim|| < \infty$. Let $U(x)$ satisfy*

$$\varepsilon \ddot{U} = N(x)(U - h(x)),$$

$$|U(x) \exp(\delta|x|)| \to 0 \quad as \ |x| \to \infty.$$

Then

$$(5.4) \qquad ||U^\sim|| \leq C||h^\sim||$$

for some C independent of ε (as will be all constants C below).

PROOF. It is an easy consequence of the maximum principle that

$$(5.5) \qquad ||U||_0 \leq C||h||_0.$$

Now let $W \equiv Ue^{-\delta x}$, $H = he^{-\delta x}$. Assume for the moment that $W \to 0$ as $x \to -\infty$. We have from the differential equation for w that

$$\varepsilon \ddot{W} = N(W - H) - 2\varepsilon(\delta \dot{W} + \delta^2 W).$$

At a local maximum of W in $(-\infty, -X)$, we have $\ddot{W} \leq 0$ and $\dot{W} = 0$, so that $W \leq CH$ at that point. Since W vanishes at $-\infty$, it cannot have a maximum there. Finally at $x = -X$, we have from (5.5) that

$$|W| = |U|e^{-\delta X} \leq C||h||_0 e^{-\delta X} \leq C||h||_0 \leq C||h^\wedge||,$$

so in all, we have on $(-\infty, -X)$, $\sup |W(x)| \leq C||h^\wedge||$. Hence

$$(5.6) \qquad ||U^\wedge|| \leq C||h^\wedge||.$$

Now relax the assumption that $W \to 0$ at $-\infty$. Let h_k be the truncation of h to be zero for $x \leq -k$, and U_k the corresponding solution with h_k in place of h. The existence of U_k follows from an upper-lower solution argument, and it approaches zero at a rapid exponential rate as $x \to -\infty$. For U_k and h_k we have the estimate

analogous to (5.5). Moreover C in that estimate is independent of k. As $k \to \infty$, the solutions U_k are equicontinuous and a subsequence approaches a solution of the same equation with (5.5) continuing to hold. Uniqueness of bounded solutions follows from the maximum principle, so the limit so obtained must be the original postulated solution. Therefore (5.6) holds in general. This, combined with (5.5) and a similar argument for (X, ∞) completes the proof.

LEMMA 5.3. *The function* w *satisfies*

$$||w^\sim|| \le \varepsilon C ||u^\sim||.$$

PROOF. By Lemma 3.14, w satisfies the hypotheses of Lemma 5.2 for small enough δ. Applying that lemma, we obtain

$$(5.7) \qquad ||w^\sim|| \le C||q^\sim||.$$

However, using the fact (Lemma 3.13) that $|\bar{V}| \le C\varepsilon$, we see from the expression for q that

$$(5.8) \qquad ||q^\sim|| \le C\varepsilon(||u^\sim|| + ||\dot{u}^\sim|| + ||\ddot{u}^\sim||) \equiv C\varepsilon ||u^\sim||_2.$$

From (5.2), (5.7), and (5.8), we obtain

$$(5.9) \qquad ||\ddot{w}^\sim|| \le C||u^\sim||_2,$$

and an easily derived interpolation inequality applied to (5.7) and (5.9) yields

$$(5.10) \qquad ||\dot{w}^\sim|| \le C\varepsilon^{1/2}||u^\sim||_2.$$

By the definitions of V and w, we have $w = \gamma - F'(\tilde{n})u$ for some intermediate $\tilde{n}$, hence

$$(5.11) \qquad ||\gamma^\sim|| \le ||w^\sim|| + C||u^\sim||, \qquad ||\dot{\gamma}^\sim|| \le ||\dot{w}^\sim|| + C||u^\sim||_1.$$

Simply taking differences in the equation $\dot{n} = \bar{g}(n, E)$ and its derivative, we obtain

$$||u^\sim||_1 \le C(||u^\sim|| + ||\gamma^\sim||), \quad ||u^\sim||_2 \le C(||u^\sim||_1 + ||\gamma^\sim||_1).$$

From this and (5.11),

$$||u^\sim||_2 \le C(||u^\sim||_1 + ||\dot{w}^\sim||) \le C(||u^\sim|| + ||w^\sim|| + ||\dot{w}^\sim||).$$

We now estimate the last term on the right by (5.10) and eliminate the term in $\varepsilon^{1/2}$ to obtain

$$||u^\sim||_2 \le C(||u^\sim|| + ||w^\sim||).$$

This, together with (5.7) and (5.8), now yields

$$(5.12) \qquad ||w^\sim|| \le \varepsilon C||u^\sim||,$$

which completes the proof.

PROOF OF THEOREM 5.1. Let $p(n, V) \equiv \tilde{g}(n, F(n) + V)$, where this function is given by (4.2). Then $\dot{n} = p(n, V)$, and the same is true with n and V replaced by $\bar{n}$ and $\bar{V}$. Taking differences results in the equation

$$(5.13) \qquad \dot{u} = p_1(\hat{n}, \hat{V})u + p_2(\hat{n}, \hat{V})w,$$

where p_1 and p_2 are the derivatives of p with respect to n and V respectively, $\hat{n} = n_1 + \tau u$, $0 < \tau(x) < 1$, and similarly for $\hat{V}$. Denote the matrix coefficient p_1

on the right of (5.13) by $A^*(x)$; it has the same essential properties as did A_1 in the last section. Also denote the last term in (5.13) by $h^*(x)$; because of the factor w, it of course decays exponentially as $|x| \to \infty$. Now (5.13) becomes

$$(5.14) \qquad \dot{u} = A^*(x)u + h^*(x).$$

The coefficient of w in h^* being bounded, we have from Lemma 5.3 that

$$(5.15) \qquad ||h^{*\sim}|| \le ||w^\sim|| \le C\varepsilon ||u^\sim||.$$

Repeating the arguments used in §4, we obtain the result that the solution u of (5.14) decaying exponentially as $x \to -\infty$ and satisfying $u_1(0) = u_2(0)$ is unique, i.e. $u \equiv n - \bar{n}$.

Consider separately

$$A^*(x) = A_- + R_-(x); \quad A^*(x) = A_+ + R_+(x).$$

By the same type of argument as in §4, the solution of (5.14) can be represented by the solution $\tilde{u}(x)$ of an integral equation on $(-\infty, -X)$. Now $\tilde{u}$ is the solution obtained in §4. Therefore $u = \tilde{u}(x) + c_0 w_+(x)$. It is known from the above arguments that

$$||\tilde{u}^\wedge|| \le C||h^*(x)^\wedge||, \quad ||\tilde{u}||_0 \le C||h^{*\sim}||, \quad ||w_+^\wedge|| \le C,$$

$$|c_0| = \frac{|\tilde{u}_2(0) - \tilde{u}_1(0)|}{|w_1^+(0) - w_2^+(0)|} \le C||\tilde{u}||_0 \le C||h^{*\sim}||.$$

Therefore

$$||u^\wedge|| \le ||\tilde{u}^\wedge|| + ||c_0 w_+^\wedge|| \le ||\tilde{u}^\wedge|| + |C_0|\,||w_+^\sim|| \le C||h^{*\sim}||.$$

Now on (X, ∞), all solutions decaying exponentially as $x \to \infty$ can be represented as solutions of an integral equation as follows. With no loss of generality, take A_+, as before, to be diagonal. Then

$$u(x) = J_2^+ h^*(x) + J_2^+ R_+(x)u(x) + U_+(x),$$

where the operator J_2^+ and the function $U_+(x)$ are defined componentwise as follows:

$$(J_2^+ h^*(x))_1 = \int_\infty^x h_1^*(s)\, ds,$$

$$(J_2^+ h^*(x))_2 = \int_X^x e^{(b_2-1)(x-s)} h_2^*(s)\, ds,$$

$$(J_2^+ h^*(x))_3 = \int_X^x e^{(-(b_2+b_3)/2)(x-s)} h_3^*(s)\, ds,$$

$$U_+(x) = \begin{pmatrix} 0 \\ e^{(b_2-1)(x-X)} u_2(X) \\ e^{(-(b_2+b_3)/2)(x-X)} u_3(X) \end{pmatrix}.$$

Here $u_2(X)$ and $u_3(X)$ are arbitrary; there is a two-parameter family of solutions with these properties. In particular, this representation holds for the function u being considered here.

By an argument similar to that in §4, we have

$$||J_2^+ R(x)u^\wedge||_+ \le \tfrac{1}{2}||u^\wedge||_+$$

if X is sufficiently large. Therefore

$$\|u^\wedge\|_+ \le 2\|J_2^+ h^{*\wedge}\|_+ + 2\|U_0^\wedge\|_+ \le C\|h^{*\sim}\|_+ + C|u(X)|$$
$$\le C(\|h^{*\sim}\|_+ + \|h^{*\sim}\|) \le C'\|h^{*\sim}\|.$$

In the above, it was shown that $\|u^\wedge\| + \|u\|_0 \le C\|h^{*\sim}\|$. Therefore $\|u^\sim\| \le C\|h^{*\sim}\|$. From (5.15) we have $\|\tilde{u}\| \le C\|h^{*\sim}\| \le C\varepsilon\|\tilde{u}\|$ for small enough ε. This implies that $u \equiv 0$, i.e. $n \equiv \bar{n}$, which completes the proof.

Discussion. We have investigated the most basic traveling wave separation process, namely that separating two positive ions, with a common negative counterion. The existence and uniqueness of such traveling waves has been given. Neither of these results is obvious or easy when $\varepsilon > 0$, although an explicit unique expression can be given for the solution when $\varepsilon = 0$. The uniqueness proof turns out to be a deeper result than the existence, and the method used may be applicable to other problems.

The emphasis has been on the case when ε is small, and on obtaining qualitative results independent of ε. It is shown, as expected, that the unique solution with positive ε is uniformly $O(\varepsilon)$ close to the electrically neutral solution for $\varepsilon = 0$.

The method of constructing approximate solutions (on a finite interval) given in §3 is suggestive of a comparable numerical method for calculating the solutions. And indeed it does yield a reasonable approach; these results will be reported in a separate paper.

ACKNOWLEDGMENTS. We are grateful to Wu Ya-ping for correcting some of the proofs in §§4 and 5.

REFERENCES

[**Ag**] D. Agin, *Electroneutrality and electrodiffusion in the squid axon*, Proc. Nat. Acad. Sci. U.S.A. **57** (1967), 1232–1238.
[**Al**] R. A. Alberty, *Moving boundary systems formed by weak electrolytes: Theory of simple systems formed by weak acids and bases*, J. Amer. Chem. Soc. **72** (1950), 361.
[**ABR**] R. A. Arandt, J. D. Bond, and L. D. Roper, *Electrical approximate solutions of steady-state electrodiffusion for a simple membrane*, J. Theoret. Biol. **34** (1972), 265–276.
[**B**] M. Bier et al., *Electrophoresis: Mathematical modeling and computer simulation*, Science **219** (1983), 281.
[**CB**] M. Coxon and M. J. Binder, *Isotachophoresis theory*, J. Chromatogr. **95** (1974), 133.
[**CL**] E. Coddington and N. Levinson, *Theory of ordinary differential equations*, McGraw-Hill, New York, 1955.
[**De**] Z. Deyl, ed., *Electrophoresis: A survey of techniques and applications*, Elsevier, Amsterdam, 1979.
[**Do**] V. P. Dole, *A theory of moving boundary systems formed by strong electrolytes*, J. Amer. Chem. Soc. **67** (1945), 119.
[**K**] F. Kohlrausch, *Ueber Concentrations-Verschiebungen durch Electrolyse im Inneren von Losungen und Losungsgemischen*, Ann. Physik **62** (1897), 209.
[**SP**] D. A. Saville and O. A. Palusinski, *Theory of electrophoretic separations*, Parts 1 and 2, AIChE J. **32** (1986).

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ARIZONA, TUCSON, ARIZONA 85721

Current address (P. C. Fife and Y. Su): Department of Mathematics, University of Utah, Salt Lake City, Utah 84112

Current address (O. A. Palusinski): Department of Electrical and Computer Engineering, University of Arizona, Tucson, Arizona 85721