Analysis of an Upwind-Mixed Finite Element Method
for Nonlinear Contaminant Transport Equations

Clint N. Dawson

November, 1993

TR93-57
ANALYSIS OF AN UPWIND-MIXED FINITE ELEMENT METHOD FOR NONLINEAR CONTAMINANT TRANSPORT EQUATIONS

CLINT DAWSON *

Abstract. In this paper, the numerical approximation of a nonlinear diffusion equation arising in contaminant transport is studied. The equation is characterized by advection, diffusion, and adsorption. Assuming the adsorption term is modeled by a Freundlich isotherm, it can be nonlinear in concentration and nondifferentiable as the concentration approaches zero. We consider the approximation of this equation by a method which upwinds the advection and incorporates diffusion using a mixed finite element method. Error estimates for a semi-discrete formulation are derived, and numerical results for a fully discrete formulation are given.

Keywords: Advection, diffusion, adsorption, mixed finite elements, upwinding

AMS(MOS) subject classification: 65N30, 65N06, 65N22

1. Introduction. In this paper, we consider the flow of an incompressible fluid through a homogeneous, saturated porous medium, where the fluid is contaminated by a solute, with concentration \( c \geq 0 \). We will assume flow is at steady-state, and that transport is described by advection, molecular diffusion and mechanical dispersion, and chemical reactions (adsorption) between the solute and the surrounding solid porous skeleton. We will also assume flow is one-dimensional and in the direction of the positive \( x \)-axis. Under these assumptions, mass-conservation of the contaminant gives the equation [4]

\[
ct + \frac{\rho_b}{n} \frac{s_t}{s} + vc_x - Dc_{xx} = 0, \quad 0 < x < L, \quad t > 0,
\]

where \( s(x,t) \) is the mass of contaminant adsorbed on the solid matrix per unit mass of solid, \( \rho_b \neq 0 \) is the bulk density of the porous medium, \( n > 0 \) is the porosity, \( v \) is the effective fluid velocity, and \( D > 0 \) accounts for molecular diffusion and mechanical dispersion. For simplicity, \( \rho_b, n, v, \) and \( D \) are assumed constant.

The chemical reactions describing adsorption may be fast (equilibrium) or slow (non-equilibrium) depending on the rate of reaction with respect to the rate of flow. In the case of equilibrium adsorption reactions, the contaminant adsorbed by the solid is generally assumed to be a function of the concentration in the fluid; that is,

\[
s = f(c).
\]

In the case of non-equilibrium adsorption reactions, \( s \) is assumed to satisfy

\[
s_t = k(f(c) - s),
\]

where \( k \) is a rate parameter. Thus (2) is obtained in the limit as \( k \to \infty \). The function \( f \) in (2) and (3) is called an adsorption isotherm. Common isotherms are the Langmuir isotherm

\[
f(c) = \frac{K_1 c}{1 + K_2 c}, \quad K_1, K_2 > 0,
\]

* Dept. of Computational and Applied Mathematics, Rice University, Houston, TX 77251.
and the Freundlich isotherm

\begin{equation}
 f(c) = K_3 c^p, \quad K_3 > 0.
\end{equation}

In (5), \( p \in (0, 1] \) is commonly chosen.

In this paper, we consider the case of equilibrium adsorption. Substituting (2) into (1), letting

\begin{equation}
 \phi(c) = \frac{\rho_k}{n} f(c),
\end{equation}

we obtain

\begin{equation}
 c_t + \phi(c)c_t + vc_x - Dc_{xx} = 0, \quad 0 < x < L, \quad t > 0.
\end{equation}

We will augment (7) with the initial and boundary conditions

\begin{align}
 c(x, 0) &= c^0(x), \quad 0 < x < L, \\
 c(0, t) &= c_0(t), \quad t > 0, \\
 c(L, t) &= c_L(t), \quad t > 0.
\end{align}

We note that (7) is uniformly parabolic if \( \phi(c) \) is smooth for \( c \geq 0 \). If \( f(c) \) is of Freundlich type, however, then (7) is degenerate parabolic for \( p < 1 \), since \( f'(c) \) blows-up as \( c \downarrow 0 \). In this case, the set \( \{ c > 0 \} \) has a free boundary, and expands with finite speed. Across this free boundary, the smoothness of the solution is limited. A much more detailed description of the adsorption process and the mathematical ramifications are given in [11].

We discuss the numerical approximation of (7)-(10) using a method based on upwinding the advective term in (7), and adding diffusion using a mixed finite element method. We note that in many important cases, (7) is “advection-dominated,” meaning diffusion is small relative to advection. In these cases, it is well-known that advection should be approximated by some upwind method which incorporates characteristic information. In this paper, we will derive estimates assuming a one-point upwinding method for approximating advection. In practice, higher-order upwinding schemes such as those described in [5] can be applied, but the error estimates are limited by the smoothness of the solution. Our analysis allows for higher-order approximations to the advective flux when the solution is sufficiently smooth.

Numerous publications on the numerical approximation of (1) have appeared, see, for example, [2, 3, 7, 8, 12, 13, 16]. In [2] and [3], Barrett and Knabner describe and analyze a piecewise-linear Galerkin finite element approximation to a model similar to (1) but with no advection term \( (v = 0) \), assuming both non-equilibrium and equilibrium adsorption kinetics. They solve a perturbed system, replacing \( \phi(c) \) by a smooth function \( \phi_\epsilon(c) \) which converges to \( \phi \) as \( \epsilon \to 0 \). In the equilibrium case, they also assume (1) and (3) hold and allow \( k \to \infty \). The approach described here differs from their approach in that it is better-suited to advection-dominated problems,
and uses a mixed finite element approximation rather than a Galerkin approximation. In [8], a Characteristic-Galerkin approximation to contaminant transport with non-equilibrium adsorption kinetics is described and analyzed. The methodology described here also differs substantially from this approach, and in fact, the analysis given in [8] does not directly extend to the problem considered here.

The upwind-mixed method described below has been implemented numerically for both equilibrium and non-equilibrium adsorption kinetics [7], and has been used successfully to model long-time behavior of solutions, assuming the Freundlich isotherm with various values of $p > 0$ [12]. More recently, the method has been extended to multidimensions and is now being used to study long-time behavior of solutions to a two-dimensional contaminant transport model, again assuming equilibrium kinetics [9].

The methodology described below and the error estimates are not restricted to one space dimension. Our primary reasons for assuming one-dimensional flow is simplicity. The mathematical techniques used to analyze the numerical method require a special treatment of the time-derivative term $\phi(c)_t$. Since this term is not spatially dependent, the same techniques can be extended to multi-dimensions, assuming the solution is sufficiently smooth.

This paper is divided into the following additional sections. In the next section, we give notation and state regularity assumptions on the coefficients and solution to (7)-(10). In Section 3, we derive the discrete-time and semi-discrete upwind-mixed finite element methods for solving (7)-(10), and in Section 4, we analyze the semi-discrete method and state an error estimate. Finally, in Section 5, we present numerical results for the discrete-time scheme described in Section 3.

2. Notation and assumptions. In this section, we define notation and state the regularity assumptions which will be used in the error estimates below.

Let $\Omega = (0, L)$. Let $(\cdot, \cdot)$ denote the $L^2(\Omega)$ inner product, and $|| \cdot ||$ the corresponding norm. Let $C^\beta(\Omega)$ denote the space of functions defined on $\Omega$ which are Hölder continuous with exponent $\beta > 0$.

Let $[0, T]$ denote a time interval, where $T > 0$ is a fixed constant, and let $\mathcal{X} = \mathcal{X}(\Omega)$ denote a normed space. Denote by $||| \cdot |||_{L^\infty(0,T;\mathcal{X})}$ the norm of $\mathcal{X}$-valued functions $f$ with the map $t \rightarrow ||f(t)||_{\mathcal{X}}$ belonging to $L^\infty(0,T)$. Define $||| \cdot |||_{H^1(0,T;\mathcal{X})}$ similarly.

We will make the following assumptions on the coefficients and data in (7)-(10):

(A1) $\phi(c)$ is Hölder continuous on $[0, \infty)$ with exponent $p \in (0, 1)$, and $\phi$ is monotone increasing.

(A2) $v > 0$ and $D > 0$.

(A3) $e^\theta(x)$ is continuously differentiable on $\Omega$ and nonnegative.

(A4) $c_0(t)$ and $c_L(t)$ are continuous and nonnegative.

(A5) $c \in H^1(0,T; L^2(\Omega)) \cap L^\infty(0,T; C^{1+\alpha}(\Omega))$, where

$$\alpha = \frac{p}{1-p},$$

and $p$ is the exponent in (A1).
The first smoothness assumption on $c$ in (A5) is proven in [3]. The second smoothness assumption in (A5), though not formally proven there, is suggested by results given in [10], where travelling wave solutions to (1) are studied. The results in [10] can be used to show that, for travelling wave solutions $c$ propagating in the positive $x$ direction, $c$ approaches zero like $(s(t) - x)^{\frac{1}{1-r}}$, where $s(t)$ denotes the boundary of the free surface separating $\{c > 0\}$ from $\{c = 0\}$. Extension of this estimate to more realistic models, that is, multidimensional equations with variable coefficients, is not known at this time. However, the existence of travelling wave solutions to multidimensional extensions of (1) has been studied [11, 17, 18].

We remark that for exponents $p \geq 1$, the Freundlich isotherm (4) is a smooth function, and the solution $c$ can also be expected to be smooth.

Finally, we note that the presence of a Dirichlet boundary condition at $x = L$ may cause the formation of boundary layers in this region. In practice, we normally simulate situations where $c$ has compact support in some interval $[0, L]$ for all time $0 < t < T$, and set $c_L(t) = 0$. We may instead enforce an “outflow” condition at $x = L$, but we do not consider that case here.

3. Formulation of the method. In order to derive the upwind-mixed method, we rewrite (7) in mixed form. Let $g$ and $z$ satisfy

\begin{equation}
(11) \quad g = uc,
\end{equation}

and

\begin{equation}
(12) \quad D^{-1}z + c_x = 0.
\end{equation}

Then (7) can be written as

\begin{equation}
(13) \quad c_t + \phi(c)_t + g_x + z_x = 0.
\end{equation}

Let

\[ \delta_x : 0 = x_{-\frac{1}{2}} < x_{\frac{1}{2}} < \ldots < x_{j+\frac{1}{2}} = L \]

denote a partition of $[0, L]$ into intervals $B_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, $j = 1, \ldots, J$. Let $h_j = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}$, and let $h = \max_j h_j$. Let $W_h$ and $V_h$ denote the lowest-order Raviart-Thomas approximating spaces for $c$ and $z$ [15]; that is, $W_h$ is the space of piecewise constant functions defined on $\delta_x$, with basis functions $\{w_j\}$, where

\begin{equation}
(14) \quad w_j(x) = \begin{cases} 1, & x \in B_j, \\ 0, & \text{otherwise}, \end{cases}
\end{equation}

and $V_h$ is the space of continuous, piecewise linear functions on $\delta_x$. With $V_h$, we associate the standard basis $\{v_{j+\frac{1}{2}}\}$ of “hat” functions, where $v_{j+\frac{1}{2}}$ is the continuous, piecewise linear function on $[0, L]$ satisfying

\begin{equation}
(15) \quad v_{j+\frac{1}{2}}(x_{i+\frac{1}{2}}) = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}
\end{equation}
Note that any function $v$ in $V_h$ has the representation

\begin{equation}
(16)\quad v(x) = \sum_{j=0}^{J} v(x, j+\frac{1}{2}) v_{j+\frac{1}{2}}(x).
\end{equation}

We first describe a scheme which is discrete in time. Let $\Delta t > 0$ denote a time step, and $t^n = n\Delta t$, $n = 0, 1, \ldots$. Let $f^n = f(t^n)$. Discretizing (13) in time, using an explicit approximation to the advective flux $g$ and an implicit approximation to the diffusive flux $z$, we obtain the following scheme. For $n = 0$,

\begin{equation}
(17)\quad (C^0, w) = (c^0, w), \quad w \in W_h.
\end{equation}

For $n = 1, 2, \ldots$,

\begin{align}
(18)\quad (\partial_t C^n, w) + (\phi(C^n), w) + (Z^n, w) = -(C^n, w), \quad w \in W_h,
\end{align}

\begin{align}
(19)\quad (D^{-1} Z^n, v) - (C^n, v_x) = c_0(t^n)v(0) - c_L(t^n)v(L), \quad v \in V_h,
\end{align}

where

\begin{align}
(20)\quad G(x, j+\frac{1}{2}, t^n) = \begin{cases} vC^{n-1}(x, j), & j \neq 0, \\ v_c(t^n), & j = 0, \end{cases}
\end{align}

and

\begin{align}
(21)\quad \partial_t C^n = \frac{C^n - C^{n-1}}{\Delta t}.
\end{align}

Write

\begin{align}
C^n(x) = \sum_{j=1}^{J} C^n_j w_j(x),
\end{align}

\begin{align}
Z^n(x) = \sum_{j=0}^{J} Z^n_{j+\frac{1}{2}} v_{j+\frac{1}{2}}(x),
\end{align}

and

\begin{align}
G^n(x) = \sum_{j=0}^{J} G^n_{j+\frac{1}{2}} v_{j+\frac{1}{2}}(x).
\end{align}

Let $C^n$, $Z^n$, and $G^n$ denote the vectors of unknowns corresponding to $C^n$, $Z^n$, and $G^n$, respectively. Let $A$ be the diagonal matrix with $A_{i,j} = (w_j, w_i)$, $M$ the symmetric, positive definite matrix with $M_{i,j} = (v_j, v_i)$, and $N$ the rectangular matrix with $N_{i,j} = (w_j, (v_i)_x)$. Let $B^n$ and $\Phi^n$ denote vectors with $B^n_i = c_0^n v_i(0) - c_L^n v_i(L)$, and $\Phi^n_i = \phi(C^n_i) h_i$. Then, (19) leads to

\begin{align}
Z^n = M^{-1}(B^n + NC^n),
\end{align}
and (18) leads to

\[ AC^n + \Phi^n + \Delta t N^T Z^n = AC^{n-1} + \Phi^{n-1} - \Delta t N^T G^{n-1} \]

or

\[ (A + \Delta t N^T M^{-1} N) C^n + \Phi^n = AC^{n-1} + \Phi^{n-1} - \Delta t N^T [G^{n-1} - M^{-1} B^n]. \]

Since \( A + \Delta t N^T M^{-1} N \) is symmetric and positive definite, and \( \phi \) is continuous and monotone, the existence and uniqueness of \( C^n \) follows from standard monotone operator theory [14].

The semi-discrete analog of (17)-(20) is the following. Initially, set

(22) \[ (C(\cdot, 0) - \phi(\cdot), w) = 0, \quad w \in W_h. \]

Then, for each \( t \in (0, T] \), find \( C(\cdot, t) \in W_h, Z(\cdot, t) \in V_h, \) and \( G(\cdot, t) \in V_h \) which satisfy

(23) \[ (C_t + \phi(C) + G_x + Z_x, w) = 0, \quad w \in W_h, \]

(24) \[ (D^{-1} Z, v) - (C, v_x) = c_0(t)v(0) - c_L(t)v(L), \quad v \in V_h, \]

where

(25) \[ G(x_{j+\frac{1}{2}}, t) = \begin{cases} \nu C(x_j, t), & j \neq 0, \\ \nu c_0(t), & j = 0. \end{cases} \]

4. **Analysis of the semi-discrete scheme.** In the arguments below, let \( K \) denote a generic, positive constant. We will also use the inequality

(26) \[ ab \leq \frac{\epsilon}{2} a^2 + \frac{1}{2\epsilon} b^2, \quad a, b, \epsilon \in \mathbb{R}, \quad \epsilon > 0. \]

Let \( \tilde{c} \) be the \( L^2 \)-projection of \( c \) into \( W_h \):

(27) \[ (\tilde{c}(\cdot, t) - c(\cdot, t), w) = 0, \quad w \in W_h, \]

and let \( \tilde{z}, \tilde{g} \in V_h \) satisfy

(28) \[ (\tilde{z}_x - z_x, w) = (\tilde{g}_x - g_x, w) = 0, \quad w \in W_h. \]

In one space dimension, \( \tilde{z} (\tilde{g}) \) can be chosen as the interpolant of \( z (g) \) in \( V_h \). Analogous projections exist in multidimensions. Under the regularity assumptions [A5], we have the following estimates:

(29) \[ |(c - \tilde{c})(x, t)| \leq K h, \]

(30) \[ |(g - \tilde{g})(x, t)| \leq K h, \]
and
\begin{equation}
|(z - \tilde{z})(x, t)| \leq Kh^{\sigma(p)},
\end{equation}
where
\begin{equation}
\sigma(p) = \begin{cases} 
\alpha, & 0 < p < \frac{1}{2}, \\
1, & \frac{1}{2} \leq p \leq \frac{3}{4}, \\
2, & p \geq \frac{3}{4}.
\end{cases}
\end{equation}

Moreover,
\begin{equation}
|G(x_{j+\frac{1}{2}}, t) - g(x_{j+\frac{1}{2}}, t)| \leq v|C(x_j, t) - c(x_{j+\frac{1}{2}}, t)| \\
\leq v|(C - \tilde{c})(x_j, t)| + v|(\tilde{c} - c)(x_j, t)| \\
+ v|c(x_j, t) - c(x_{j+\frac{1}{2}}, t)| \\
\leq K \left[|(C - \tilde{c})(x_j, t)| + h\right],
\end{equation}
by (29). Noting that \( \tilde{g} \) agrees with \( g \) at \( x_{j+\frac{1}{2}} \), multiplying above by \( h_{j+\frac{1}{2}} = \frac{1}{2}[h_j + h_{j+1}] \) and summing on \( j \), we obtain the discrete \( L^2(\Omega) \) estimate
\begin{equation}
\sum_{j=1}^{J} |(G - \tilde{g})(x_{j+\frac{1}{2}}, t)|^2 h_{j+\frac{1}{2}} \leq K \left[ \sum_{j=1}^{J} |(C - \tilde{c})(x_j, t)|^2 h_{j+\frac{1}{2}} + h^2 \right].
\end{equation}

Since these discrete \( L^2 \) norms are equivalent to the continuous \( L^2 \) norm, with equivalence constants independent of \( h \), we find
\begin{equation}
\|(G - \tilde{g})(\cdot, t)\| \leq K^* [h + \|(C - \tilde{c})(\cdot, t)\|],
\end{equation}
where \( K^* \) is a positive constant independent of \( h \).

For \( w = w(x, t) \) continuously differentiable in \( t \), integrating (23) in time and integrating by parts gives
\begin{align}
- \int_0^T (C, w_t)dt & - \int_0^T (\phi(C), w_t)dt + \int_0^T (Z_x, w)dt \\
= & -(C(\cdot, T) + \phi(C(\cdot, T)), w(\cdot, T)) + (C(\cdot, 0) + \phi(C(\cdot, 0)), w(\cdot, 0)) \\
& - \int_0^T (G_x, w)dt.
\end{align}

Let \( \zeta = C - \tilde{c} \), \( \xi = c - \tilde{c} \), \( \kappa = Z - \tilde{z} \), and \( \gamma = z - \tilde{z} \). Then, multiplying (13) by \( w(\cdot, t) \in W_h \) with \( w \) smooth in \( t \), integrating in time, integrating by parts, subtracting the result from (36), and applying (27), (28), and (22), we find
\begin{align}
- \int_0^T (\zeta, w_t)dt & - \int_0^T (\phi(C) - \phi(\tilde{c}), w_t)dt + \int_0^T (\kappa, w)dt \\
= & \int_0^T (\tilde{g}_x - G_x, w)dt - \int_0^T (\phi(c) - \phi(\tilde{c}), w_t) \\
& - (\zeta(\cdot, T), w(\cdot, T)) + (\xi(\cdot, T), w(\cdot, T)) \\
& - (\phi(C(\cdot, T)) - \phi(c(\cdot, T)), w(\cdot, T)) \\
& + (\phi(\tilde{c}(\cdot, 0)) - \phi(c(\cdot, 0)), w(\cdot, 0)) \\
& + (\phi(\tilde{c}(\cdot, 0)) - \phi(c(\cdot, 0)), w(\cdot, 0))
\end{align}
Multiplying (12) by \( v \in V_h \), integrating by parts, and subtracting the result from (24), we find

\[
(D^{-1} \kappa(\cdot, s), v) - (\zeta(\cdot, s), v_x) = (D^{-1} \gamma(\cdot, s), v), \quad v \in V_h,
\]

for each \( s \in (0, T] \).

In (37), let

\[
w(\cdot, t) = \int_t^T e^{-rs} \zeta(\cdot, s) ds, \quad 0 \leq t \leq T,
\]

where \( r > 0 \) will be chosen below. This gives

\[
\int_0^T e^{-rt} ||\zeta(\cdot, t)||^2 dt + \int_0^T e^{-rt} (\phi(C) - \phi(\bar{c}), \zeta) dt
\]

\[
+ \int_0^T (\kappa_x(\cdot, t), \int_t^T e^{-rs} \zeta(\cdot, s) ds) dt
\]

\[
= \int_0^T (\bar{g}_x - G_x, \int_t^T e^{-rs} \zeta(\cdot, s) ds) dt + \int_0^T e^{-rt} (\phi(c) - \phi(\bar{c}), \zeta) dt
\]

\[
+ (\phi(\bar{c}(\cdot, 0)) - \phi(c), \int_0^T e^{-rs} \zeta(\cdot, s) ds).
\]

A test function similar to (39) was used by Arbogast [1] to analyze a Galerkin approximation to Richard’s equation.

Multiplying (38) by \( e^{-rs} \) and integrating in time (holding \( v \) fixed), we find

\[
(D^{-1} \int_t^T e^{-rs} \kappa(\cdot, s) ds, v) - (\int_t^T e^{-rs} \zeta(\cdot, s) ds, v_x)
\]

\[
= (D^{-1} \int_t^T e^{-rs} \gamma(\cdot, s) ds, v), \quad v \in V_h.
\]

Set \( v = \kappa(\cdot, t) \), then integrating by parts

\[
\int_0^T D^{-1} (\int_t^T e^{-rs} \kappa(\cdot, s) ds, \kappa(\cdot, t)) dt
\]

\[
= \frac{D^{-1}}{2} \left( \int_0^T e^{-rs} \kappa(\cdot, s) ds, \int_0^T e^{-rs} \kappa(\cdot, s) ds \right)
\]

\[
+ \frac{D^{-1}}{2} \int_0^T re^{rt} (\int_t^T e^{-rs} \kappa(\cdot, s) ds, \int_t^T e^{-rs} \kappa(\cdot, s) ds) dt.
\]

Thus,

\[
\frac{D^{-1}}{2} \left| \int_0^T e^{-rs} \kappa(\cdot, s) ds \right|^2 + \frac{D^{-1}}{2} \int_0^T re^{rt} \left| \int_t^T e^{-rs} \kappa(\cdot, s) ds \right|^2 dt
\]

\[
= \int_0^T (\int_t^T e^{-rs} \gamma(\cdot, s) ds, \kappa(\cdot, t)) dt
\]

\[
+ \int_0^T (\int_t^T e^{-rs} \zeta(\cdot, s) ds, \kappa_x(\cdot, t)) dt.
\]
Adding (40) and (43) we find

\[
\int_0^T e^{-rt}||\zeta(\cdot, t)||^2dt + \int_0^T e^{-rt}(\phi(C) - \phi(\tilde{c}), \zeta(\cdot, t))dt
+ \frac{D^{-1}}{2} \left[ \int_0^T e^{-rt}\kappa(\cdot, s)ds \int_t^T e^{-rs}\kappa(\cdot, s)ds \right] + \int_0^T e^{-rt}(\phi(c) - \phi(\tilde{c}), \zeta(\cdot, t))dt
+ \int_0^T (\tilde{\gamma}(\cdot, t) - G_x, \int_t^T e^{-rs}\zeta(\cdot, s)ds)dt
+ \int_0^T (\int_t^T e^{-rs}\gamma(\cdot, s)ds, \kappa(\cdot, t))dt
\]

(44) \equiv T_1 + T_2 + T_3 + T_4.

Consider

(45) \quad T_1 = \int_0^T (\tilde{g}_x - G_x, \int_t^T e^{-rs}\zeta(\cdot, s)ds)dt.

By (41), setting \( v = (\tilde{g} - G)(\cdot, t) \),

\[
T_1 = \int_0^T D^{-1} \left( \int_t^T e^{-rs}[\kappa(\cdot, s) - \gamma(\cdot, s)]ds, (\tilde{g} - G)(\cdot, t) \right)dt
= \int_0^T D^{-1} e^{\frac{1}{2}rt} \int_t^T e^{-rs}\kappa(\cdot, s)ds, e^{\frac{1}{2}rt}(\tilde{g} - G)(\cdot, t)dt
- \int_0^T D^{-1} e^{\frac{1}{2}rt} \int_t^T e^{-rs}\gamma(\cdot, s)ds, e^{\frac{1}{2}rt}(\tilde{g} - G)(\cdot, t)dt
\]

(46) \equiv T_{1,1} - T_{1,2}.

Using (26) and (35)

\[
T_{1,1} \leq \frac{\epsilon}{2} \int_0^T D^{-2} e^{rt} ||\int_t^T e^{-rs}\kappa(\cdot, s)ds ||^2 dt + \frac{1}{2\epsilon} \int_0^T e^{-rt} ||\tilde{g} - G||^2 dt
\leq \frac{D^{-1}\epsilon}{2} \int_0^T D^{-1} e^{rt} ||\int_t^T e^{-rs}\kappa(\cdot, s)ds ||^2 dt
+ \frac{K^*}{2\epsilon} \int_0^T e^{-rt} ||\zeta(\cdot, t)||^2 dt + h^2 dt.
\]

For the second term above

\[
T_{1,2} \leq \int_0^T D^{-2} e^{rt} ||\int_t^T e^{-rs}\gamma(\cdot, s)ds ||^2 dt + \frac{1}{8} \int_0^T e^{-rt} ||\zeta(\cdot, t)||^2 dt + Kh^2.
\]

By the Hölder continuity of \( \phi \) and the estimate (29),

\[
T_2 = \int_0^T e^{-rt}(\phi(c) - \phi(\tilde{c}), \zeta)dt
\leq K \int_0^T e^{-rt} ||\phi(c) - \phi(\tilde{c})||^2 dt + \frac{1}{8} \int_0^T e^{-rt} ||\zeta(\cdot, t)||^2 dt
\]
\[
\leq K \int_0^T e^{-rt} ||c - \hat{c}|| dt + \frac{1}{8} \int_0^T e^{-rt} ||\zeta(\cdot, t)||^2 dt
\]
\[
\leq K h^{2p} \int_0^T e^{-rt} dt + \frac{1}{8} \int_0^T e^{-rt} ||\zeta(\cdot, t)||^2 dt
\]
\[
\leq K rh^{2p} + \frac{1}{8} \int_0^T e^{-rt} ||\zeta(\cdot, t)||^2 dt.
\]

The estimate of $T_3$ is similar:

\[
T_3 = (\phi(\hat{c}(\cdot, 0)) - \phi(c^0(\cdot)), \int_0^T e^{-rt} \zeta(\cdot, t) dt)
\]
\[
= \int_0^T (e^{-\frac{1}{2}rt}(\phi(\hat{c}(\cdot, 0)) - \phi(c^0(\cdot)))e^{-\frac{1}{2}rt}\zeta(\cdot, t)) dt
\]
\[
\leq K rh^{2p} + \frac{1}{8} \int_0^T e^{-rt} ||\zeta(\cdot, t)||^2 dt.
\]

Finally, consider

\[
T_4 = \int_0^T D_1^{-1} \left( \int_t^T e^{-rs} \gamma(\cdot, s) ds, \kappa(\cdot, t) \right) dt
\]
\[
= - \int_0^T D_1^{-1} e^{rt} \left( \int_t^T e^{-rs} \gamma(\cdot, s) ds, \frac{d}{dt} \int_t^T e^{-rs} \kappa(\cdot, s) ds \right) dt.
\]

Integrating by parts, we find

\[
T_4 = D_1^{-1} \left( \int_0^T e^{-rs} \gamma(\cdot, s) ds, \int_0^T e^{-rs} \kappa(\cdot, s) ds \right)
\]
\[
+ D_1^{-1} \int_0^T \left( [re^{rt} \int_t^T e^{-rs} \gamma(\cdot, s) ds - \gamma(\cdot, t)], \int_t^T e^{-rs} \kappa(\cdot, s) ds \right) dt
\]
\[
\leq K || \int_0^T e^{-rs} \gamma(\cdot, s) ds ||^2 + \frac{D_1^{-1}}{4} || \int_0^T e^{-rs} \kappa(\cdot, s) ds ||^2
\]
\[
+ \frac{D_1^{-1}}{8} \int_0^T re^{rt} || \int_t^T e^{-rs} \kappa(\cdot, s) ds ||^2 dt
\]
\[
+ K \int_0^T \left[ re^{rt} || \int_t^T e^{-rs} \gamma(\cdot, s) ds ||^2 + ||\gamma(\cdot, t)||^2 \right] dt.
\]

Substituting the estimates for $T_1 - T_4$ in (44) we find

\[
\int_0^T e^{-rt} ||\zeta(\cdot, t)||^2 dt + \int_0^T e^{-rt}(\phi(C) - \phi(\hat{c}), \zeta(\cdot, t)) dt
\]
\[
+ \frac{D_1^{-1}}{2} \left[ || \int_0^T e^{-rs} \kappa(\cdot, s) ds ||^2 + r \int_0^T e^{rt} || \int_t^T e^{-rs} \kappa(\cdot, s) ds ||^2 dt \right]
\]
\[
\leq \left[ \frac{D_1^{-1}}{2} + \frac{r}{8} \right] D_1^{-1} \int_0^T e^{rt} || \int_t^T e^{-rs} \kappa(\cdot, s) ds ||^2 dt
\]
\[
+ \left[ \frac{K^*}{2} + \frac{3}{8} \right] \int_0^T e^{-rt} ||\zeta(\cdot, t)||^2 dt
\]

10
\[ +K(h^2 + h^{2p}) + \frac{D^{-1}}{4} \left| \int_0^T e^{-rs} \kappa(\cdot,s)ds \right|^2 \\
+K \int_0^T (D^{-2} + r) e^{rt} \left| \int_0^T e^{-rs} \gamma(\cdot,s)ds \right|^2 dt \\
+K \left| \int_0^T e^{-rs} \gamma(\cdot,s) ds \right|^2 + K \int_0^T \left| \gamma(\cdot,t) \right|^2 dt. \]

(51)

Note that the second term on the left of (51) is positive by the monotonicity of \( \phi \).

We now choose \( \varepsilon \) and \( r \) so that

(52) \[ \frac{K^*}{2\varepsilon} + \frac{3}{8} \leq \frac{1}{2} \]

and

(53) \[ \frac{3r}{8} \geq \frac{D^{-1} \varepsilon}{2}, \]

and apply the estimate (31) for \( \gamma \) to obtain

\[ \int_0^T e^{-rt} \left| \zeta(\cdot,t) \right|^2 dt + \int_0^T e^{-rt}(\phi(C) - \phi(c), \zeta(\cdot,t)) dt \]

\[ + \left| \int_0^T e^{-rs} \kappa(\cdot,s) ds \right|^2 \leq K(h^{2p} + h^2 + h^{2\sigma(p)}). \]

(54)

We obtain the following result:

**Theorem:** Let \( c \) and the coefficients and data satisfy (A1)-(A5). Then

(55) \[ \int_0^T \left| C - c \right|^2 dt + \int_0^T (\phi(C) - \phi(c), C - c) dt \leq K h^\mu. \]

where \( K \) is a constant independent of \( h \), and

(56) \[ \mu = \min(1, \sigma(p)), \]

for \( 0 < p \leq 1 \) and \( \sigma(p) \) given by (32).

**Remark.** For \( p > 1 \), assuming \( \phi(c) \) is Lipschitz continuous and monotone increasing, and \( c \) is sufficiently smooth, the estimate given above is first order accurate in \( h \).

5. **Numerical results for the discrete scheme.** In this section, we give some numerical results for the discrete-time scheme discussed in Section 3. We also discuss the approximation of the advective flux \( G \) by a higher-order method.

The scheme (17)-(20) has been implemented under the following assumptions:

(1) \( 0 < x < 1 \),
(1) \( v = 1, D = .01 \),
(2) \( \phi(c) = c^\frac{1}{2} \),
(3) \( c^0(x) \equiv 0, c_0(t) = 1, c_L(t) = 0 \).
The numerical solutions for 25, 50, and 100 evenly spaced mesh points are given at time $T = .8$ in Figure 1. In these runs, $\Delta t = 3h/2$. Figure 1 indicates that the numerical solution converges to a travelling wave as $h$ and $\Delta t$ approach zero.

In Figure 2, we compare the solution for $p = \frac{1}{2}$ computed above with the solution for $p = 1$, both using 100 mesh points. This figure demonstrates the effect of the exponent $p$. For $p < 1$, the "effective" diffusion in the system, which is

\begin{equation}
\frac{D}{1 + \phi'(c)},
\end{equation}

decreases as $c \downarrow 0$, since $\phi'(c)$ blows up. This causes the sharper front for $p = \frac{1}{2}$ seen in the figure.

As mentioned in the introduction, higher-order approximations to the advective flux $G$ can easily be incorporated in our algorithm. Here we construct a higher-order flux approximation as in [5]. Consider the pure advection equation

\begin{equation}
c_t + \phi(c)_t + uc_x = 0.
\end{equation}
Fig. 2. *Comparison of $p = 1$ (−) and $p = .5$ (+) solutions*
Let $\mu = c + \phi(c)$, and $\eta(\mu) = c$, then (58) becomes

$$\mu_t + \nu \eta(\mu)_x = 0,$$

which is a conservation law for $\mu$. Approximating (59) by finite differences in space and time, and using Crank-Nicholson time-stepping, we find

$$\partial_t \mu(x_j, t^n) + \frac{\nu}{h_j} [\eta(\mu(x_{j+\frac{1}{2}}, t^{n+\frac{1}{2}})) - \eta(\mu(x_{j-\frac{1}{2}}, t^{n+\frac{1}{2}}))] \approx 0.$$  

The flux term $\eta(\mu(x_{j+\frac{1}{2}}, t^{n+\frac{1}{2}}))$ can be approximated using a Taylor expansion. Assuming $\mu$ is smooth and using (59), we have

$$\mu(x_{j+\frac{1}{2}}, t^{n+\frac{1}{2}}) \approx \mu(x_j, t^n) + \frac{\Delta t}{2} \mu_t(x_j, t^n) + \frac{h_j}{2} \mu_x(x_j, t^n)$$  

where $\Delta t = t^{n+1} - t^n$ and $\mu_x = \frac{\partial \mu}{\partial x}$.

$$= \mu(x_j, t^n) + \frac{h_j}{2} (1 - \frac{\nu \Delta t}{h_j} \eta'(\mu(x_j, t^n))) \mu_x(x_j, t^n).$$  

Noting that $G(x_{j+\frac{1}{2}}, t^n)$ can be thought of as an approximation to $\nu \eta(\mu(x_{j+\frac{1}{2}}, t^{n+\frac{1}{2}}))$, we can set

$$G(x_{j+\frac{1}{2}}, t^n) = \nu \eta(\hat{\mu}^{n+\frac{1}{2}}_{j+\frac{1}{2}}), \quad j = 1, 2, \ldots,$$

where

$$\hat{\mu}^{n+\frac{1}{2}}_{j+\frac{1}{2}} = \hat{\mu}^n_j + \frac{h_j}{2} (1 - \frac{\nu \Delta t}{h_j} \eta'(\hat{\mu}^n_j)) \delta_x \hat{\mu}^n_j,$$

$$\hat{\mu}^n_j = C^n(x_j) + \phi(C^n(x_j)),$$

and $\delta_x \hat{\mu}^n_j$ is an approximate slope. The approximate slope can be determined by a slope limiting procedure outlined in [5]. Namely, we define $\delta_x \hat{\mu}^n_j$ by

$$\delta_x \hat{\mu}^n_j = \delta_{lim} \hat{\mu}^n_j \ast sgn(\hat{\mu}^n_{j+1} - \hat{\mu}^n_{j-1})$$

where

$$\delta_{lim} \hat{\mu}^n_j = \begin{cases} \min(\frac{\hat{\mu}^n_{j+1} - \hat{\mu}^n_j}{x_{j+1} - x_j}, \frac{\hat{\mu}^n_j - \hat{\mu}^n_{j-1}}{x_j - x_{j-1}}), & \text{if } (\hat{\mu}^n_{j+1} - \hat{\mu}^n_j) \ast (\hat{\mu}^n_j - \hat{\mu}^n_{j-1}) > 0, \\ 0, & \text{otherwise.} \end{cases}$$

The effect of the slope-limiting procedure is to produce a higher-order approximation to the solution, except where the solution has a local maximum or minimum. This allows for a better approximation to sharp fronts than that obtained by first-order upwinding, without introducing spurious oscillations into the numerical solution.

We note that, forcing $\delta_x \hat{\mu}^n_j = 0$, then $\hat{\mu}^{n+\frac{1}{2}}_{j+\frac{1}{2}} = \hat{\mu}^n_j$, and $\eta(\hat{\mu}^n_j) = C^n(x_j)$, which is simply the first-order upwinding scheme given above.
The semi-discrete analog of (63) is

\( \tilde{\mu}_{j+\frac{1}{2}}(t) = \tilde{\mu}_j(t) + \frac{h}{2} \delta_x \tilde{\mu}_j(t), \)

Substituting (66) into (62), we note that for \( \mu \) sufficiently smooth and \( \delta_x \tilde{\mu}_j(t) \neq 0 \), the estimate (33) is improved to

\[ |G(x_{j+\frac{1}{2}}, t) - g(x_{j+\frac{1}{2}}, t)| \leq K \left[ |(C - \bar{c})(x_j, t)| + h^2 \right], \]

To demonstrate the effects of a higher-order flux approximation, we consider the data

1. \( 0 < x < 1 \),
2. \( v = 1, D = 0 \),
3. \( \phi(c) = c^\frac{3}{2} \),
4. \( c_0(t) = 0, c_L(t) = 0 \).
with initial condition

\[
(68) \quad c^0(x) = \begin{cases} 
1, & .25 \leq x \leq .3, \\
0, & \text{otherwise}.
\end{cases}
\]

Simulation results at \( T = .8 \) for the different flux approximations are given in Figure 3. Here we used 20 and 100 mesh points. As seen in the figure, the height of the profile is better preserved for the higher-order flux approximation. The lower-order flux approximation gives a substantial amount of numerical diffusion, especially for 20 mesh points.

**Remark.** We conclude by briefly discussing the extension of the problem described above to multidimensions. In two and three spatial dimensions, the generalization of (8) is

\[
(69) \quad c_t + \phi(c)_t + \nabla \cdot (v c - D \nabla c) = 0,
\]

where \( v \) is a velocity vector, and \( D \) is a diffusion/dispersion tensor. The techniques described and analyzed above can be extended to this model. A multidimensional extension of the upwind-mixed method is described in [7]. Assuming similar smoothness assumptions on the solution, with \( D \) positive definite, one can extend the semi-discrete analysis to the multidimensional case by combining the techniques given above with techniques used to analyze the linear analog of (69) as given in [7]. As mentioned earlier, this scheme has been implemented and is being used to study the long-time behavior of solutions to (69) under various assumptions on \( \phi, v \) and \( D \). These results will be reported in [9].

**REFERENCES**


