AD-A070 194 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
EFFICIENT TIME-STEPPING PROCEDURES FOR MISCIBLE DISPLACEMENT PR--ETC(U)
MAR 79 R E EWING
UNCLASSIFIED MRC-TSR-1934 NL

1 OF 1
AD
A070194

END
DATE
FILMED
7-79
DDC

MRC Technical Summary Report #1934

EFFICIENT TIME-STEPPING PROCEDURES FOR MISCIBLE DISPLACEMENT PROBLEMS IN POROUS MEDIA

Richard E. Ewing

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

March 1979

(Received December 18, 1978)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, D. C. 20550

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

(6) EFFICIENT TIME-STEPPING PROCEDURES FOR MISCIBLE DISPLACEMENT

PROBLEMS IN POROUS MEDIA.

(10) Richard E. Ewing

(9) Technical Summary Report #1934
(11) March 1979 (12) 44p.
(14) MRC-TSR-1934

ABSTRACT

A model system of equations which has been used to describe the miscible displacement of one incompressible fluid by another in a porous medium is the coupled quasilinear system for $c = c(x,t)$ and $p = p(x,t)$ for $x \in \Omega$, $t \in (0,T]$ given by

$$\nabla \cdot [a(x,c)\{\nabla p - \gamma(x,c)\nabla g\}] \equiv -\nabla \cdot u = f_1(x) ,$$

$$\nabla \cdot [b(x,t)\nabla c] - u(x,c,\nabla p) \cdot \nabla c = \varphi(x) \frac{\partial c}{\partial t} - f_2(x,c) ,$$

with appropriate initial and Neumann boundary conditions. Another case considered is when $b = b(x,c,\nabla p)$ above. Iterative methods are presented and analyzed which are based on using a preconditioned conjugate gradient iteration for approximately solving the systems of linear equations produced at each time step by Galerkin methods for time-stepping the above system. Optimal order convergence rates are obtained for the iterative methods. The iterative methods are computationally more efficient than Galerkin methods previously proposed to solve the above system. The use of different time increments in the time-stepping procedures for the different variables is also presented and analyzed. The use of unequal time increments takes advantage of different smoothnesses in time of the physical variables p and c and greatly reduces the work done in the computation of the approximate solution.

AMS (MOS) Subject Classifications: 65M15, 65N15, 65N30, 76.35

Key Words: Galerkin methods, Error estimates, Iterative methods,
Conjugate gradient methods, Fluid flow

Work Unit Number 7 (Numerical Analysis)

(15)

Sponsored by the United States Army under Contract Numbers DAAG29-75-C-0024 and DAAG29-78-G-0161. ~~This material is based upon work supported by the~~ National Science Foundation under Grant Number MCS78-09525.

-page -A-

SIGNIFICANCE AND EXPLANATION

The numerical approximation of a system of partial differential equations used to model the miscible displacement of one incompressible fluid by another in porous media is considered. For example, the model system has been used to describe the variables of pressure and the changing concentration of a chemical solvent in oil used to flood oil wells to push the oil through the porous media toward production wells to get greater recovery of oil from underground reservoirs.

Wheeler and the author have recently presented some numerical methods for approximating the solution of the model system and have obtained optimal order error estimates for these methods. This paper presents and analyzes methods which are computationally more efficient than earlier methods.

In time-stepping problems with time-dependent coefficients, numerical methods produce different systems of linear equations which must be solved at each time step. Factoring a different matrix at each time step in the solution process can be very computationally expensive since the matrices are frequently of the order $1,000 \times 1,000$. The methods presented in this paper require the factorization of only two matrices. Then an iterative procedure which compensates for the fact that new matrices are not factored at each time step is used to obtain approximate solutions. Significant amounts of computation are saved by these methods. Optimal order estimates on rates of convergence for these methods are obtained.

EFFICIENT TIME-STEPPING PROCEDURES FOR MISCIBLE DISPLACEMENT PROBLEMS IN POROUS MEDIA

Richard E. Ewing

1. Introduction

In [10] numerical approximations by Galerkin methods to a problem arising in the miscible displacement of one incompressible fluid by another in a porous medium are presented and analyzed. A brief discussion of the physical problem is given in [10] and a mathematical model which is sufficiently general to incorporate most of the major features of the physical problem (see [10,15,17]) is presented. In this paper, we shall present and analyze some methods for time-stepping these model equations which are much more efficient computationally than the methods discussed in [10]. We shall use an iterative method based on a preconditioned conjugate gradient iteration to approximate the solution of the systems of linear equations which arise from a Galerkin approximation of the model equations. The iterative methods presented preserve the accuracy inherent in the underlying Galerkin method and let us obtain very nearly optimal possible orders for the work involved in solving the linear systems of equations.

We first present the model equations for our physical problem. Find $c = c(x,t)$ and $p = p(x,t)$, solutions of

$$\nabla\cdot[a\{\nabla p - \gamma\nabla g\}] \equiv -\nabla\cdot u = f_1 , \tag{1.1}$$

$$\nabla\cdot[b\nabla c] - u\cdot\nabla c = \phi\frac{\partial c}{\partial t} - f_2 , \tag{1.2}$$

for $x \in \Omega$, $t \in J \equiv (0,T]$, where Ω is a bounded domain in $\mathbb{R}^d$, $d = 2$, with boundary $\partial\Omega$ and $u(x,c,\nabla p) = -a\{\nabla p - \gamma\nabla g\}$ is a vector in $\mathbb{R}^2$. Here $a = a(x,c)$, $\gamma = \gamma(x,c)$, $g = g(x)$, $f_1 = f_1(x)$, $\phi = \phi(x)$, $f_2 = f_2(x,c)$, and b are specified. For ease of exposition, we shall consider two cases for b: Case I, with $b = b(x,t)$, and Case II, with $b = b(x,c,p_x,p_y)$. The cases of greatest physical interest (see 15, 17) are Case II and a subcase of Case I with $b = b(x)$ (special results for this subcase are obtained in Corollary 4.2). We assume that the following boundary and initial conditions hold:

Sponsored by the United States Army under Contract Numbers DAAG29-75-C-0024 and DAAG29-78-G-0161. This material is based upon work supported by the National Science Foundation under Grant Number MCS78-09525.

$$(1.3) \qquad a\left\{\frac{\partial p}{\partial \nu} - \gamma\frac{\partial q}{\partial \nu}\right\} = 0\,, \qquad x \in \partial\Omega,\quad t \in J\,,$$

$$(1.4) \qquad b\,\frac{\partial c}{\partial \nu} = 0\,, \qquad x \in \partial\Omega,\quad t \in J\,,$$

$$(1.5) \qquad c(x,0) = c_0(x)\,, \qquad x \in \Omega\,,$$

where $\frac{\partial F}{\partial \nu}$ is the normal derivative of F on the boundary of Ω. Note that (1.1)-(1.5) will define $p(x,t)$ only to within an arbitrary constant. We shall normalize p by the condition that

$$(1.6) \qquad \frac{1}{|\Omega|}\int_\Omega p(x,t)dx = 1, \qquad t \in J\,,$$

where $|\Omega|$ is the measure of the domain Ω.

We note that the analysis that follows would easily treat forcing functions f_1 and f_2 which are smoothly distributed over Ω. If singular functions are used to model the effect of small injection and production wells, the analysis will fail. Thus we shall make the assumption that for our problem, the sources and sinks are smoothly distributed and shall then, without loss of generality, assume that $f_1 \equiv 0$ and $f_2 \equiv 0$ *for the remainder of the paper.*

Continuous time Galerkin approximations for (1.1)-(1.5) are presented and analyzed in [10]. By lagging or extrapolating the coefficients in discrete-time versions of these methods, we are able to linearize and uncouple the systems of equations required in the approximations. However, the discrete-time versions of these methods require that different systems of linear equations be solved at each time step; this is a computationally expensive process. In this paper, we shall present and analyze iterative methods which require the factorization of only two matrices for the total solution process. The use of iterative methods to approximate the solution of the linear equations arising from the parabolic equation is an extension of the techniques developed in [6,9] for quasilinear time-dependent problems. The use of iterative methods to approximate the solution of an elliptic, basically time-independent equation, only to within the accuracy of the time truncation error from an associated time-dependent problem seems to be new. We emphasize that unlike standard iterative procedures for elliptic equations, all the methods presented will only require a number of iterations which is independent of Δt, h, and n to obtain a norm reduction sufficient to feed the associated parabolic problem. Although a preconditioned conjugate gradient iterative method will be presented, any method which achieves the specified norm reductions will suffice in all of the analysis to follow.

In the physical problem which motivates our consideration of (1.1)-(1.5), the pressure p is much smoother in time than the concentration c. In order to take advantage of this difference in smoothness of p and c, we shall use different time increments Δt_c and Δt_p in our analysis for the time-stepping of the systems of equations arising from the equations for concentration and pressure. By using these unequal time increments, we shall need to update the pressure variable much less frequently than the concentration variable and thus avoid considerable unnecessary computation.

In Section 2 we introduce two families of finite element spaces which we use to approximate our unknown functions p and c. We present the hypotheses on (1.1)-(1.5) and the solution (c,p), discuss elliptic projections for c and for p, and present our basic Galerkin approximation of (1.1)-(1.5) together with several modifications. In Section 3 we present our preconditioned conjugate gradient modifications of the methods described in Section 2 and analyze the effect of the iterative approximation on a single time step. In Section 4 we obtain global error estimates for the various methods described in Sections 2 and 3. Section 5 contains a brief discussion of the estimates of the work of computation for the methods presented in this paper.

2. Preliminaries and Description of Basic Galerkin Approximations

Let $(u,v) = \int_\Omega uv dx$, $\langle u,v\rangle = \int_{\partial\Omega} uv d\sigma$, $\|u\|^2 = (u,u)$, and $|u|^2 = \langle u,u\rangle$. Let $W_s^k(\Omega)$ be the Sobolev space on Ω with norm

$$(2.1) \qquad \|\psi\|_{W_s^k} = \left(\sum_{|\alpha|\le k} \left\|\frac{\partial^\alpha \psi}{\partial x^\alpha}\right\|^s_{L^s(\Omega)} \right)^{\frac{1}{s}},$$

with the usual modification for $s = \infty$. When $s = 2$, denote $\|\psi\|_{W_2^k} \equiv \|\psi\|_{H^k} \equiv \|\psi\|_k$.

If $\nabla F = (F_1, F_2)$, write $\|\nabla F\|_{W_s^k}$ in place of $(\|F_1\|^s_{W_s^k} + \|F_2\|^s_{W_s^k})^{\frac{1}{s}}$. Also $H^s(\partial\Omega)$ will denote the Sobolev space on $\partial\Omega$.

Let $\{M_h\}$ be a family of finite-dimensional subspaces of $H^1(\Omega)$ with the following property:

For $p = 2$ or ∞, there exist an integer $r \ge 2$ and a constant K_0 such that, for $1 \le q \le r$ and $\psi \in W_p^q(\Omega)$,

$$(2.2) \qquad \inf_{\chi\in M_h} \{\|\psi - \chi\|_{W_p^0} + h\|\psi - \chi\|_{W_p^1}\} \le K_0 \|\psi\|_{W_p^q} h^q .$$

We also define a family of finite-dimensional subspaces of $H^1(\Omega)$ called $\{N_h\}$ which satisfies the same property as $\{M_h\}$ with r replaced by s. We also assume that the families $\{M_h\}$ and $\{N_h\}$ satisfy the following so-called "inverse hypotheses": if $\varphi \in M_h$ and $\psi \in N_h$,

$$(2.3) \qquad \begin{aligned} &a) \quad \|\varphi\|_{L^\infty(\Omega)} \le K_0 h^{-\frac{d}{2}} \|\varphi\| = K_0 h^{-1} \|\varphi\| , \\ &b) \quad \|\nabla\psi\|_{L^\infty(\Omega)} \le K_0 h^{-1} \|\nabla\psi\| , \\ &c) \quad \|\varphi\|_1 \le K_0 h^{-1} \|\varphi\| . \end{aligned}$$

Restrict Ω as follows (with (S) denoting the collection of restrictions):

(S): 1. Ω is H^2-regular, i.e., if

$$-\Delta v + \theta v = \zeta , \qquad x \in \Omega, \quad \theta = 0 \text{ or } 1 ,$$

$$\frac{\partial v}{\partial \nu} = \eta , \qquad x \in \partial\Omega ,$$

and $$(\zeta,1) + \langle \eta,1 \rangle = 0 \quad \text{if} \quad \theta = 0 ,$$

then $\|v\|_2 \le K(\Omega)\{\|\zeta\| + \|\eta\|_{H^{1/2}(\partial\Omega)}\}$;

2. $\partial\Omega$ is Lipschitz.

For the following assumptions, we shall restrict the variable q_1 to lie between two physically determined constants, i.e.

(2.4) $$-M_* \le q_1 \le M^*, \quad M_* > 0, \quad M^* > 1 ,$$

(e.g., $M_* = \hat{\epsilon}$ and $M^* = 1 + \hat{\epsilon}$, for some $\hat{\epsilon} > 0$ (see [10])). Assume the following regularity for a, γ, b, u, and φ:

(Q): 1. There exist uniform constants such that

(2.5)

a) $0 < a_* \le a(x,q_1) \le a^* \le K_1$,

b) $0 < \varphi_* \le \varphi(x) \le K_1$,

c) $|\gamma(x,q_1)| \le K_1$,

d) $|\nabla q| \le K_1$,

e) $0 < b_* \le b(x,t)$ (case I) ,

f) $|u_i(x,q_1,q_2)| \le K_1(1 + |q_2|)$, $\quad i = 1,2, \quad q_2 \in \mathbb{R}$.

2. Let the derivatives of $a = a(x,c)$, $b = b(x,t)$, $\gamma = \gamma(x,c)$, and $u_i = u_i(x,c,q)$ satisfy the following assumptions: for $i = 1,2$, and $q_2 \in \mathbb{R}$,

(2.6) $$\left|\frac{\partial a}{\partial x_i}(x,q_1)\right| + \left|\frac{\partial a}{\partial c}(x,q_1)\right| + \left|\frac{\partial^2}{\partial x_i \partial c} a(x,q_1)\right| + \left|\frac{\partial^2 a}{\partial c^2}(x,q_1)\right|$$
$$+ \left|\frac{\partial \gamma}{\partial c}(x,q_1)\right| + \left|\frac{\partial b}{\partial t}(x,t)\right| + \left|\frac{\partial u_i}{\partial c}(x,q_1,q_2)\right| \le M .$$

3. In Case II, when $b \equiv b(x,c,\nabla p) \equiv b(x,c,p_x,p_y)$, for $q_2,q_3 \in \mathbb{R}$,

a) $0 < b_* \le b(x,q_1,q_2,q_3)$,

b) $b(x,q_1,p_x,p_y) \le M$,

(2.7) c) $$\left|\frac{\partial b}{\partial c}(x,q_1,q_2,q_3)\right| + \left|\frac{\partial b}{\partial p_x}(x,q_1,q_2,q_3)\right| + \left|\frac{\partial b}{\partial p_y}(x,q_1,q_2,q_3)\right|$$

$$+ \left|\frac{\partial^2 b}{\partial c^2}(x,q_1,q_2,q_3)\right| \le M .$$

Define

(2.8) $$\|\psi\|_{W_p^q((a,b);X)} \equiv \big\| \|\psi(\cdot,t)\|_X \big\|_{W_p^q(a,b)} , \qquad 1 \le p,\ q \le \infty .$$

Let (p,c), the solution of (1.1)-(1.5), satisfy the following regularity assumptions:

(R): a) $$\|c\|_{L^\infty(J;H^r)} + \left\|\frac{\partial c}{\partial t}\right\|_{L^2(J;H^r)} \le K_2 ,$$

b) $$\|p\|_{L^\infty(J;H^s)} + \left\|\frac{\partial p}{\partial t}\right\|_{L^2(J;H^s)} \le K_2 ,$$

(2.9) c) $$\|c\|_{L^\infty(J;H^3)} + \left\|\frac{\partial c}{\partial t}\right\|_{L^\infty(J;H^{2+\epsilon})} + \left\|\frac{\partial c}{\partial t}\right\|_{L^2(J;W_3^2)} \le K_2 \quad \text{for some } \epsilon > 0 ,$$

d) $$\|p\|_{L^\infty(J;H^3)} + \left\|\frac{\partial p}{\partial t}\right\|_{L^\infty(J;W_3^2)} \le K_2 ,$$

e) $$\left\|\frac{\partial^2 p}{\partial t^2}\right\|_{L^\infty(J;H^1)} + \left\|\frac{\partial^2 c}{\partial t^2}\right\|_{L^\infty(J;H^1)} \le K_2 .$$

In our analysis, we shall use a couple of auxiliary elliptic problems. This technique was used by Wheeler in [18]. Let $\tilde{p} \in N_h$ be the elliptic projection of p into N_h defined by

(2.10) $$(a(\cdot,c(\cdot,t))\nabla\tilde{p},\nabla v) = (a(\cdot,c(\cdot,t))\nabla p,\nabla v)$$

$$= (a(\cdot,c(\cdot,t))\gamma(\cdot,c(t))\nabla q,\nabla v), \qquad v \in N_h ,$$

for each $t \in J$, where

$$(2.11)\qquad \frac{1}{|\Omega|}\int_{\Omega}(\tilde{p}(x,t) - p(x,t))dx = 0, \quad \text{for each} \quad t \in J ,$$

and where (p,c) is the solution of (1.1)-(1.5). The restrictions (S) imply the following result [7,18].

Lemma 2.1: There exists a constant $K_3 = K_3(\Omega,a_*,K_0,K_1,K_2)$ such that for each $t \in J$,

$$(2.12)\qquad \|p - \tilde{p}\| + h\left\{\|\nabla(p - \tilde{p})\| + \left\|\nabla \frac{\partial(p - \tilde{p})}{\partial t}\right\|\right\} \le K_3 h^s\{\|p\|_s + \|\frac{\partial p}{\partial t}\|_s\} .$$

For Case II, let $\lambda > 0$ be taken sufficiently large that the bilinear form

$$B(\psi_1,\psi_2) \equiv (b(c,\nabla p)\nabla\psi_1,\nabla\psi_2) + (u(c,\nabla p)\cdot\nabla\psi_1,\psi_2) + \lambda(\psi_1,\psi_2)$$

is coercive over $H^1(\Omega)$. Then let $\tilde{c} \in M_h$ be the elliptic projection of c into M_h defined by

$$(2.13)\qquad B(\tilde{c},w) = B(c,w) = -(\varphi \frac{\partial c}{\partial t}, w) + \lambda(c,w), \qquad w \in M_h ,$$

for each $t \in J$. For Case I, where $b = b(x,t)$, the coefficients $b(c,\nabla p)$ in $B(\psi_1,\psi_2)$ are replaced by $b(t)$. Then, as in [7,18] we can obtain the following lemma.

Lemma 2.2: There exists a constant $K_4 = K_4(\Omega,b_*,\lambda,K_0,K_1,K_2,M)$ such that for $\ell = 0$ or 1,

$$(2.14)\qquad \begin{aligned} &\text{a)}\ \|\tilde{c} - c\|_{L^2(J;H^\ell)} + \left\|\frac{\partial(\tilde{c} - c)}{\partial t}\right\|_{L^2(J;H^\ell)} \le K_4 h^{r-\ell}\{\|c\|_{L^2(J;H^r)} + \|\frac{\partial c}{\partial t}\|_{L^2(J;H^r)}\} , \\ &\text{b)}\ \|\tilde{c} - c\|_{L^\infty(J;L^2)} \le K_4 h^r \|c\|_{L^\infty(J;H^r)} . \end{aligned}$$

We also make the assumptions on $\{M_h\}$, $\{N_h\}$, c, and p that there exists a constant K_5 such that

$$(2.15)\qquad \begin{aligned} &\|\nabla\tilde{p}\|_{L^\infty(J;L^\infty)} + \|\frac{\partial\tilde{p}}{\partial t}\|_{L^\infty(J;H^1)} + \|\tilde{c}\|_{L^\infty(J;W^1_\infty)} + \|\frac{\partial\tilde{c}}{\partial t}\|_{L^\infty(J;H^1)} \\ &\quad + \|\frac{\partial\tilde{c}}{\partial t}\|_{L^1(J;W^2_\infty)} \le K_5 . \end{aligned}$$

Sufficient conditions for (2.15) can be found, as in [6] and [18]. Finally, as in [2,6,7], we can obtain the following lemma.

Lemma 2.3: There exists a constant K_6, depending upon K_0, K_1, and K_2 such that

$$(2.16) \qquad \left\|\frac{\partial^2 \tilde{c}}{\partial t^2}\right\|_{L^\infty(J;H^1)} + \left\|\frac{\partial^2 \tilde{p}}{\partial t^2}\right\|_{L^\infty(J;H^1)} \le K_6 .$$

The continuous time approximation of (1.1)-(1.5) can be defined as follows: let $P : [0,T] \to N_h$, the approximation for p, and $C : [0,T] \to M_h$, the approximation for c, be defined by (supressing the dependence of the coefficients on x)

$$(2.17) \qquad (a(C)\nabla P,\nabla v) = (a(C)\gamma(C)\nabla g,\nabla v), \qquad v \in N_h ,$$

and

$$(2.18) \qquad (b(t)\nabla C,\nabla w) + (u(C,\nabla P)\cdot\nabla C,w) = -(\varphi \frac{\partial C}{\partial t}, w), \qquad w \in M_h ,$$

with

$$(2.19) \qquad (\varphi(\cdot)C(\cdot,0),w) = (\varphi(\cdot)\tilde{c}(\cdot,0),w), \qquad w \in M_h ,$$

where $\tilde{c}$ is the elliptic projection of c defined in (2.13). In [10] it was shown that the appropriate assumptions from (R) and (Q) yield:

Case I: $(b = b(x,t))$

$$(2.20) \qquad h\|\nabla(P - p)\|_{L^\infty(J;L^2)} + \|C - c\|_{L^\infty(J;L^2)} + h\|\nabla(C - c)\|_{L^2(J;L^2)} \le K_7(h^r + h^s + h^{r+s-3}) ,$$

Case II: $(b = b(x,c,\nabla p))$

$$(2.21) \qquad h\|\nabla(P - p)\|_{L^\infty(J;L^2)} + \|C - c\|_{L^\infty(J;L^2)} + h\|\nabla(C - c)\|_{L^2(J;L^2)} \le K_8(h^r + h^{s-1}) .$$

Let $\Delta t > 0$, $N = T/\Delta t \in \mathbb{Z}$, and $t^\sigma = \sigma\Delta t$, $\sigma \in \mathbb{R}$. Also let $\psi^n \equiv \psi^n(x) \equiv \psi(x,t^n)$, and $d_t\psi^n \equiv (\psi^{n+1} - \psi^n)/\Delta t$.

In [10], the following discrete time approximation with discretization error of the order Δt was analyzed and an extrapolated coefficient Crank-Nicolson-Galerkin scheme was stated to have $(\Delta t)^2$ time discretization error. Denote the approximation of p by $W : \{0 = t_0,t_1,\ldots,t_N = T\} \to N_h$ and the approximation of c by

$Z : \{0 = t_0, t_1, \ldots, t_N = T\} \to M_h$. Assuming that Z^n and W^n are known, we determine Z^{n+1} and then W^{n+1} as follows:

$$(2.22)\qquad \left(\varphi\,\frac{Z^{n+1} - Z^n}{\Delta t},\chi\right) + (b(t^{n+1})\nabla Z^{n+1},\nabla\chi) = -(u(Z^n,\nabla W^n)\cdot\nabla Z^{n+j},\chi), \qquad \chi \in M_h ,$$

where $j = 0$ or 1 and

$$(2.23)\qquad (Z^0,\chi) = (c_0(\cdot),\chi), \qquad \chi \in M_h$$

and

$$(2.24)\qquad (a(Z^{n+1})\nabla W^{n+1},\nabla y) = (a(Z^{n+1})\gamma(Z^{n+1})\nabla g,\nabla y), \qquad y \in N_h .$$

We note that the coefficient matrix arising from the algebraic system (2.22) with $j = 0$ is symmetric. However, in many problems, the transport term is large compared to the diffusion term and it may be numerically advantageous to use (2.22) with $j = 1$ even though the coefficient matrix is no longer symmetric. The remainder of this paper will consider the case $j = 0$.

Since W^{n+1} does not appear in (2.22), we can separate (2.22) and (2.24) by first solving (2.22) at time t^{n+1} and using that solution in the coefficients for the solution of (2.24) at the time t^{n+1}. In this way we have uncoupled (2.22) and (2.24) and now must only solve two separate linear systems. This greatly reduces the size of our problem and, correspondingly, the work needed to obtain a solution.

In the physical problem which motivates our consideration of (1.1)-(1.5), the pressure p is much smoother in time than the concentration c. Thus W^{n+1} from (2.24) does not differ radically from the W^n determined at the previous time level. This motivates various modifications of (2.22)-(2.24) where equation (2.24) is solved only at every k^{th} time step with k determined from the relative smoothnesses in time of p and c. We shall consider several of these modifications.

Let $\Delta t_1 = k\Delta t$ with k to be chosen later in different ways. Consider Δt_1-time levels, which coincide with Δt-time levels but have separation Δt_1 instead of Δt. Let $D^* = D^*(c,p)$ be a particular ratio of the norms which enter into the truncation errors for various derivatives of c and p. (We shall describe this ratio in more detail for specific examples later.)

One modification of (2.22)-(2.24) is to choose k such that

(2.25) $$k \approx (D^{*})^{-1} .$$

Then

(2.26) $$\Delta t_1 D^{*} = k\Delta t D^{*} \approx \Delta t .$$

In this case, letting $[\![n/k]\!]$ be the greatest integer less than or equal to n/k, we replace (2.22) (e.g. in Case II where $b = b(x,c,\nabla p)$) by

(2.27) $$\left(\varphi \frac{Z^{n+1} - Z^{n}}{\Delta t}, \chi\right) + (b(Z^{n},\nabla W^{[\![n/k]\!]})\nabla Z^{n+1},\nabla\chi) = -(u(Z^{n},\nabla W^{[\![n/k]\!]})\cdot\nabla Z^{n},\chi), \qquad \chi \in M_h .$$

Then W^{k} is determined by (2.24) only at the Δt_1 time levels. The errors made in the coefficients by lagging the pressure in this fashion is $O(\Delta t_1 D^{*}) = O(\Delta t)$. Another modification which has the same order truncation error, suggested by Todd Dupont, is to replace $\nabla W^{[\![n/k]\!]}$ by $\nabla EW^{[\![n/k]\!]} \equiv \frac{3}{2}\nabla W^{[\![n/k]\!]} - \frac{1}{2}\nabla W^{[\![n/k]\!]-1}$ in the coefficients of (2.27). This linear extrapolation to the midpoint of the current Δt_1-time interval would be as easy as (2.27) to implement (once two values at Δt_1-time levels have been determined) and probably more accurate (especially for Δt-time levels near the upper end of the Δt_1-time interval).

A more accurate modification of (2.22) can be defined by evaluating the coefficients b and u in (2.27) at appropriate linear extrapolations of the ∇W^{k} from the two previous Δt_1-time levels to the Δt-time level t^{n+1}. For example, we can write $n = [\![n/k]\!] + \nu/k$ for some $\nu = 0,1,\ldots,k-1$. Then for $n \geq k$ and ν defined above, we define $E_k^{\nu}F^{n}$ to be the linear extrapolation for the time level t^{n+1} from the values of F at the two previous Δt_1 time levels. Thus for $\nu = 0,1,\ldots,k-1$, and $\theta = (\nu+1)/k$, we define

(2.28) $$E_k^{\nu}F^{n} \equiv \begin{cases} (1+\theta)F^{[\![n/k]\!]} - \theta F^{[\![n/k]\!]-1}, & n \geq k \\ F^{0}, & n < k . \end{cases}$$

Then if $\left\|\frac{d^2F}{dt^2}\right\|_{L^2(0,T;L^2)} \leq K$, we are making an error of $O((\Delta t_1)^2)$ by approximating

$F(t^{n+1})$ by $E_k^{\nu}F^n$. This clearly defines k different extrapolations E_k^{ν} to be repeated in sequence between two of the Δt_1-time levels.

Using this extrapolation, if we choose k such that

$$(2.29) \qquad k \approx (\Delta t D^*)^{-\frac{1}{2}} ,$$

then we see that

$$(2.30) \qquad (\Delta t_1)^2 D^* = k^2 (\Delta t)^2 D^* \approx \Delta t .$$

Clearly (2.29) allows k to be chosen quite large and therefore W^k must be computed much less frequently than before.

With the notation of (2.28), our new modification of (2.22) (in Case I when $b = b(x,t)$) can be described as

$$(2.31) \quad (\varphi d_t Z^n, \chi) + (b(t^{n+1})\nabla Z^{n+1}, \nabla\chi) = -(u(Z^n, E_k^{\nu}\nabla W^n)\cdot\nabla Z^n, \chi), \qquad \chi \in M_h, \quad n \geq k .$$

In Case II, replace $b(t^{n+1})$ by $b(Z^n, E_k^{\nu}\nabla W^n)$. Again W^k is then determined only at the Δt_1-time levels by (2.24).

We also consider another modification of both (2.27) and (2.31). We noted in (Q) that if the argument corresponding to c in our coefficients satisfies (2.4), then the coefficients satisfy the bounds in (Q). Since c satisfies $0 \leq c \leq 1$ (and thus (2.4)) and since Z^n is an approximation to c^n, when evaluating the coefficients, if $Z^n < 0$, we replace the argument for Z^n by 0 and if $Z^n > 1$, we replace the argument for Z^n by 1. This type of truncation of arguments has been discussed earlier in [4]. The analysis of the error that this truncation causes will be contained in the proofs of our major results. We shall use the notation $u(Z^{*n}, \nabla W^{[\![n/k]\!]})$ in (2.27) to note that the Z^n argument of u has been truncated to lie between 0 and 1. In the same manner, the modification of (2.31) with truncated coefficients is written as

$$(2.32) \qquad (\varphi d_t Z^n, \chi) + (b(t^{n+1})\nabla Z^{n+1}, \nabla\chi) = -(u(Z^{*n}, E_k^{\nu}\nabla W^n)\cdot\nabla Z^n, \chi), \qquad \chi \in M_h .$$

We note that Crank-Nicolson-Galerkin methods can be defined as another modification of (2.32) which will have $O((\Delta t)^2)$ time truncation errors. See [6,9,10,16] for methods of this form.

3. Approximate Solution of the Linear Equations by Iteration

In this section we shall present the linear equations arising from (2.32) and the corresponding (2.24). (We note that by replacing $E_k^\nu \nabla W^n$ by $\nabla W^{[\![n/k]\!]}$ in the coefficients of b and u in what follows will yield the linear equations arising from (2.27).) We also present an iterative method for approximating their solution.

The conjugate gradient procedure presented here provides only one example of the possible modifications of (2.32) and (2.24) that fall under the analysis given in the next section. Any method which provides the norm reduction defined in this section will preserve the results of Section 4.

Let $\{\mu_i\}_{i=1}^{M_1}$ be a basis for M_h and $\{\psi_i\}_{i=1}^{M_2}$ be a basis for N_h. We then denote the solution of (2.32) and (2.24) by (Z^m, W^m), where

$$\text{a)}\quad Z^m = \sum_{i=1}^{M_1} \xi_i^m \mu_i ,$$

(3.1)

$$\text{b)}\quad W^m = \sum_{i=1}^{M_2} \omega_i^m \psi_i , \quad \text{for } m \text{ such that } m/k \in \mathbb{Z} .$$

We next define several matrices and vectors. For Case II let

$$\text{a)}\quad \Phi = (\phi_{ij}) = ((\varphi \mu_j, \mu_i)) ,$$

$$\text{b)}\quad B^m(\theta,\sigma) = (b_{ij}^m(\theta,\sigma)) = ((b((\sum_{\ell=1}^{M_1} \theta_\ell^m \mu_\ell)^*, E_k^\nu \sum_{\ell=1}^{M_2} \sigma_\ell^m \nabla\psi_\ell)\nabla\mu_j, \nabla\mu_i)) ,$$

$$\text{c)}\quad U^m(\theta,\sigma) = (u_{ij}^m(\theta,\sigma)) = ((u((\sum_{\ell=1}^{M_1} \theta_\ell^m \mu_\ell)^*, E_k^\nu \sum_{\ell=1}^{M_2} \sigma_\ell^m \nabla\psi_\ell)\cdot\nabla\mu_j, \mu_i)) ,$$

(3.2)

$$\text{d)}\quad A^m(\theta) = (a_{\alpha\beta}^m(\theta)) = ((a((\sum_{\ell=1}^{M_1} \theta_\ell^m \mu_\ell)^*)\nabla\psi_\beta, \nabla\psi_\alpha)) ,$$

$$\text{e)}\quad \Gamma^m(\theta) = (\gamma_\alpha^m(\theta)) = ((a((\sum_{\ell=1}^{M_1} \theta_\ell^m \mu_\ell)^*)\gamma((\sum_{\ell=1}^{M_1} \theta_\ell^m \mu_\ell)^*)\nabla q, \nabla\psi_\alpha)) ,$$

$$\text{f)}\quad B_0 = ((b_0 \nabla\mu_j, \nabla\mu_i)) ,$$

$$\text{g)}\quad A_0 = ((a_0 \nabla\psi_\beta, \nabla\psi_\alpha)) .$$

for $i,j = 1,\ldots,M_1$ and $\alpha,\beta = 1,\ldots,M_2$. The matrices for Case I are correspondingly simpler. Here b_0 and a_0 can be chosen in a very arbitrary way. A good choice might be $b_0 = b(x,c_0(x),\nabla W^0)$ and $a_0 = a(x,c_0(x))$ or, if average values $\bar{c}$ and $\nabla\bar{p}$ are more or less known, use these values in the coefficients to replace c_0 and ∇W^0.

We can write (2.32) and (2.24) in the form

$$
\begin{aligned}
L^n(\xi)(\xi^{n+1} - \xi^n) &\equiv (\Phi + \Delta t B^n(\xi,\omega))(\xi^{n+1} - \xi^n) \\
&= -\Delta t B^n(\xi,\omega)\xi^n + \Delta t U^n(\xi,\omega)\xi^n, \qquad n \geq 1 ,
\end{aligned} \tag{3.3}
$$

and for m such that $m/k \in \mathbb{Z}$,

$$A^m(\xi)\omega^m = \Gamma^m(\xi) . \tag{3.4}$$

We shall not solve (3.3) and (3.4) exactly; instead we shall use a pre-determined number of preconditioned conjugate gradient [1,5,6,8,9] iterations to advance the solution one time step. The preconditioning matrices will be chosen to be independent of n. Specifically, we shall use

$$L_0 \equiv \Phi + \Delta t B_0 \tag{3.5}$$

and A_0 for the preconditioners for (3.3) and (3.4) respectively.

Denote by

$$
\begin{aligned}
&\text{a)} \quad C^m = \sum_{i=1}^{M_1} \alpha_i^m \mu_i \quad \text{and} \\
&\text{b)} \quad P^m = \sum_{i=1}^{M_2} \beta_i^m \psi_i , \quad \text{where} \quad m/k \in \mathbb{Z} ,
\end{aligned} \tag{3.6}
$$

the approximations to Z^m and W^m, respectively, produced by only approximately solving (3.3) and (3.4). A starting procedure for obtaining C^0 and P^0 will be discussed later. Assuming that these quantities are known, we shall find α^{n+1} (and thus C^{n+1}) using a preconditioned conjugate gradient iteration to approximate $\xi^{n+1} - \xi^n$ from (3.3). We shall use different initial guesses for $\xi^{n+1} - \xi^n$ for $n = 0$ and for $n \geq 1$. We shall use linear extrapolation for $n \geq 1$. Specifically, we initialize our iteration for (3.3) as follows:

$$
\text{(3.7)}\quad
\begin{aligned}
&\text{a)}\ \begin{cases} n = 0: & x_0 \equiv x_0^1 = 0 \\ n \geq 1: & x_0 \equiv x_0^{n+1} = \alpha^n - \alpha^{n-1}\ , \end{cases}\\
&\text{b)}\quad n \geq 0:\ \ q_0 \equiv q_0^{n+1} = L_0 s_0 \equiv L_0 s_0^{n+1}\\
&\qquad = L^n(\alpha,\beta)x_0 + \Delta t B^n(\alpha,\beta)\alpha^n - \Delta t U^n(\alpha,\beta)\alpha^n\ .
\end{aligned}
$$

Then, using the initialization x_0, q_0, and s_0 from (3.7), for $j = 1,2,\ldots,\kappa_1 - 1$, where the number of iterations κ_1 will be chosen later, independently of n, set

$$
\text{(3.8)}\quad
\begin{aligned}
&\text{a)}\ x_{j+1} = x_j + \theta_{1j}s_j,\ \text{ where }\ \theta_{1j} = \frac{-(L_0^{-1}q_j,q_j)_e}{(s_j,L^n(\alpha,\beta)s_j)_e}\ ,\\
&\text{b)}\ q_{j+1} = q_j + \theta_{1j}L^n(\alpha,\beta)s_j\ ,\\
&\text{c)}\ s_{j+1} = L_0^{-1}q_{j+1} + \theta_{2j}s_j,\ \text{ where }\ \theta_{2j} = \frac{(L_0^{-1}q_{j+1},q_{j+1})_e}{(L_0^{-1}q_j,q_j)_e}\ ,
\end{aligned}
$$

where $(\cdot,\cdot)_e$ is the Euclidean inner product. Finally, set

$$
\text{(3.9)}\qquad \alpha^{n+1} = \alpha^n + x_{\kappa_1}\ .
$$

In a similar fashion, we define the iterative approximation of (3.4) as follows:

$$
\text{(3.10)}\quad
\begin{aligned}
&\text{a)}\ \begin{aligned} m = 1:&\ \ \tilde{x}_0 \equiv \tilde{x}_0^k = \beta^0\ ,\\ m \geq 2:&\ \ \tilde{x}_0 \equiv \tilde{x}_0^{2k} = 2\beta^k - \beta^0\ , \end{aligned}\\
&\text{b)}\ \begin{aligned} m \geq 0:&\ \ \tilde{q}_0 \equiv \tilde{q}_0^{(m+1)k} = A_0\tilde{s}_0 \equiv A_0\tilde{s}_0^{(m+1)k}\\ &= A^{mk}(\alpha)\tilde{x}_0 - \Gamma^{mk}(\alpha)\ . \end{aligned}
\end{aligned}
$$

Using (3.10) as an initialization for $\tilde{x}_0$, $\tilde{q}_0$, and $\tilde{s}_0$, for $j = 1,2,\ldots,\kappa_2 - 1$, where the number of iterations κ_2 will be chosen later, independently of m and k, determine $\tilde{x}_{j+1}$, $\tilde{q}_{j+1}$, and $\tilde{s}_{j+1}$ from (3.8) with L_0 and $L^n(\alpha)$ everywhere replaced by A_0 and $A^{mk}(\alpha)$ respectively. Finally set

$$
\text{(3.11)}\qquad \beta^{(m+1)k} = \tilde{x}_{\kappa_2}\ .
$$

Clearly since the two iterations are interrelated, the order of computation must be $\alpha^1, \alpha^2, \alpha^3, \ldots, \alpha^k, \beta^k, \alpha^{k+1}, \ldots, \alpha^{2k}, \beta^{2k}, \alpha^{2k+1}, \ldots$.

We define $\bar{\alpha}^{n+1}$ to be the solution of (3.3) with ξ^n replaced by α^n, i.e., let $\bar{\alpha}^{n+1}$ satisfy

$$(3.12) \qquad L^n(\alpha)(\bar{\alpha}^{n+1} - \alpha^n) = -\Delta t B^n(\alpha,\beta)\alpha^n + \Delta t U^n(\alpha,\beta)\alpha^n, \qquad n \geq 0 .$$

Similarly, we define $\bar{\beta}^{mk}$ to be the solution of

$$(3.13) \qquad A^{mk}(\alpha)\bar{\beta}^{mk} = \Gamma^{mk}(\alpha), \qquad m \geq 1 .$$

It is well known [1,6,8] that there exist constants $0 < \rho_1 < 1$ and $0 < \rho_2 < 1$ such that

$$(3.14) \quad \begin{aligned}
&\text{a)} \quad \|L^0(\alpha)^{\frac{1}{2}}(\bar{\alpha}^1 - \alpha^1)\|_e \leq \rho_1 \|L^0(\alpha)^{\frac{1}{2}}(\bar{\alpha}^1 - \alpha^0)\|_e , \\
&\text{b)} \quad \|L^n(\alpha)^{\frac{1}{2}}(\bar{\alpha}^{n+1} - \alpha^{n+1})\|_e \leq \rho_1 \|L^n(\alpha)^{\frac{1}{2}}(\bar{\alpha}^{n+1} - 2\alpha^n + \alpha^{n-1})\|_e, \qquad n \geq 1 , \\
&\text{c)} \quad \|A^k(\alpha)^{\frac{1}{2}}(\bar{\beta}^k - \beta^k)\|_e \leq \rho_2 \|A^k(\alpha)^{\frac{1}{2}}(\bar{\beta}^k - \beta^0)\|_e , \qquad m = 1 , \\
&\text{d)} \quad \|A^{mk}(\alpha)^{\frac{1}{2}}(\bar{\beta}^{mk} - \beta^{mk})\|_e \leq \rho_2 \|A^{mk}(\alpha)^{\frac{1}{2}}(\bar{\beta}^{mk} - 2\beta^{(m-1)k} + \beta^{(m-2)k})\|_e, \qquad m \geq 2 ,
\end{aligned}$$

where the subscript e indicates the Euclidean norm of the vector. Given a_0 and b_0 there exist constants ψ_0, ψ_1, $\tilde{\psi}_0$, and $\tilde{\psi}_1$ such that

$$(3.15) \quad \begin{aligned}
&\text{a)} \quad 0 < \psi_0 \leq \frac{x^T L^n(\alpha) x}{x^T L_0 x} \leq \psi_1, \qquad 0 \neq x \in \mathbb{R}^{M_1} , \\
&\text{b)} \quad 0 < \tilde{\psi}_0 \leq \frac{y^T A^{mk}(\alpha) y}{y^T A_0 y} \leq \tilde{\psi}_1 , \qquad 0 \neq y \in \mathbb{R}^{M_2} ,
\end{aligned}$$

where the constants are independent of h and depend only on the bounds for the coefficients in (2.5). Letting

$$\text{(3.16)}\qquad \text{a)}\quad Q_1 = \frac{1-(\psi_0/\psi_1)^{\frac{1}{2}}}{1+(\psi_0/\psi_1)^{\frac{1}{2}}}\,,$$

$$\text{b)}\quad Q_2 = \frac{1-(\tilde\psi_0/\tilde\psi_1)^{\frac{1}{2}}}{1+(\tilde\psi_0/\tilde\psi_1)^{\frac{1}{2}}}\,,$$

we know from [1,5,6,8] that $\rho_i \le 2Q_i^{\kappa_i}$, $i = 1,2$. If $\delta > 0$ and

$$\text{(3.17)}\qquad \kappa_i \ge \delta \log \frac{1}{\Delta t} / \log \frac{1}{Q_i}\,, \qquad i = 1,2\,,$$

then

$$\text{(3.18)}\qquad \rho_i < 2(\Delta t)^{\delta}\,, \qquad i = 1,2\,.$$

Note that

$$\text{a)}\quad \bar C^{n+1} = \sum_{i=1}^{M_1} \bar\alpha_i^{n+1}\mu_i$$

(3.19) and

$$\text{b)}\quad \bar P^{mk} = \sum_{i=1}^{M_2} \bar\beta_i^{mk}\psi_i$$

satisfy, respectively, for Case I,

$$\text{a)}\quad \left(\varphi\, \frac{\bar C^{n+1} - C^n}{\Delta t},\chi\right) + \left(b(t^{n+1})\nabla \bar C^{n+1},\nabla\chi\right) = -\left(u(C^{*n},\nabla E_k^{\nu}P^n)\cdot\nabla C^n,\chi\right) \qquad \chi \in M_h\,,$$

(3.20) and, for $(n+1)/k \in \mathbb{Z}$,

$$\text{b)}\quad \left(a(C^{*n+1})\nabla \bar P^{n+1},\nabla y\right) = \left(a(C^{*n+1})\gamma(C^{*n+1})\nabla g,\nabla y\right), \qquad y \in N_h\,.$$

We note that for $n = 0,1,\ldots,N$ we have

$$\text{(3.21)}\qquad C^{*n}(x) \equiv \begin{cases} 0\,, & \text{if } C(x) \le 0\,, \\ C^n(x)\,, & \text{if } 0 < C^n(x) < 1\,, \\ 1\,, & \text{if } C^n(x) \ge 1\,, \end{cases}$$

so that, since $0 \leq c \leq 1$,

$$\|c^n - C^{*n}\| \leq \|c^n - C^n\| \quad . \tag{3.22}$$

At this point we shall define some special norms and semi-norms for our analysis. Let

$$\begin{aligned}
&\text{a)} \quad \|\chi\|_{\varphi}^2 \equiv (\varphi\chi,\chi) \ , \\
&\text{b)} \quad \|\chi\|_{b^n}^2 \equiv \begin{cases} (b(t^{n+1})\nabla\chi,\nabla\chi), & \text{for case I} \ , \\ (b(C^n, E_k^{\nu}\nabla P^n)\nabla\chi,\nabla\chi), & \text{for case II} \ , \end{cases} \\
&\text{c)} \quad \|\chi\|_{a^n}^2 \equiv (a(C^n)\nabla\chi,\nabla\chi) \ .
\end{aligned} \tag{3.23}$$

By (2.5), $\|\cdot\|_{\varphi}$ is equivalent to $\|\cdot\|$ and $\|\cdot\|_{a^n}$ and $\|\cdot\|_{b^n}$ are uniformly equivalent to $\|\nabla.\|$ for all n.

We note that in terms of this norm-notation, (3.14) and the triangle inequality yield

$$\begin{aligned}
&\text{a)} \quad \|\bar{C}^1 - C^1\|_{\varphi} + (\Delta t)^{\frac{1}{2}}\|\bar{C}^1 - C^1\|_{b^0} \\
&\qquad \leq \rho_1' \{\|\delta C^0\|_{\varphi} + (\Delta t)^{\frac{1}{2}}\|\delta C^0\|_{b^0}\} \ , \\
&\text{b)} \quad \|\bar{C}^{n+1} - C^{n+1}\|_{\varphi} + (\Delta t)^{\frac{1}{2}}\|\bar{C}^{n+1} - C^{n+1}\|_{b^n} \\
&\qquad \leq \rho_1' \{\|\delta^2 C^n\|_{\varphi} + (\Delta t)^{\frac{1}{2}}\|\delta^2 C^n\|_{b^n}\} \ , \qquad n \geq 1 \ , \\
&\text{c)} \quad \|\bar{P}^k - P^k\|_{a^k} \leq \rho_2\|\bar{P}^k - P^0\|_{a^k} \ , \qquad m = 1 \ , \\
&\text{d)} \quad \|\bar{P}^{mk} - P^{mk}\|_{a^{mk}} \leq \rho_2\|\bar{P}^{mk} - 2P^{(m-1)k} + P^{(m-2)k}\|_{a^{mk}} \ , \qquad m \geq 2 \ ,
\end{aligned} \tag{3.24}$$

where we define

$$\begin{aligned}
&\text{a)} \quad \delta C^n = C^{n+1} - C^n \ , \\
&\text{b)} \quad \delta^2 C^n = C^{n+1} - 2C^n + C^{n-1} \\
&\text{c)} \quad \rho_1' = \frac{\rho_1}{1 - \rho_1} \ .
\end{aligned} \tag{3.25}$$

The convergence results of Section 4 depend only upon the norm reductions (3.24) and not upon the particular iterative method used to achieve these norm reductions.

We shall now define a starting procedure to obtain C^0 and P^0. First we must compute a C^0 which satisfies

$$\|C^0 - \tilde{C}^0\|_1 \leq K_9\{h^r + (\Delta t)^2\} . \tag{3.26}$$

One way to obtain such an estimate is to factor one additional matrix and solve (2.13) directly. Once C^0 is obtained, we can use the factored A_0 from (3.2.g), a starting guess of $\tilde{x}_0 = 0$ (or anything closer) in (3.10.a) and iterate the conjugate gradient procedure κ times where κ satisfies (3.17) with $\delta = 1$ to obtain

$$\|P^0 - \bar{P}^0\|_{a^0} \leq K_{10}(\Delta t) . \tag{3.27}$$

<u>Remark</u>: Although (3.27) will retain the proper convergence rates through all the analysis to follow, for better practical results, one should iterate κ times where κ satisfies (3.17) with $\delta = 2$ to obtain

$$\|P^0 - \bar{P}^0\|_{a^0} \leq K(\Delta t)^2 . \tag{3.28}$$

4. A Priori Error Estimates

In this section we obtain a priori bounds for the errors $C^n - c^n$ and $P^n - p^n$ for the procedures defined in Section 3. The first result, Theorem 4.1, states that given a starting procedure for obtaining C^0, P^0, C^1, and P^k, we obtain optimal order H^1 and L^2 error estimates if the iterative process reduces the error in the solution of the algebraic problems by a fixed factor (independent of n, h, and Δt). Thus we need only a fixed number of iterations of both the concentration and pressure equations at the appropriate time steps. We emphasize that the elliptic problem must be approximately solved only at every k^{th} time step where k is determined by (2.29) and the relative smoothnesses of c and p. Corollary 4.2 extends the optimal order error results of Theorem 4.1 to the physical case where $b = b(x)$ depends solely on x by replacing the iteration in the equations for the concentration by a simple backsolve at each time step. Theorem 4.3, the third result, states that, in Case II where $b = b(x,c,\nabla p)$, by iterating

$$\kappa_1 = \log(\Delta t)^{-1}/\log(\rho_1)^{-1}$$

times on the equations for concentration at each time step and a fixed number of times for the equations for pressure, we can still obtain quasi-optimal L^2 error estimates. If a space of piecewise polynomials of one higher degree is used for the pressure than for the concentration, optimal order L^2 estimates are obtained. Corollary 4.4 states that by updating the preconditioning matrix L_0 from (3.5) every $(\Delta t)^{-1/2}$ time steps, two iterations of the equations for concentration will suffice to obtain our results. Finally extensions to methods with time truncation errors of order $O((\Delta t)^2)$ are discussed briefly.

Theorem 4.1: Let (c,p) satisfy (1.1)-(1.5) and (C^n,P^n) satisfy (3.20), (3.26), and (3.27) and let k be fixed as in (2.29). If we obtain norm reductions in (3.24) of the form

(4.0)

a) $\rho_1' \le (\Delta t)^{\frac{1}{2}}$ for C^1,

b) $\rho_1' \le \min\left\{\frac{1}{8}, \frac{1}{16}\left(\frac{b_*}{b^*}\right)\right\}$ for C^n, $n \ge 2$,

c) $\rho_2 \le (\Delta t)^{\frac{1}{2}}$ for P^k,

d) $\rho_2 \le \frac{1}{12}\left(\frac{a_*}{a^*}\right)^{\frac{1}{2}}$ for P^{mk}, $m \ge 2$,

then there exist positive constants $K_{11} = K_{11}(\lambda, a_*, b_*, \varphi_*, K_i, i \le 10)$ h_0 and τ_0, such that, if $\Delta t \le \tau_0$ and $h \le h_0$,

$$\sup_{t^n}\{\|C - c\|^2 + h\|C - c\|_1^2\} + \sum_{n=0}^{N-1} \|d_t(C - c)^n\|^2 \Delta t \tag{4.1}$$

$$\le K_{11}\{h^{2r} + h^{2s} + h^{2r+2s-6} + (\Delta t)^2 + D^{*2}(\Delta t_1)^4\}$$

where D^* is given in (4.36).

Proof: Let $\zeta^n = C^n - \tilde{c}^n$ and $\eta^n = P^n - \tilde{p}^n$. We shall first obtain an estimate on $\nabla\eta$ at the Δt_1-time levels in terms of ζ and the error induced at each time level from the iterative approximation to (3.4). Subtract (2.10) from (3.20.b) to obtain at each Δt_1-time level,

$$(a(C^{*mk})\nabla\eta^{mk}, \nabla y) = ([a(c^{mk}) - a(C^{*mk})]\nabla\tilde{p}^{mk}, \nabla y)$$

$$+ ([a(C^{*mk})\gamma(C^{*mk}) - a(c^{mk})\gamma(c^{mk})]\nabla g, \nabla y) \tag{4.2}$$

$$+ (a(C^{*mk})\nabla(P^{mk} - \bar{P}^{mk}), \nabla y), \qquad y \in N_h.$$

Let $y = \eta^{mk} \in N_h$. Since C^{*mk} satisfies (3.21) and thus (2.4), we can use (2.5), (2.14), (2.15), (3.22), and (3.23) to obtain

$$\|\eta^{mk}\|^2_{a^{mk}} \le M\|\nabla\tilde{p}^{mk}\|_{L^\infty}\{\|c^{mk} - \tilde{c}^{mk}\| + \|\zeta^{mk}\|\}\|\nabla\eta^{mk}\|$$

$$+ 2MK_1\|\nabla g\|_{L^\infty}\{\|c^{mk} - \tilde{c}^{mk}\| + \|\zeta^{mk}\|\}\|\nabla\eta^{mk}\| + \|\bar{P}^{mk} - P^{mk}\|_{a^{mk}}\|\eta^{mk}\|_{a^{mk}} \tag{4.3}$$

$$\le \frac{3}{4}\|\eta^{mk}\|^2_{a^{mk}} + K_{13}\{\|\zeta^{mk}\|^2 + h^{2r}\} + \frac{1}{2}\|\bar{P}^{mk} - P^{mk}\|^2_{a^{mk}}.$$

We then see that

$$\|\eta^{mk}\|_{a^{mk}} \leq K_{14}\{\|\zeta^{mk}\| + h^r\} + 2\|\bar{P}^{mk} - P^{mk}\|_{a^{mk}} . \tag{4.4}$$

We next consider the concentration equations. Recall that $d_t\zeta^n \equiv \dfrac{\zeta^{n+1} - \zeta^n}{\Delta t}$. Next subtract (2.13) from (3.20.a) to obtain for $n = 0,1,\ldots,\ell - 1$,

$$\begin{aligned}(\varphi d_t\zeta^n,\chi) + (b^{n+1}\nabla\zeta^{n+1},\nabla\chi) &= \left(\varphi\left[\frac{\partial c^{n+1}}{\partial t} - d_t\tilde{c}^n\right],\chi\right) + \lambda((\tilde{c} - c)^{n+1},\chi) \\ &+ ([u(c^{n+1},\nabla p^{n+1})\cdot\nabla\tilde{c}^{n+1} - u(C^{*n},E_k^{\nu}\nabla P^n)\cdot\nabla C^n],\chi) \\ &+ \left(\varphi\,\frac{C^{n+1} - \bar{C}^{n+1}}{\Delta t},\chi\right) + (b(t^{n+1})\nabla(C^{n+1} - \bar{C}^{n+1}),\nabla\chi), \qquad \chi \in M_h .\end{aligned} \tag{4.5}$$

As in [6,9], in order to most efficiently make use of the starting procedures (3.7.a) of the conjugate gradient iteration, we need to use the test function $\chi = \zeta^{n+1} - \zeta^n = \Delta t d_t\zeta^n$ in (4.5). Then we see that with a slight rearrangement, the left side of (4.5) becomes

$$\begin{aligned}&\Delta t(\varphi d_t\zeta^n,d_t\zeta^n) + \frac{\Delta t}{2}(b^{n+1}\nabla(\zeta^{n+1} - \zeta^n),\nabla d_t\zeta^n) + \frac{1}{2}(b^{n+1}\nabla(\zeta^{n+1} + \zeta^n),\nabla(\zeta^{n+1} - \zeta^n)) \\ &= \Delta t\|d_t\zeta^n\|_{\varphi}^2 + \frac{(\Delta t)^2}{2}\|d_t\zeta^n\|_{b^n}^2 + \frac{1}{2}\{\|\zeta^{n+1}\|_{b^n}^2 - \|\zeta^n\|_{b^n}^2\} ,\end{aligned} \tag{4.6}$$

where we have used the norms and semi-norms defined in (3.23). Using techniques like those used in obtaining (4.4), we sum the first two terms on the right of (4.5) from $n = 0$ to $n = \ell - 1$ and use (2.14) to see that

$$\begin{aligned}&\left|\sum_{n=0}^{\ell-1}\left(\left\{\varphi\left[\frac{\partial c^n}{\partial t} - d_t\tilde{c}^n\right] + \lambda(\tilde{c} - c)^{n+1}\right\},d_t\zeta^n\right)\Delta t\right| \\ &\leq \frac{1}{16}\sum_{n=0}^{\ell-1}\|d_t\zeta^n\|_{\varphi}^2\Delta t + K_{15}\left\{h^{2r} + \left\|\frac{\partial^2 c}{\partial t^2}\right\|_{L^2(J;L^2)}^2(\Delta t)^2\right\} .\end{aligned} \tag{4.7}$$

Next, the third term on the right of (4.5) can be summed and split as follows:

$$\begin{aligned}&\left|\sum_{n=0}^{\ell-1}([u(c^{n+1},\nabla p^{n+1})\cdot\nabla\tilde{c}^{n+1} - u(C^{*n},E_k^{\nu}\nabla P^n)\cdot\nabla C^n],d_t\zeta^n)\Delta t\right| \\ &\leq \left|\sum_{n=0}^{\ell-1}([u(c^{n+1},\nabla p^{n+1}) - u(C^{*n},E_k^{\nu}\nabla P^n)]\cdot\nabla\tilde{c}^{n+1},d_t\zeta^n)\Delta t\right| \\ &+ \left|\sum_{n=1}^{\ell-1}(u(C^{*n},E_k^{\nu}\nabla P^n)\cdot\nabla[\tilde{c}^{n+1} - C^n],d_t\zeta^n)\Delta t\right| \equiv T_1 + T_2 .\end{aligned} \tag{4.8}$$

We note that from (2.5.f) and (2.16), we can bound the second term on the right of (4.8) by

$$
\begin{aligned}
(4.9)\qquad T_2 \le \sum_{n=0}^{\ell-1} K_1^2(1 + \|E_k^\nu \nabla P^n\|_{L^\infty})^2 K_{16}\{\|\zeta^n\|^2 + (\Delta t)^2\|\frac{\partial \tilde{c}}{\partial t}\|^2_{L^\infty(J;H^1)}\}\,\Delta t \\
+ \frac{1}{16}\sum_{n=1}^{\ell-1} \|d_t\zeta^n\|_\varphi^2 \Delta t .
\end{aligned}
$$

We shall now state an induction hypothesis. Assume that for $m = 0,1,\ldots,[\![\ell - 1/k]\!]$,

$$
\text{(I1)}\qquad \|\nabla P^{mk}\|_{L^\infty} \le 2K_5
$$

where K_5 is given by (2.15). Clearly from (2.15) and (3.27) we know that (I1) is satisfied for $m = 0$. Then from (2.28), we see that under (I1),

$$
(4.10)\qquad \|E_k^\nu \nabla P^{mk}\|_{L^\infty} \le 4K_5, \qquad m = 1,2,\ldots,[\![\ell-1/k]\!]
$$

for use in (4.9). We next bound the first term on the right of (4.8) as follows:

$$
\begin{aligned}
T_1 \le \Big|\sum_{n=0}^{\ell-1} ([u(c^{n+1},\nabla p^{n+1}) - u(c^{n+1},\nabla\tilde{p}^{n+1})]\cdot\nabla\tilde{c}^{n+1}, d_t\zeta^n)\Big|\Delta t \\
(4.11)\qquad + \Big|\sum_{n=0}^{\ell-1} ([u(c^{n+1},\nabla\tilde{p}^{n+1}) - u(c^{n+1},E_k^\nu\nabla\tilde{p}^n) + u(c^{n+1},E_k^\nu\nabla\tilde{p}^n) \\
- u(c^{n+1},E_k^\nu\nabla P^n) + u(c^{n+1},E_k^\nu\nabla P^n) - u(C^{*n},E_k^\nu\nabla P^n)]\cdot\nabla\tilde{c}^{n+1}, d_t\zeta^n)\Big|\Delta t \equiv T_3 + T_4 .
\end{aligned}
$$

We then use (2.14) and (2.16) to bound the second term on the right of (4.11). Note that

$$
\begin{aligned}
T_4 \le K_{17}\Big\{(\Delta t_1)^4\Big\|\frac{\partial^2\tilde{p}}{\partial t^2}\Big\|^2_{L^2(J;H^1)} + (\Delta t)^2\|\frac{\partial c}{\partial t}\|^2_{L^\infty(J;L^2)} + h^{2r} + \sum_{n=0}^{\ell-1}\|\zeta^n\|^2\Delta t\Big\} \\
(4.12)\qquad + \frac{1}{16}\sum_{n=0}^{\ell-1}\|d_t\zeta^n\|_\varphi^2\Delta t + K_{18}\sum_{m=0}^{[\![\ell-1/k]\!]}\|\nabla\eta^{mk}\|^2\Delta t .
\end{aligned}
$$

We next use summation by parts to treat the first term on the right of (4.11). Note that

$$T_3 \le \Big|\sum_{n=1}^{\ell-1} \big([\{u(c^{n+1},\nabla p^{n+1}) - u(c^{n+1},\nabla\tilde{p}^{n+1})\}\cdot\nabla\tilde{c}^{n+1}$$

$$- \{u(c^{n},\nabla p^{n}) - u(c^{n},\nabla\tilde{p}^{n})\}\cdot\nabla\tilde{c}^{n}], \zeta^n\big)\Big|$$

$$+ \big|([u(c^1,\nabla p^1) - u(c^1,\nabla\tilde{p}^1)]\cdot\nabla\tilde{c}^1,\zeta^0)\big| + \big|([u(c^\ell,\nabla p^\ell) - u(c^\ell,\nabla\tilde{p}^\ell)]\cdot\nabla\tilde{c}^\ell,\zeta^\ell)\big|$$

$$\le \Big|\sum_{n=1}^{\ell-1} ([u(c^{n+1},\nabla p^{n+1}) - u(c^{n+1},\nabla\tilde{p}^{n+1})]\cdot\nabla(\tilde{c}^{n+1} - \tilde{c}^{n}),\zeta^n)\Big| \tag{4.13}$$

$$+ \Big|\sum_{n=1}^{\ell-1} ([\{u(c^{n+1},\nabla p^{n+1}) - u(c^{n+1},\nabla\tilde{p}^{n+1})\} - \{u(c^{n},\nabla p^{n}) - u(c^{n},\nabla\tilde{p}^{n})\}]\cdot\nabla\tilde{c}^{n},\zeta^n)\Big|$$

$$+ \big|([u(c^1,\nabla p^1) - u(c^1,\nabla\tilde{p}^1)]\cdot\nabla\tilde{c}^1,\zeta^0)\big|$$

$$+ \big|([u(c^\ell,\nabla p^\ell) - u(c^\ell,\nabla\tilde{p}^\ell)]\cdot\nabla\tilde{c}^\ell,\zeta^\ell)\big| \equiv T_5 + T_6 + T_7 + T_8\,.$$

We note that since $\dfrac{\partial u_i}{\partial q_i} = a(c)$, for $i = 1,2$, we have

$$T_5 + T_7 + T_8 = \Big|\sum_{n=1}^{\ell-1} (a(c^{n+1})\nabla(p - \tilde{p})^{n+1}\cdot\nabla d_t\tilde{c}^n,\zeta^n)\Delta t\Big|$$

$$+ \big|(a(c^1)\nabla(p - \tilde{p})^1\cdot\nabla\tilde{c}^1,\zeta^0)\big| + \big|(a(c^\ell)\nabla(p - \tilde{p})^\ell\cdot\nabla\tilde{c}^\ell,\zeta^\ell)\big|$$

$$\le \Big|\sum_{n=1}^{\ell-1} (a(c^{n+1})\nabla(p - \tilde{p})^{n+1}\cdot\nabla d_t(\tilde{c} - c)^n,\zeta^n)\Delta t\Big|$$

$$+ \Big|\sum_{n=1}^{\ell-1} (a(c^{n+1})\nabla(p - \tilde{p})^{n+1}\cdot\nabla d_t c^n,\zeta^n)\Delta t\Big| \tag{4.14}$$

$$+ \big|(a(c^1,\nabla(p - \tilde{p})^1\cdot\nabla(c - \tilde{c})^1,\zeta^0)\big| + \big|(a(c^1)\nabla(p - \tilde{p})^1\cdot\nabla c^1,\zeta^0)\big|$$

$$+ \big|(a(c^\ell)\nabla(p - \tilde{p})^\ell\cdot\nabla(c - \tilde{c})^\ell,\zeta^\ell)\big| + \big|(a(c^\ell)\nabla(p - \tilde{p})^\ell\cdot\nabla c^\ell,\zeta^\ell)\big|\,.$$

We next use (2.3), (2.12), and (2.14) to bound the first term on the right of (4.14),

$$\Big|\sum_{n=1}^{\ell-1} (a(c^{n+1})\nabla(p - \tilde{p})^{n+1}\cdot\nabla d_t(\tilde{c} - c)^n,\zeta^n)\Delta t\Big|$$

$$\le K_{19}\sum_{n=1}^{\ell-1} \|\nabla(p - \tilde{p})^{n+1}\|\,\|\nabla d_t(c - \tilde{c})^n\|\,\|\zeta^n\|_{L^\infty}\Delta t$$

$$\le K_{20}\sum_{n=1}^{\ell-1} h^{s+r-3}\|\zeta^n\|\Delta t \tag{4.15}$$

$$\le K_{21}\sum_{n=1}^{\ell-1} \|\zeta^n\|^2\Delta t + K_{22}h^{2r+2s-6}\,.$$

The third and fifth terms on the right of (4.14) can be bounded in a similar fashion to yield

$$(4.16)\qquad \begin{aligned} &|(a(c^1)\nabla(p-\tilde{p})^1\cdot\nabla(c-\tilde{c})^1,\zeta^0)| + |(a(c^\ell)\nabla(p-\tilde{p})^\ell\cdot\nabla(c-\tilde{c})^\ell,\zeta^\ell)| \\ &\le \frac{1}{16}\|\zeta^\ell\|^2_{b^{\ell-1}} + K_{23}\{\|\zeta^0\|_1^2 + h^{2r+2s-6}\}\ . \end{aligned}$$

We integrate by parts and use (2.12) and (2.15) to bound the second term on the right of (4.14). Denote $a(c^n)$ by a^n in what follows.

$$\begin{aligned} &\Big|\sum_{n=1}^{\ell-1}(a^{n+1}\nabla(p-\tilde{p})^{n+1}\cdot\nabla d_t c^n,\zeta^n)\Delta t\Big| \\ &= \Big|-\sum_{n=1}^{\ell-1}((p-\tilde{p})^{n+1},\nabla\cdot\{a^{n+1}\zeta^n\nabla d_t c^n\})\Delta t \\ (4.17)\qquad &\quad + \sum_{n=1}^{\ell-1}\langle(p-\tilde{p})^{n+1},a^{n+1}\zeta^n\nabla d_t c^n\cdot\nu\rangle\Delta t\Big| \\ &\le K_{24}\sum_{n=1}^{\ell-1}\{\|\zeta^n\|_1^2 + h^{2s}\}\sigma_{1,n} + \Big|\sum_{n=1}^{\ell-1}\langle(p-\tilde{p})^{n+1},a^{n+1}\zeta^n\nabla d_t c^n\cdot\nu\rangle\Delta t\Big| \end{aligned}$$

where

$$(4.18)\qquad \sigma_{1,n} = \int_{t^n}^{t^{n+1}} \Big\|\frac{\partial c}{\partial t}(\cdot,s)\Big\|_{W_3^2} ds\ .$$

Note that by (2.9), $\sum_{n=0}^{N}\sigma_{1,n} \le K_5$. In order to bound the second term on the right of (4.17) we shall need to introduce an auxiliary problem to perform a Nitsche lift [14,10]. Let $\psi \in H^2(\Omega)$ be the solution of the elliptic problem

$$(4.19)\qquad \begin{aligned} &\text{a)}\quad -\nabla\cdot[a(x,c^n)\nabla\psi] + \psi = 0, && x\in\Omega,\quad t^n\in[0,T]\ , \\ &\text{b)}\quad a(c^n)\frac{\partial\psi^n}{\partial\nu} = \gamma^n, && x\in\partial\Omega\ , \end{aligned}$$

where

$$(4.20)\qquad \begin{aligned} \gamma^n &= a^n\zeta^{n-1}\nabla d_t c^{n-1}\cdot\nu \\ &\equiv G_1\zeta^{n-1}\in H^{1/2}(\partial\Omega), \qquad n = 1,\dots,N\ . \end{aligned}$$

For a definition of $H^{1/2}(\partial\Omega)$ and its norm $|\cdot|_{1/2}$, see [13]. Note that since $\frac{\partial c}{\partial t} \in L^\infty(J;H^{2+\epsilon}(\Omega))$, we have

$$\|G_1\|_{L^\infty(J;H^{1/2+\epsilon}(\partial\Omega))} \le K_{25} . \tag{4.21}$$

Then under the smoothness assumptions on a and c, we use (S) and Lemma 2.2 of [3] to obtain

$$\begin{aligned} \|\psi\|_2 &\le K_{26}|\gamma|_{1/2} \\ &\le K_{27}\|G_1\|_{L^\infty(J;H^{1/2+\epsilon}(\partial\Omega))}|\zeta|_{1/2} \\ &\le K_{28}\|\zeta\|_1 . \end{aligned} \tag{4.22}$$

From (4.19) we see that for any $V \in H^1(\Omega)$,

$$(a(c^n)\nabla\psi^n,\nabla V) + (\psi^n,V) = \langle\gamma^n,V\rangle, \qquad n = 1,2,\ldots,N . \tag{4.23}$$

We then use $V = (p - \tilde{p})^{n+1}$ in (4.23) and (2.10) to see that from (4.17),

$$\begin{aligned} T_9 &= \Big|\sum_{n=1}^{\ell-1} \langle(p - \tilde{p})^{n+1}, a^{n+1}\zeta^n\nabla d_t\zeta^n\cdot\nu\rangle\Delta t\Big| = \Big|\sum_{n=1}^{\ell-1} \langle(p - \tilde{p})^{n+1},\gamma^{n+1}\rangle\Delta t\Big| \\ &= \Big|\sum_{n=1}^{\ell-1} (a^{n+1}\nabla\psi^{n+1},\nabla(p - \tilde{p})^{n+1})\Delta t\Big| + \Big|\sum_{n=1}^{\ell-1} (\psi^{n+1},(p - \tilde{p})^{n+1})\Delta t\Big| \\ &= \Big|\sum_{n=1}^{\ell-1} (a^{n+1}\nabla(\psi^{n+1} - \psi^*),\nabla(p - \tilde{p})^{n+1})\Delta t\Big| + \Big|\sum_{n=1}^{\ell-1} (\psi^{n+1},(p - \tilde{p})^{n+1})\Delta t\Big| \\ &\le \sum_{n=1}^{\ell-1} K_{31}\{\|\psi^{n+1} - \psi^*\|_1\,\|(p - \tilde{p})^{n+1}\|_1 + \|\psi^{n+1}\|\,\|(p - \tilde{p})^{n+1}\|\}\Delta t . \end{aligned} \tag{4.24}$$

We then use the approximability assumption on N_h like (2.2), (2.12) and (4.22) to obtain

$$T_9 \le \sum_{n=1}^{\ell-1} K_{32} \|\psi^{n+1}\|_2 h^s \|p^{n+1}\|_s \Delta t$$

$$(4.25) \qquad \le \sum_{n=1}^{\ell-1} K_{33} |\gamma^{n+1}|_{1/2} h^s \Delta t$$

$$\le K_{34}\{h^{2s} + \sum_{n=1}^{\ell-1} \|\zeta^n\|_1^2 \Delta t\} .$$

Similar estimates yield the bounds

$$|((p - \tilde{p})^1, a^1 \zeta^0 \nabla c^1 \cdot \nu| + |((p - \tilde{p})^\ell, a^\ell \zeta^\ell \nabla \tilde{c}^\ell \cdot \nu)|$$

$$(4.26) \qquad \le \frac{1}{16} \|\zeta^\ell\|^2_{b^{\ell-1}} + K_{35}\{h^{2s} + \|\zeta^0\|_1^2\} .$$

Next, for the vector-valued function $u = u(x,c,q) : \Omega \times [0,1] \times \mathbb{R} \to \mathbb{R}^2$, we define

$$\text{a)} \quad u'_{1,n}(x) \equiv \sum_{i=1}^{2} \int_0^1 \frac{\partial u_i}{\partial c}(x, \theta c^n + (1-\theta)c^{n-1}, \frac{\partial p^n}{\partial x_i}) d\theta ,$$

$$(4.27) \qquad \text{b)} \quad u'_{2,n}(x) \equiv \sum_{i=1}^{2} \int_0^1 \frac{\partial u_i}{\partial c}(x, \theta c^n + (1-\theta)c^{n-1}, \frac{\partial \tilde{p}^n}{\partial x_i}) d\theta ,$$

$$\text{c)} \quad \frac{\partial u_i}{\partial q_i} = a(c), \qquad i = 1,2 .$$

Using (4.27), we treat the second term in (4.13) as follows:

$$T_6 = |\sum_{n=1}^{\ell-1} ([\{u(c^{n+1},\nabla p^{n+1}) - u(c^n,\nabla p^{n+1}) + u(c^n,\nabla p^{n+1}) - u(c^n,\nabla p^n)\}$$

$$- \{u(c^{n+1},\nabla \tilde{p}^{n+1}) - u(c^n,\nabla \tilde{p}^{n+1}) + u(c^n,\nabla \tilde{p}^{n+1}) - u(c^n,\nabla \tilde{p}^n)\}] \cdot \nabla \tilde{c}^n, \zeta^n)|$$

$$= |\sum_{n=1}^{\ell-1} ([u'_{1,n+1}\{c^{n+1} - c^n\} + a^n\{\nabla p^{n+1} - \nabla p^n\} + u'_{2,n+1}\{c^{n+1} - c^n\}$$

$$+ a^n\{\nabla \tilde{p}^{n+1} - \nabla \tilde{p}^n\}] \cdot \nabla \tilde{c}^n, \zeta^n)|$$

$$(4.28) \qquad = |\sum_{n=1}^{\ell-1} ([\{u'_{1,n+1} - u'_{2,n+1}\}\{c^{n+1} - c^n\} + a^n\{\nabla(p - \tilde{p})^{n+1} - \nabla(p - \tilde{p})^n\}] \cdot \nabla \tilde{c}^n, \zeta^n)|$$

$$\le \left|\sum_{n=1}^{\ell-1} \left(\left[\int_0^1 \frac{\partial a}{\partial c} d\theta \nabla(p - \tilde{p})^{n+1} d_t c^n + a^n \nabla d_t (p - \tilde{p})^n\right] \cdot \nabla \tilde{c}^n, \zeta^n\right) \Delta t\right| .$$

Then using the same techniques as above, if

$$(4.29)\qquad \|c\|_{L^\infty(J;H^r)} + \|c\|_{L^\infty(J;H^3)} + \left\|\frac{\partial c}{\partial t}\right\|_{L^\infty(J;H^2)} + \|p\|_{L^\infty(J;H^s)} + \left\|\frac{\partial p}{\partial t}\right\|_{L^2(H^s)} \le K_{36}$$

we can obtain

$$(4.30)\qquad T_6 \le K_{37}\{h^{2s} + h^{2r+2s-6} + \sum_{n=1}^{\ell-1} \|\zeta^n\|_1^2 \Delta t\} .$$

We next obtain bounds for the last two terms on the right side of (4.5) using (3.24). Since different starting procedures were used in the conjugate gradient iteration to obtain C^1 and C^n for $n \ge 2$, we shall estimate each case separately. From (3.24.a) we see that for $n = 0$,

$$(4.31)\qquad \begin{aligned} &\left|\left(\varphi \frac{C^1 - \bar{C}^1}{\Delta t}, \Delta t d_t\zeta^0\right) + (b(t^1)\nabla(C^1 - \bar{C}^1), \nabla \Delta t d_t\zeta^0)\right| \\ &\le \|C^1 - \bar{C}^1\|_\varphi \|d_t\zeta^0\|_\varphi + \|C^1 - \bar{C}^1\|_{b^0}\|d_t\zeta^0\|_{b^0}\Delta t \\ &\le \rho_1'\{\|\delta C^0\|_\varphi + (\Delta t)^{\frac{1}{2}}\|\delta C^0\|_{b^0}\}\{\|d_t\zeta^0\|_\varphi + (\Delta t)^{\frac{1}{2}}\|d_t\zeta^0\|_{b^0}\} \\ &\le \rho_1'\Delta t\{\|d_t\zeta^0\|_\varphi + (\Delta t)^{\frac{1}{2}}\|d_t\zeta^0\|_{b^0} \\ &\quad + \left\|\frac{\partial \tilde{c}}{\partial t}\right\|_{L^\infty(J;H^1)}\}\{\|d_t\zeta^0\|_\varphi + (\Delta t)^{\frac{1}{2}}\|d_t\zeta^0\|_{b^0}\} . \end{aligned}$$

We then note that in order to obtain a $(\Delta t)^2$ term, we need, for $n = 0$,

$$(4.32)\qquad \rho_1' \le K_{38}(\Delta t)^{\frac{1}{2}}$$

to obtain

$$(4.33)\qquad \begin{aligned} &|(\varphi(C^1 - \bar{C}^1), d_t\zeta^0) + \Delta t(b(t^1)\nabla(C^1 - \bar{C}^1), \nabla d_t\zeta^0)| \\ &\le \frac{\Delta t}{16}\|d_t\zeta^0\|_\varphi^2 + \frac{(\Delta t)^2}{16}\|d_t\zeta^0\|_{b^0}^2 + K_{39}\left\|\frac{\partial \tilde{c}}{\partial t}\right\|_{L^\infty(J;H^1)}^2 (\Delta t)^2 . \end{aligned}$$

Similarly, if for $n \ge 1$,

$$\rho_1' \le \min\left\{\frac{1}{8}, \frac{1}{16}\left(\frac{b_*}{b^*}\right)\right\},$$

we obtain

$$|(\varphi(C^{n+1} - \bar{C}^{n+1}), d_t\zeta^n) + (b(t^{n+1})\nabla(C^{n+1} - \bar{C}^{n+1}), \nabla d_t\zeta^n)\Delta t|$$

$$\le \rho_1'\{\|\delta^2 C^n\|_\varphi + (\Delta t)^{\frac{1}{2}}\|\delta^2 C^n\|_{b^n}\}\{\|d_t\zeta^n\|_\varphi + (\Delta t)^{\frac{1}{2}}\|d_t\zeta^n\|_{b^n}\}$$

(4.34)
$$\le \rho_1'\Delta t\Big\{\|d_t\zeta^n\|_\varphi + \|d_t\zeta^{n-1}\|_\varphi + (\Delta t)^{\frac{1}{2}}\Big[\|d_t\zeta^n\|_{b^n} + \left(\frac{b^*}{b_*}\right)^{\frac{1}{2}}\|d_t\zeta^{n-1}\|_{b^{n-1}}\Big]$$

$$+ \Delta t\left\|\frac{\partial^2\tilde{c}}{\partial t^2}\right\|_{L^\infty(J,H^1)}\Big\}\Big\{\|d_t\zeta^n\|_\varphi + (\Delta t)^{\frac{1}{2}}\|d_t\zeta^n\|_{b^n}\Big\}$$

$$\le \frac{1}{8}\{\Delta t\|d_t\zeta^{n-1}\|_\varphi^2 + (\Delta t)^2\|d_t\zeta^{n-1}\|_{b^{n-1}}^2\} + \frac{7}{16}\{\Delta t\|d_t\zeta^n\|_\varphi^2$$

$$+ (\Delta t)^2\|d_t\zeta^n\|_{b^n}^2\} + (\Delta t)^2\left\|\frac{\partial^2\tilde{c}}{\partial t^2}\right\|_{L^\infty(J;H^1)}^2 .$$

Combining the above estimates, we obtain

$$\frac{1}{4}\sum_{n=0}^{\ell-1}\|d_t\zeta^n\|_\varphi^2\Delta t + \frac{1}{2}\sum_{n=0}^{\ell-1}\{\|\zeta^{n+1}\|_{b^n}^2 - \|\zeta^n\|_{b^n}^2\}$$

$$\le \frac{1}{8}\|\zeta^\ell\|_{b^{\ell-1}}^2 + K_{40}\sum_{n=1}^{\ell-1}(\sigma_{1,n} + \Delta t)\|\zeta^n\|_1^2$$

(4.35)
$$+ K_{41}\{\|\zeta^0\|_1^2 + h^{2r} + h^{2s} + h^{2r+2s-6} + (\Delta t)^2 + (\Delta t_1)^4 D^{*2}\}$$

$$+ K_{42}\sum_{m=0}^{[\![\ell-1/k]\!]}\|\nabla(\bar{P}^{mk} - P^{mk})\|^2\Delta t ,$$

where in this case

(4.36)
$$D^{*2} = \|\tilde{p}\|_{W_\infty^2(J;H^1)}/\|\tilde{c}\|_{W_\infty^2(J;H^1)}$$

is to be used in (2.29) to determine k given Δt. We next consider the preconditioned conjugate gradient iteration on the pressure equations to treat the last term on the right of (4.35). From (3.23), (3.24), and (4.4), with $m \ge 2$, we see that

$$\|\nabla(\bar{P}^{mk} - P^{mk})\| \le a_*^{-\frac{1}{2}} \|\bar{P}^{mk} - P^{mk}\|_{a^{mk}} \le a_*^{-\frac{1}{2}} \rho_2 \|\bar{P}^{mk} - 2P^{(m-1)k} + P^{(m-2)k}\|_{a^{mk}}$$

$$\le (a^*/a_*)^{\frac{1}{2}} \rho_2 \Big\{ \|\eta^{mk}\|_{a^{mk}} + 2\|\eta^{(m-1)k}\|_{a^{(m-1)k}} + \|\eta^{(m-2)k}\|_{a^{(m-2)k}}$$

$$+ (\Delta t_1)^2 \left\|\frac{\partial^2 \bar{p}}{\partial t^2}\right\|_{L^\infty(J;H^1)} \Big\} + a_*^{-\frac{1}{2}} \rho_2 \|\bar{P}^{mk} - P^{mk}\|_{a^{mk}}$$

$$(4.37) \qquad \le K_{43}\{\|\zeta^{mk}\| + \|\zeta^{(m-1)k}\| + \|\zeta^{(m-2)k}\| + (\Delta t_1)^2 + h^r\}$$

$$+ \rho_2 \Bigg[\left\{ a_*^{-\frac{1}{2}} + 2\left(\frac{a^*}{a_*}\right)^{\frac{1}{2}} \right\} \|\bar{P}^{mk} - P^{mk}\|_{a^{mk}}$$

$$+ 4\left(\frac{a^*}{a_*}\right)^{\frac{1}{2}} \|\bar{P}^{(m-1)k} - P^{(m-1)k}\|_{a^{(m-1)k}}$$

$$+ 2\left(\frac{a^*}{a_*}\right)^{\frac{1}{2}} \|\bar{P}^{(m-2)k} - P^{(m-2)k}\|_{a^{(m-2)k}} \Bigg] .$$

We then see that if

$$(4.38) \qquad \rho_2 \le \frac{1}{12} (a_*/a^*)^{\frac{1}{2}}, \qquad m \ge 2 ,$$

then

$$(4.39) \qquad \sum_{m=2}^{[\![\ell-1/k]\!]} \Delta t \|\nabla \eta^{mk}\|^2 \le K_{44} \left\{ \sum_{n=0}^{\ell-1} \Delta t \|\zeta^n\|^2 + h^{2r} + (\Delta t_1)^4 D^{*2} \right\}.$$

A similar result will hold for $m = 1$ if we iterate sufficiently many times that

$$(4.40) \qquad \rho_2 \le (\Delta t)^{\frac{1}{2}}, \qquad m = 1 .$$

In order to apply a discrete version of the Gronwall Lemma in the H^1 norm, we must first shift the coefficients in the third term on the right of (4.6) to obtain a telescoping sum in the $\|\cdot\|_{b^n}$ semi-norm and then introduce an L^2 term into the telescoping sum to obtain H^1 terms. Note that

$$\|\zeta^n\|^2_{b^n} = \|\zeta^n\|^2_{b^{n-1}} + ([b(t^{n+1}) - b(t^n)]\nabla\zeta^n, \nabla\zeta^n)$$

(4.41)

$$\leq \|\zeta^n\|^2_{b^{n-1}} + K_{45}\Delta t\|\zeta^n\|^2_1 .$$

Also note that

$$\|\zeta^{n+1}\|^2 - \|\zeta^n\|^2 = 2\Delta t(d_t\zeta^n, \zeta^n) + (\Delta t)^2\|d_t\zeta^n\|^2$$

(4.42)

$$\leq \frac{1}{16}\|d_t\zeta^n\|^2_\varphi + \Delta t K_{16}\|\zeta^n\|^2 .$$

We now sum (4.41) and (4.42) from $n = 0$ to $n = \ell$, add the results to (4.35), use (3.26), (4.39) and (4.40), and apply a version of the discrete Gronwall Lemma in the $\|\cdot\| + \|\cdot\|_b$ norm (equivalent to the H^1 norm), to obtain

$$\sum_{n=0}^{\ell-1} \{\Delta t\|d_t\zeta^n\|^2_\varphi + (\Delta t)^2\|d_t\zeta^n\|^2_{b^n}\} + \|\zeta^\ell\|^2_1$$

(4.43)

$$\leq K_{47}\{h^{2r} + h^{2s} + h^{2r+2s-6} + (\Delta t)^2 + (\Delta t_1)^4 D^{*2}\} .$$

Since (4.43) holds for any $\ell = 0,\ldots,N$, we can replace $\|\zeta^\ell\|^2_1$ by $\sup_{0\leq n\leq N}\|\zeta^n\|^2_1$.

Then we see from (2.3), (2.15), (4.4), (4.37) and (4.43) that

$$\|\nabla p^{mk}\|_{L^\infty} \leq \|\nabla\eta^{mk}\|_{L^\infty} + \|\nabla\tilde{p}\|_{L^\infty(J;L^\infty)} \leq K_3 + K_0 h^{-1}\|\nabla\eta^{mk}\|$$

(4.44)

$$\leq K_5 + K_0 h^{-1} K\{h^r + (\Delta t_1)^2 D^* + \|\zeta^{mk}\| + \|\zeta^{(m-1)k}\| + \|\zeta^{(m-2)k}\|\}$$

$$\leq K_5 + Kh^{-1}\{h^r + h^s + \Delta t + (\Delta t_1)^2 D^*\} .$$

Then if $r,s \geq 2$, $\Delta t = (\Delta t_1)^2 D^*$, and $\Delta t \leq h^{1+\epsilon}$ for $\epsilon > 0$, we see that for Δt and h sufficiently small our induction hypothesis will be satisfied for all $m \leq [\![N/k]\!]$. Finally (2.14), (4.43), and the triangle inequality yield (4.1), the desired result. //

We note that if we replaced $\nabla E_k^\nu p^n$ in the coefficient u in (3.20.a) by $\nabla p^{[\![n/k]\!]}$ or $\nabla E p^{[\![n/k]\!]}$ (as in (2.27) and the discussion following (2.27)) we would replace the term $D^{*2}(\Delta t_1)^4$ in (4.1) by $\tilde{D}^{*2}(\Delta t_1)^2$ where $\tilde{D}^*$ is slightly different

than D^*. However, the use of the preconditioned conjugate gradient iteration with this modification would still involve bounds on $\left\|\frac{\partial^2\tilde{p}}{\partial t^2}\right\|_{L^\infty(J;H^1)}$ and $\left\|\frac{\partial^2\tilde{c}}{\partial t^2}\right\|_{L^\infty(J;H^1)}$.
Since these terms would probably dominate both D^* and $\tilde{D}^*$, the ability to choose a much larger k and thus solve (3.20.b) much less frequently would motivate the use of the different extrapolations defined in (2.28). However, the repeated evaluation of the coefficients u (and b in Case II) is often computationally expensive. By considering the u appearing in (1.1) and (1.2) as a separate variable to be determined from the elliptic equation (1.2) by a mixed method, we could possibly extrapolate the variable u instead of just ∇p to achieve both the accuracy and the computational efficiency of not having to evaluate a new coefficient at each Δt-time step. This idea is being considered elsewhere.

We note that if $b = b(x)$ only, then the L^n defined in (3.3) will be independent of t and n. In this case $L \equiv \Phi + \Delta tB$ will be factored and the preconditioned conjugate gradient iteration for the parabolic equation will be replaced at each step by a simple back-solve of the factored L. The iteration will still be used for the elliptic equations. We obtain the following simplification of Theorem 4.1.

Corollary 4.2: Assume $b = b(x)$ and k is fixed as in (2.29). If we obtain norm reductions in (3.24) (c) and (d) of the form

a) $\rho_2 \le (\Delta t)^{\frac{1}{2}}$ for P^k, and

b) $\rho_2 \le \frac{1}{12}\left(\frac{a_*}{a^*}\right)^{\frac{1}{2}}$ for P^{m-k}, $m \ge 2$,

then there exist positive constants $K_{48} = K_{48}(\lambda, a_*, b_*, \varphi_*, K_i, i \le 10)$, h_0, and τ_0, such that, if $\Delta t \le \tau_0$ and $h \le h_0$,

$$\sup_{t^n}\{\|C - c\|^2 + h\|C - c\|_1^2\} + \sum_{n=0}^{N-1}\|d_t(C - c)^n\|^2\Delta t \tag{4.45}$$

$$\le K_{48}\{h^{2r} + h^{2s} + h^{2r+2s-6} + (\Delta t)^2 + D^{*2}(\Delta t_1)^4\}$$

where

$$D^* = \|\tilde{p}\|_{W^2_\infty(J;H^1)} / \|\tilde{c}\|_{W^1_\infty(J;H^1)} .$$

We next consider the version of (4.5) for Case II, where $b = b(x,t)$ is replaced by $b = b(x,c,\nabla p)$. In this case (4.5) is replaced by

$$(\varphi d_t \zeta^n, \chi) + (b(C^{*n}, E_k^\nu \nabla P^n)\nabla\zeta^{n+1}, \nabla\chi) = \left(\varphi\left[\frac{\partial c^{n+1}}{\partial t} - d_t \tilde{c}^n\right], \chi\right) + \lambda((\tilde{c}-c)^{n+1}, \chi)$$

(4.46)
$$+ ([b(c^{n+1}, \nabla p^{n+1}) - b(C^{*n}, E_k^\nu \nabla P^n)]\nabla\tilde{c}^{n+1}, \nabla\chi)$$
$$+ ([u(c^{n+1}, \nabla p^{n+1})\cdot\nabla\tilde{c}^{n+1} - u(C^{*n}, E_k^\nu \nabla P^n)\cdot\nabla C^n, \chi)$$
$$+ (\varphi \frac{C^{n+1} - \tilde{C}^{n+1}}{\Delta t}, \chi) + (b(C^{*n}, E_k^\nu \nabla P^n)\nabla(C^{n+1} - \tilde{C}^{n+1}), \nabla\chi), \qquad \chi \in M_h .$$

A test function of the form $\zeta^{n+1} - \zeta^n$ is required to make most efficient use of the iterative procedure defined in Section 2. However, this choice of test function causes serious problems in the treatment of the third term on the right of (4.46). Standard techniques for treating a term of this type like *summation by parts in time* (see [6,9]) will not work due to the inability to treat the resulting terms of the form

$$\sum_{n=0}^{\ell-1} ([b(c^{n+1}, \nabla p^{n+1}) - b(c^{n+1}, \nabla E_k^\nu P^n) - \{b(c^n, \nabla p^n) - b(c^n, \nabla E_k^\nu P^{n-1})\}]\nabla\tilde{c}^n, \nabla d_t \zeta^n)\Delta t .$$

For this reason, in the proof of the following result, we shall use a test function of the form $\chi = \zeta^{n+1}$.

<u>Theorem 4.3</u>: Let $b = b(x,c,\nabla p)$ and (C^n, P^n) satisfy (3.20) (with b^{n+1} replaced by $b(C^{*n}, E_k^\nu \nabla P^n)$), (3.26) (with $\|\cdot\|_1$ replaced by $\|\cdot\|$), and (3.27) and let k be fixed as in (2.29). If we obtain norm reductions in (3.24) of the form

a) $\rho_1' \le \frac{1}{4}\Delta t$ for $n \ge 1$,

b) $\rho_2 \le (\Delta t)^{\frac{1}{2}}$ for P^k ,

c) $\rho_2 \le \frac{1}{12}\left(\frac{a_*}{a^*}\right)^{\frac{1}{2}}$ for P^{mk}, $\qquad m \ge 2$, then there exist positive constants $K_{49} = K_{49}(\lambda, a_*, b_*, \varphi_*, K_i, i \le 10)$ such that, if $\Delta t \le \tau_0$ and $h \le h_0$,

$$\sup_{t^n} \{\|C - c\|^2 + h\|C - c\|_1^2\} \le K_{49}\{h^{2r} + h^{2s-2} + (\Delta t)^2 + (\Delta t_1)^4 D^{*2}\} , \tag{4.47}$$

where D^* is given by (4.36).

Proof: We shall use the same notation as in the proof of Theorem 4.1. We obtain (4.4) just as before and then consider (4.46). With $\chi = \zeta^{n+1}$, the left hand side of (4.46) becomes

$$\begin{aligned} &\frac{1}{2\Delta t}\{\|\zeta^{n+1}\|_\varphi^2 - \|\zeta^n\|_\varphi^2\} + \|\zeta^{n+1}\|_{b^n}^2 \\ &= (\varphi d_t \zeta^n, \zeta^{n+1}) + (b(C^{*n}, E_k^\nu \nabla P^n)\nabla\zeta^{n+1}, \nabla\zeta^{n+1}) . \end{aligned} \tag{4.48}$$

We bound the first and second terms on the right of (4.46) as follows

$$\begin{aligned} &\left|\left(\varphi\left[\frac{\partial c^{n+1}}{\partial t} - d_t\tilde{c}^n\right] + \lambda(\tilde{c} - c)^{n+1}, \zeta^{n+1}\right)\right| \\ &\le K_{50}\left\{h^{2r} + \left\|\frac{\partial^2 c}{\partial t^2}\right\|^2_{L^\infty(J;L^2)}(\Delta t)^2\right\} + \frac{1}{32}\|\zeta^{n+1}\|_\varphi^2 . \end{aligned} \tag{4.49}$$

We then bound the third and fourth terms on the right of (4.46) by

$$\begin{aligned} &|([b(c^{n+1}, \nabla p^{n+1}) - b(C^{*n}, E_k^\nu \nabla P^n)]\nabla\tilde{c}^{n+1}, \nabla\zeta^{n+1}) + ([u(c^{n+1}, \nabla p^{n+1})\cdot\nabla\tilde{c}^{n+1} \\ &\quad - u(C^{*n}, E_k^\nu \nabla P^n)\cdot\nabla C^n], \zeta^{n+1})| \\ &\le \frac{1}{32}\{\|\zeta^{n+1}\|^2 + \|\zeta^{n+1}\|_{b^n}^2\} + K_{51}\left\{(\Delta t)^2\left[\left\|\frac{\partial c}{\partial t}\right\|^2_{L^\infty(J;H^0)} + \left\|\frac{\partial\tilde{c}}{\partial t}\right\|^2_{L^\infty(J;H^1)}\right]\right. \\ &\quad \left. + (\Delta t_1)^4\left\|\frac{\partial^2 p}{\partial t^2}\right\|^2_{L^\infty(J;H^1)} + h^{2r} + h^{2s-2} + \|\zeta^n\|^2 + \|E_k^\nu \nabla \eta^n\|^2\right\} . \end{aligned} \tag{4.50}$$

Note we have used our induction hypothesis (I1) in (4.50) as before. We next use (3.24) to bound the last two terms on the right of (4.46). For $n \ge 0$, we multiply by Δt and obtain

$$\left|\left(\varphi\,\frac{C^{n+1}-\bar{C}^{n+1}}{\Delta t},\zeta^{n+1}\right)+(b(C^{*n},E_k^{\nu}\nabla P^n)\nabla(C^{n+1}-\bar{C}^{n+1}),\nabla\zeta^{n+1})\right|\Delta t$$

$$\leq \|C^{n+1}-\bar{C}^{n+1}\|_{\varphi}\|\zeta^{n+1}\|_{\varphi}+\|C^{n+1}-\bar{C}^{n+1}\|_{b^n}\|\zeta^{n+1}\|_{b^n}\Delta t$$

(4.51)

$$\leq \{\|\zeta^{n+1}\|_{\varphi}+(\Delta t)^{\frac{1}{2}}\|\zeta^{n+1}\|_{b^n}\}\rho_1'\{\|\delta C^n\|_{\varphi}+(\Delta t)^{\frac{1}{2}}\|\delta C^n\|_{b^n}\}$$

$$\leq \{\|\zeta^{n+1}\|_{\varphi}+(\Delta t)^{\frac{1}{2}}\|\zeta^{n+1}\|_{b^n}\}\rho_1'\Big\{\|\zeta^{n+1}\|_{\varphi}+\|\zeta^n\|_{\varphi}+\Delta t\,\Big\|\frac{\partial\tilde{c}}{\partial t}\Big\|_{L^\infty(J;H^1)}$$

$$+(\Delta t)^{\frac{1}{2}}\Big[\|\zeta^{n+1}\|_{b^n}+\|\zeta^n\|_{b^n}\Big]\Big\}.$$

Then if we iterate sufficiently many times that

(4.52) $$\rho_1'\leq\frac{1}{4}(\Delta t),$$

we see that (4.51) is bounded by

(4.53) $$\frac{11\Delta t}{16}\{\|\zeta^{n+1}\|_{\varphi}^2+\Delta t\|\zeta^{n+1}\|_{b^n}^2\}+\frac{(\Delta t)^2}{4}\|\zeta^n\|_{b^{n-1}}^2\frac{b^*}{b_*}+K_{52}\Delta t\{\|\zeta^n\|_{\varphi}^2+(\Delta t)^2\}.$$

We then sum the above inequalities on n, combine the results, and let D^* be as in (4.36) to obtain

$$\left(\frac{1}{2}-\frac{3}{4}\Delta t\right)\|\zeta^{\ell}\|_{\varphi}^2+\sum_{n=0}^{\ell-1}\Delta t\left[\frac{31}{32}-\left\{\frac{11}{16}+\frac{1}{4}\frac{b^*}{b_*}\right\}\Delta t\right]\|\zeta^{n+1}\|_{b^n}^2$$

(4.54)

$$\leq K_{53}\left[h^{2r}+h^{2s-2}+(\Delta t)^2+D^{*2}(\Delta t_1)^4+\sum_{n=0}^{\ell-1}\Delta t\|\zeta^n\|^2+\sum_{m=0}^{[\ell-1/k]}\Delta t\|\nabla(\bar{P}^{mk}-P^{mk})\|^2\right].$$

The rest of the proof follows as in the proof of Theorem 4.1. //

We note that (see [6]) if we update the preconditioning matrix, each $(\Delta t)^{-1/2}$-time steps, we obtain a norm reduction of $O(\sqrt{\Delta t})$ with one iteration and $O(\Delta t)$ with two. Using this idea we obtain the following corollary to Theorem 4.3.

Corollary 4.4: Let (C^n,P^n) satisfy (3.20), (3.26) and (3.27) as in Theorem 4.3. By updating $L_0=\Phi+\frac{\Delta t}{2}B^n$ at each $(\Delta t)^{-1/2}$-time steps, we obtain hypothesis a) of Theorem 4.3 with only two iterations of (3.7)-(3.9) per time step.

Using techniques like those above we can obtain corresponding results for the Crank-Nicolson-Galerkin approximations to (1.1)-(1.5). These methods do not have the same stability properties as the methods analyzed above unless we make a restriction like

(4.55) $$\Delta t \leq K_{54} h^2 .$$

Under the assumption (4.55) and slightly stronger smoothness assumptions on p and c, we would obtain the same types of results as above with $(\Delta t)^2$ replaced by $(\Delta t)^4$. Without (4.55), we would need to iterate more to obtain the necessary norm reductions from (3.24). See [6] for particulars.

5. Computational Considerations

In this section we shall consider some rough operation counts to estimate the computational complexity of the methods presented here. We shall see that the preconditioned conjugate gradient iterative methods presented allow us to obtain near optimal order work estimates. Thus these methods are very efficient computationally.

Recall that we have two space variables $(d = 2)$. George [11] has shown in some special cases that with $M_1 = M_1(h) = \dim \mathcal{M}_h$, the procedure of setting up and factoring $\Phi + \Delta t B_n$ requires $O(M_1^{3/2})$ operations and that the solution of (3.3), given the factorization, requires $O(M_1 \log M_1)$ operations. Similarly, the work involved in setting up and factoring A_n and solving (3.4) using this factorization are $O(M_2^{3/2})$ and $O(M_2 \log M_2)$ respectively. Hoffman, Martin and Rose [12] have shown that such bounds are minimal. Thus, if we conjecture the validity of the above estimates for our problem and refactor $\Phi + \Delta t B_n$ and A_n and solve (3.3) and (3.4) at each time step, the total amount of work done using the method presented in [10] is

$$(5.1) \qquad O(N_T\{(M_1^{3/2} + M_1 \log M_1) + (M_2^{3/2} + M_2 \log M_2)\}) = O(N_T\{M_1^{3/2} + M_2^{3/2}\}) ,$$

where N_T is the total number of time steps $(N_T \approx \frac{1}{\Delta t})$. We note that the work of refactorization dominates the estimates.

Using the preconditioned conjugate gradient iterative method presented here, one does not need to refactor at every time step. Instead only one factorization of $L_0 = \Phi + \Delta t B_0$ and A_0 need be done. Also, using the different time increments for pressure and concentration, we need only solve (3.4) at every k^{th} time step; thus the total number of times (3.4) need be solved is N_T/k. Thus letting κ_1 and κ_2 be the number of pre-conditioned conjugate gradient iterations needed to achieve the norm reductions given in (4.0) (κ_1 and κ_2 are constants, independent of h, n and Δt), we see that the total work required for Case I is

$$(5.2) \qquad O(M_1^{3/2} + N_T \kappa_1 M_1 \log M_1 + M_2^{3/2} + \frac{N_T}{k} \kappa_2 M_2 \log M_2) .$$

Then since $N_T \approx \frac{1}{\Delta t} \approx \frac{1}{h^r} = O(M_1^{r/2})$, we see that even for $r = 2$ (piecewise linear elements), the second and fourth terms in (5.2) dominate and the work is

$$(5.3) \qquad O(N_T \kappa_1 M_1 \log M_1 + \frac{N_T}{k} \kappa_2 M_2 \log M_2) .$$

We note that the number of unknowns in the problem is

$$(5.4) \qquad O(N_T M_1 + \frac{N_T}{k} M_2) ,$$

so (5.3) represents nearly best possible order work estimates.

It is computationally wasteful to iterate exactly κ_1 times at each time step (respectively κ_2 times at every k_{th} time step) in order to achieve the pessimistic bounds given in (4.0). Instead, one can monitor the norm reduction actually produced at each step of the iteration and stop iterating when sufficient norm reduction is achieved. Additional stopping criteria can be imposed in this monitoring process. See [6] for a discussion stopping criteria for related problems.

REFERENCES

1. O. Axelsson, "On preconditioning and convergence acceleration in sparse matrix problems," CERN European Organization for Nuclear Research, Geneva, 1974.

2. J. E. Dendy, Jr., "An analysis of some Galerkin Schemes for the solution of non-linear time-dependent problems," SIAM J. Numer. Anal. 12 (1975), pp. 541-565.

3. J. Douglas, Jr., and T. Dupont, "Galerkin methods for parabolic equations with nonlinear boundary conditions," Numer. Math. 20 (1973), pp. 213-237.

4. J. Douglas, Jr., and T. Dupont, "The effect of interpolating the coefficients in nonlinear parabolic Galerkin procedures," Math. of Comp. 29 (1975), pp. 360-389.

5. J. Douglas, Jr., and T. Dupont, "Preconditioned conjugate gradient iteration applied to Galerkin methods for a mildly nonlinear Dirichlet problem," Sparse Matrix Computations, Academic Press, Inc., New York, 1976, pp. 333-348.

6. J. Douglas, Jr., T. Dupont, and R. E. Ewing, "Incomplete iteration for time-stepping a Galerkin method for a quasilinear parabolic problem," SIAM J. Numer. Anal. (to appear).

7. T. Dupont, G. Fairweather, and J. P. Johnson, "Three-level Galerkin methods for parabolic equations," SIAM J. Numer. Anal. 11 (1974), pp. 392-410.

8. M. Engeli, Th. Ginsburg, H. Rutishauser, and E. Stiefel, "Refined iterative methods for the computation of the solution and the eigenvalues of self-adjoint boundary value problems," Mitteilungen aus dem Institut für Angewandte Mathematik, nr. 8, ETH, Zurich, 1950.

9. R. E. Ewing, "Time-stepping Galerkin methods for nonlinear Sobolev partial differential equations," SIAM J. Numer. Anal. 15 (1978), pp. 1125-1150.

10. R. E. Ewing and M. F. Wheeler, "Galerkin methods for miscible displacement problems in porous media," (to appear).

11. A. George, "Nested dissection on a regular finite element mesh," SIAM J. Numer. Anal. 10 (1973), pp. 345-363.

12. A. J. Hoffman, M. S. Martin, and D. J. Rose, "Complexity bounds for regular finite difference and finite element grids," SIAM J. Numer. Anal. 10 (1973), pp. 364-369.

13. J. L. Lions and E. Magenes, Non-homogeneous Boundary Value Problems and Applications, Vol. I, Springer-Verlag, New York, 1972.

14. J. Nitsche, "Lineare spline-funktionen und die methoden von Ritz für elliptische rendwert probleme," Arch. Rat. Mech. Anal. 36 (1970), pp. 348-355.

15. D. W. Peaceman, Fundamentals of Numerical Reservoir Simulation, Elsevier Publishing Company, 1977.

16. H. H. Rachford, Jr., "Two-level discrete-time Galerkin approximations for second order nonlinear parabolic partial differential equations," SIAM J. Numer. Anal. 10 (1973), pp. 1010-1026.

17. A. Settari, H. S. Price, and T. Dupont, "Development and application of variational methods for simulation of miscible displacement in porous media," Soc. Pet. Eng. J. (June 1977), pp. 228-246.

18. M. F. Wheeler, "A priori L^2-error estimates for Galerkin approximations to parabolic partial differential equations," SIAM J. Numer. Anal. 10 (1973), pp. 723-759.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER
1934

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle)
EFFICIENT TIME-STEPPING PROCEDURES FOR MISCIBLE DISPLACEMENT PROBLEMS IN POROUS MEDIA

5. TYPE OF REPORT & PERIOD COVERED
Summary Report - no specific reporting period

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)
Richard E. Ewing

8. CONTRACT OR GRANT NUMBER(s)
DAAG29-78-G-0161
DAAG29-75-C-0024
MCS78-09525

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Mathematics Research Center, University of Wisconsin
610 Walnut Street
Madison, Wisconsin 53706

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
Work Unit Number 7 - Numerical Analysis

11. CONTROLLING OFFICE NAME AND ADDRESS
See Item 18 below.

12. REPORT DATE
March 1979

13. NUMBER OF PAGES
39

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS. (of this report)
UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES
U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, D. C. 20550

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Galerkin methods
Error estimates
Iterative methods
Conjugate gradient methods
Fluid flow

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A model system of equations which has been used to describe the miscible displacement of one incompressible fluid by another in a porous medium is the coupled quasilinear system for $c = c(x,t)$ and $p = p(x,t)$ for $x \in \Omega$, $t \in (0,T]$ given by

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. ABSTRACT - cont'd.

$$\nabla\cdot[a(x,c)\{\nabla p - \gamma(x,c)\nabla g\}] \equiv -\nabla\cdot u = f_1(x) ,$$

$$\nabla\cdot[b(x,t)\nabla c] - u(x,c,\nabla p)\cdot\nabla c = \varphi(x)\frac{\partial c}{\partial t} - f_2(x,c) ,$$

with appropriate initial and Neumann boundary conditions. Another case considered is when $b = b(x,c,\nabla p)$ above. Iterative methods are presented and analyzed which are based on using a preconditioned conjugate gradient iteration for approximately solving the systems of linear equations produced at each time step by Galerkin methods for time-stepping the above system. Optimal order convergence rates are obtained for the iterative methods. The iterative methods are computationally more efficient than Galerkin methods previously proposed to solve the above system. The use of different time increments in the time-stepping procedures for the different variables is also presented and analyzed. The use of unequal time increments takes advantage of different smoothnesses in time of the physical variables p and c and greatly reduces the work done in the computation of the approximate solution.