SIAM J. NUMER. ANAL.
Vol. 26, No. 4, pp. 904–919, August 1989

006

ALTERNATING DIRECTION MULTISTEP METHODS FOR PARABOLIC PROBLEMS–ITERATIVE STABILIZATION*

JAMES H. BRAMBLE†, RICHARD E. EWING‡, AND GANG LI‡

Abstract. Efficient multistep procedures for time-stepping Galerkin methods for parabolic partial differential equations are presented and analyzed. The procedures involve using an alternating direction preconditioned iterative method for approximately solving the linear equations arising at each timestep in a discrete Galerkin method. Optimal order convergence rates are obtained for the iterative methods. Work estimates of almost optimal order are obtained.

Key words. parabolic problems, Galerkin methods, alternating direction, multistep, preconditioned iterative methods

AMS(MOS) subject classifications. 65M15, 65N30

1. Introduction. We will consider the numerical approximation of the solution of the quasilinear parabolic problem

$$
(1.1)\quad
\begin{aligned}
&\text{(a)}\ \ c(x,u)\frac{\partial u}{\partial t} = \sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\left(a_{ij}(x,u)\frac{\partial u}{\partial x_j}\right) + f(x,t,u), && x\in\Omega,\ t\in J,\\
&\text{(b)}\ \ u(x,0) = u_0(x), && x\in\Omega,\\
&\text{(c)}\ \ u(x,t) = 0, && x\in\partial\Omega,\ t\in J,
\end{aligned}
$$

where Ω is a bounded domain in $\mathbb{R}^d$ with boundary $\partial\Omega$, $J=(0,t_0]$, and c, f, u_0, and a_{ij}, $i,j=1,\cdots,d$, are prescribed. We will consider a Galerkin discretization in the space variable and various alternating direction multistep scheme for the time-stepping procedures. For the purpose of this paper we will assume that Ω is a rectangle or rectangular solid in $\mathbb{R}^d$. Techniques for extending the analysis to the union of rectangles and more general domains appear in [5] and [12], respectively.

Backward differentiation multistep time-stepping methods have been analyzed for equations of the form (1.1) by the first author and Sammon in [4]. In this paper, we will present and analyze alternating direction variants of multistep methods which, when the region is rectangular, are much more efficient.

The nonlinearities in the coefficients a_{ij} and c in (1.1) have essentially no effect upon the alternating direction techniques which we will analyze. Since the effects of

*Received by the editors December 7, 1987; accepted for publication March 4, 1988. This work was supported in part by U.S. Army Research Office contract DAAG29-84-K-002, by U.S. Air Force Office of Scientific Research contract AFOSR-85-0117, and by the National Science Foundation grant DMS-8504360.

†Department of Mathematics, Cornell University, Ithaca, New York 14853.

‡Department of Mathematics, University of Wyoming, Laramie, Wyoming 82071.

these nonlinearities have been carefully studied in [10] and in order to emphasize the newer alternating direction techniques, we will restrict our attention to the problem

$$
(1.2)\qquad
\begin{aligned}
&\text{(a)}\ \ c(x,t)\frac{\partial u}{\partial t} = \nabla\cdot[a(x,t)\nabla u] + f(x,t,u) \\
&\qquad\qquad \equiv -L(t)u + f(x,t,u), \quad x\in\Omega,\ t\in J,\\
&\text{(b)}\ \ u(x,0) = u_0(x), \qquad\qquad x\in\Omega,\\
&\text{(c)}\ \ u(x,t) = 0, \qquad\qquad\quad x\in\partial\Omega,\ t\in J.
\end{aligned}
$$

The time dependence in the coefficients $a = a(x,t)$ and $c = c(x,t)$ preserves essentially all of the interesting properties involved in our efficient alternating direction multistep procedures.

2. Preliminaries and notation. Let $(\varphi,\psi) = \int_\Omega \varphi\psi dx$ and $\|\psi\|^2 = (\psi,\psi)$. Let $W_s^k(\Omega)$ be the Sobolev spaces on Ω with norm

$$
\|\psi\|_{W_s^k} = \left\{ \sum_{|\alpha|\le k} \left\| \frac{\partial^\alpha \psi}{\partial x^\alpha} \right\|^2_{L^s(\Omega)} \right\}^{1/s},
$$

with the usual modification for $s=\infty$. When $s=2$, let $\|\psi\|_{W_2^k} = \|\psi\|_{H^k} = \|\psi\|_k$. If $\nabla F = (F_1,F_2)$, write $\|\nabla F\|_{W_s^k}$ in place of $\{\|F_1\|^s_{W_s^k} + \|F_2\|^s_{W_s^k}\}^{1/s}$. Let $H_0^1(\Omega)$ denote the subspace of $H^1(\Omega)$ that consists of functions that vanish (in the sense of trace) on $\partial\Omega$. We note that a norm on $H_0^1(\Omega)$ which is equivalent to $\|\cdot\|_1$ on $H_0^1(\Omega)$ is $\|\nabla\cdot\|$. If $\mathbf{X}$ is a normed space with norm $\|\cdot\|_{\mathbf{X}}$ and $\varphi : [a,b]\to\mathbf{X}$, we let

$$
\|\varphi\|_{W_p^m(a,b;\mathbf{X})} \equiv \|\ \|\varphi(\cdot,t)\|_{\mathbf{X}}\|_{W_p^m(a,b)}\,.
$$

Assume that the coefficients and the boundary, $\partial\Omega$, are sufficiently smooth that the boundary value problem (1.2) is H^2-regular. Assume that the operators $L(t)$ defined in (1.2) are uniformly positive definite. In particular for (1.2), we assume the existence of constants such that for $x\in\Omega$, $t\in J$, and $-\infty < p < \infty$,

$$
(2.1)\qquad
\begin{aligned}
&\text{(a)}\ \ 0 < a_* \le a(x,t) \le K_1,\\
&\text{(b)}\ \ 0 < c_* \le c(x,t) \le K_1,\\
&\text{(c)}\ \ |f(x,t,p)| \le K_1.
\end{aligned}
$$

For each $t\in J$, consider the linear operator $T(t) : H^s \to H^{s+2}$, $s\ge 0$, defined by $T(t)F = W$, where W is the solution of the Dirichlet problem

$$
(2.2)\qquad
\begin{aligned}
&\text{(a)}\ \ L(t)W = F, \qquad x\in\Omega,\\
&\text{(b)}\ \ W = 0, \qquad\quad\ x\in\partial\Omega.
\end{aligned}
$$

It is well known that T is uniformly bounded, i.e., there is a constant C, independent of t, such that

$$
(2.3)\qquad \|T(t)F\|_{s+2} \le C\|F\|_s, \qquad t\in J.
$$

Our spatial approximation will be a Galerkin approximation to T.

For h from a sequence of small positive numbers, let $\{S_h\}$ be a family of finite-dimensional subspaces of $H_0^1(0,1)$ that have the approximation properties: for some integer $r \geq 2$ and some constant K_0 and any $\phi \in H^q(0,1) \cap H_0^1(0,1)$,

$$\inf_{\chi \in S_h} [||\phi - \chi|| + h\,||\phi - \chi||_1] \leq K_0\,||\phi||_q\, h^q, \quad \text{for } 1 \leq q \leq r+1. \tag{2.4}$$

Let $\Delta : 0 = x_0 < x_1 < \cdots < x_N = 1$ be a partition of $I = [0,1]$ and set $I_j = [x_{j-1}, x_j]$. Let

$$M_{k,r} = \{v | v \in C^k(I), \quad v \in P_r(I_j), \quad j = 1, 2, \cdots, N\},$$

where $P_r(E)$ is the class of functions defined on I with restrictions to the set E agreeing with a polynomial of degree no greater than r. Let

$$h = \max_{1 \leq j \leq N} h_j, \qquad h_j = x_j - x_{j-1}.$$

Let

$$\mathring{M}_{k,r} = \{v \in M_{k,r} | v(0) = v(1) = 0\}.$$

In our analysis, we will take $S_h = \mathring{M}_{0,r}$.

In $\mathbb{R}^d$, let $I^d = \overbrace{I \times I \times \cdots \times I}^{d \text{ times}}$ and $M_h = \overbrace{S_h \otimes S_h \otimes \cdots \otimes S_h}^{d \text{ times}}$. Then M_h has the approximation properties: for some integer $r \geq 2$ and some constant K_1 and any $u \in H^q(I^d) \cap H_0^1(I^d)$,

$$\inf_{\chi \in M_h} [\|u - \chi\| + h\|u - \chi\|_1] \leq K_1 \|u\|_q h^q, \quad \text{for } 1 \leq q \leq r+1. \tag{2.5}$$

We next define one-dimensional projection operators P_x and $P_y : H^1(0,1) \to S_h$ by

$$\text{(2.6)} \quad \begin{aligned} &\text{(a)} \quad \int_0^1 \frac{\partial}{\partial x}(P_x u - u) \frac{\partial}{\partial x}\chi \, dx = 0, \qquad \chi \in S_h, \\ &\text{(b)} \quad \int_0^1 \frac{\partial}{\partial y}(P_y u - u) \frac{\partial}{\partial y}\chi \, dy = 0, \qquad \chi \in S_h. \end{aligned}$$

For $d = 2$, we define $Z = P_x P_y u$, the two-dimensional tensor product projection in M_h. (For $d \geq 3$, we can define tensor product projections in the same fashion.) Note that the one-dimensional operators commute and thus can be taken in any order. Noting that $(\partial/\partial x) P_y u = P_y (\partial u/\partial x)$ and $(\partial/\partial y) P_x u = P_x (\partial u/\partial y)$, we can then obtain a very important orthogonality result.

LEMMA 2.1. *If* $(\partial^2 u/\partial x \partial y) \in L^2(I^2)$, *then*

$$\left(\frac{\partial^2}{\partial x \partial y}(P_x P_y u - u), \frac{\partial^2}{\partial x \partial y}\chi\right) = 0, \quad \chi \in M_h. \tag{2.7}$$

Next, we derive some approximation properties for Z.

LEMMA 2.2. *If* $u \in H_0^1(I^2)$ *and* $\partial^2 u/\partial x \partial y \in H^p(I^2)$, *then for some* $C > 0$,

$$\left\|\frac{\partial^2}{\partial x \partial y}(Z - u)\right\| \leq C \left\|\frac{\partial^2 u}{\partial x \partial y}\right\|_p h^p, \tag{2.8}$$

$$\|Z-u\|_1 \le C\left\|\frac{\partial^2 u}{\partial x \partial y}\right\|_p h^p, \quad \textit{for } 0 \le p \le r. \tag{2.9}$$

Here we take $S_h = \mathring{M}_{k,r}$, $k \ge 0$.

Proof. Consider $V = M_{k-1,r-1} \otimes M_{k-1,r-1}$. For any $\alpha, \beta \in M_{k-1,r-1}$, there exist $\varphi, \psi \in \mathring{M}_{k,r}$ and constants a and b, such that

$$\varphi'(x) = \alpha(x) + a \quad \text{and} \quad \psi'(y) = \beta(y) + b.$$

Since $u \in H_0^1(I^2)$ and $Z \in M_h$, $Z - u = 0$ on $\partial\Omega$. Then

$$\left.\frac{\partial}{\partial x}(Z-u)\right|_{y=0,1} = 0 \quad \text{and} \quad \left.\frac{\partial}{\partial y}(Z-u)\right|_{x=0,1} = 0. \tag{2.10}$$

Then

$$\left(\frac{\partial^2}{\partial x \partial y}(Z-u), \alpha(x)\right) = \left(\frac{\partial^2}{\partial x \partial y}(Z-u), \beta(y)\right) = \left(\frac{\partial^2}{\partial x \partial y}(Z-u), 1\right) = 0.$$

Thus, since $\varphi(x)\psi(y) \in M_h$, by (2.7),

$$\left(\frac{\partial^2}{\partial x \partial y}(Z-u), \alpha(x)\beta(y)\right) = \left(\frac{\partial^2}{\partial x \partial y}(Z-u), \varphi'(x)\psi'(y)\right) = 0.$$

So, for any $v \in V$,

$$\left(\frac{\partial^2}{\partial x \partial y}(Z-u), v\right) = 0.$$

Therefore,

$$\left(\frac{\partial^2}{\partial x \partial y}(Z-u), \frac{\partial^2}{\partial x \partial y}(Z-u)\right) = -\left(\frac{\partial^2}{\partial x \partial y}(Z-u), \frac{\partial^2 u}{\partial x \partial y} - v\right), \quad v \in V.$$

Then, by the approximation properties of V,

$$\left\|\frac{\partial^2}{\partial x \partial y}(Z-u)\right\| \le \inf_{v \in V}\left\|\frac{\partial^2 u}{\partial x \partial y} - v\right\| \le C\left\|\frac{\partial^2 u}{\partial x \partial y}\right\|_p h^p, \quad \text{for } \frac{\partial^2 u}{\partial x \partial y} \in H^p(I^2),$$

$$0 \le p \le r.$$

From (2.10), we have

$$\frac{\partial}{\partial x}(Z-u)(x,y) = \int_0^y \frac{\partial^2}{\partial x \partial y}(Z-u)(x,t)\,dt, \quad \text{for } 0 \le y \le 1.$$

Then

$$\left\|\frac{\partial}{\partial x}(Z-u)\right\| \le C\left\|\frac{\partial^2}{\partial x \partial y}(Z-u)\right\| \le C\left\|\frac{\partial^2 u}{\partial x \partial y}\right\|_p h^p, \quad \text{for } 0 \le p \le r.$$

Similarly,

$$\left\|\frac{\partial}{\partial y}(Z-u)\right\| \le C\left\|\frac{\partial^2}{\partial x \partial y}(Z-u)\right\| \le C\left\|\frac{\partial^2 u}{\partial x \partial y}\right\|_p h^p, \quad \text{for } 0 \le p \le r.$$

So,

$$||Z-u||_1 \le C\left|\left|\frac{\partial^2 u}{\partial x \partial y}\right|\right|_p h^p, \quad \text{for } 0 \le p \le r.$$

This completes the proof of the lemma.

We next define W_a to be the weighted elliptic projection satisfying:

$$(a\nabla(W_a - u), \nabla\chi) = 0, \qquad \chi \in M_h. \tag{2.11}$$

Then, using the idea of [8], we obtain the following important result:

LEMMA 2.3. *For $S_h = \mathring{M}_{0,r}$, we have for some $C > 0$,*

$$||Z - W_a||_1 \le C\left|\left|\frac{\partial^2 u}{\partial x \partial y}\right|\right|_r h^{r+1}. \tag{2.12}$$

Proof. From (2.11),

$$(a\nabla(Z - W_a), \nabla\chi) = (a\nabla(Z-u), \nabla\chi), \qquad \chi \in M_h. \tag{2.13}$$

$$(a\nabla(Z-u),\nabla\chi) = \sum_{i,j=1}^{N} a\left(x_{i-1/2}, y_{j-1/2}; t\right)\int_{I_i\times I_j} \nabla(Z-u)\cdot\nabla\chi \, dxdy$$
$$+O\left(\|a(t)\|_{W_1^\infty(I^2)}\|Z-u\|_1\|\chi\|_1 h\right), \qquad t > 0.$$

Since $S_h = \mathring{M}_{0,r}$, $P_x u$, $P_y u$ interpolate u at the knots, $Z = P_x P_y u$ is locally determined. Thus,

$$\int_{I_i\times I_j}\left(\frac{\partial}{\partial x}P_xP_yu - \frac{\partial u}{\partial x}\right)\frac{\partial \chi}{\partial x}dxdy = \int_{I_i\times I_j}\left(\frac{\partial}{\partial x}P_yu - \frac{\partial u}{\partial x}\right)\frac{\partial \chi}{\partial x}dxdy$$
$$= \int_{I_i\times I_j}\left(P_y\frac{\partial u}{\partial x} - \frac{\partial u}{\partial x}\right)\frac{\partial \chi}{\partial x}dxdy.$$

Similarly,

$$\int_{I_i\times I_j}\left(\frac{\partial}{\partial y}P_xP_yu - \frac{\partial u}{\partial y}\right)\frac{\partial \chi}{\partial y}dxdy = \int_{I_i\times I_j}\left(P_x\frac{\partial u}{\partial y} - \frac{\partial u}{\partial y}\right)\frac{\partial \chi}{\partial y}dxdy.$$

By the approximation properties of the one-dimensional projections P_x and P_y, and Lemma 2.2,

$$|(a\nabla(Z-u),\nabla\chi)| \le C\,||a(t)||_{L^\infty(I^2)}\left(\left|\left|\frac{\partial u}{\partial x} - P_y\frac{\partial u}{\partial x}\right|\right| + \left|\left|\frac{\partial u}{\partial y} - P_x\frac{\partial u}{\partial x}\right|\right|\right)||\chi||_1$$
$$+ C\,||a(t)||_{W_1^\infty(I^2)}\left|\left|\frac{\partial^2 u}{\partial x\partial y}\right|\right|_r ||\chi||_1\, h^{r+1}$$
$$\le C\,||a(t)||_{W_1^\infty(I^2)}\left|\left|\frac{\partial^2 u}{\partial x\partial y}\right|\right|_r ||\chi||_1\, h^{r+1}, \quad t > 0.$$

Taking $\chi = Z - W_a \in M_h$, then from (2.13),

$$||Z - W_a||_1 \le C\left|\left|\frac{\partial^2 u}{\partial x \partial y}\right|\right|_r h^{r+1}.$$

Using the above lemmas and the approximation properties of the weighted elliptic projection W_a, as in [9], we have

LEMMA 2.4. *For sufficiently smooth functions a and u, there exists $C(u) > 0$, such that*

$$\left\| \frac{\partial^k}{\partial t^k}(Z-u) \right\|_s \leq C(u)h^{r+1-s}, \text{ for } k \geq 0, \ s = 0 \text{ or } 1. \tag{2.14}$$

Also, as in [7], we have that $(\partial^k Z/\partial t^k)$, $(\partial^k W_a/\partial t^k)$ are bounded in various norms, for $k \geq 0$.

Let T_h be the Galerkin approximation of T on M_h, then we have for some $C > 0$ such that

$$\|(T-T_h)f\|_0 \leq Ch^s\|f\|_{s-2}, \qquad 2 \leq s \leq r+1.$$

In particular, the error measured in the L_2 operator norm, denoted by $\|\cdot\|_{0,0}$, satisfies

$$\|T - T_h\|_{0,0} \leq Ch^2.$$

T_h is invertible on M_h, and we can set $L_h(t) = T_h(t)^{-1}$. We then have

$$(L_h(t)\varphi, \psi) = (a(t)\nabla\varphi, \nabla\psi), \text{ for } \varphi, \psi \in M_h.$$

Let $L_{1,h}$, $L_{2,h}$ be defined as operators from M_h into M_h as follows:

$$(L_{1,h}\varphi, \psi) = \left(\frac{\partial\varphi}{\partial x}, \frac{\partial\psi}{\partial x}\right),$$

$$(L_{2,h}\varphi, \psi) = \left(\frac{\partial\varphi}{\partial y}, \frac{\partial\psi}{\partial y}\right), \text{ for } \varphi, \psi \in M_h.$$

Then

$$L_h(t) = \bar{a}\,(L_{1,h} + L_{2,h}), \text{ for } a = \bar{a} = \text{ constant},$$
$$(L_{1,h}L_{2,h}\varphi, \psi) = \left(\frac{\partial^2\varphi}{\partial x\partial y}, \frac{\partial^2\psi}{\partial x\partial y}\right), \text{ for } \varphi, \psi \in M_h.$$

Note that $L_{1,h}$, $L_{2,h}$ commute.

Let $k > 0$ and $N = t_0/k \in \mathbb{Z}$ and $t^\sigma = \sigma k$, $\sigma \in \mathbb{R}$. Also let $\varphi^n \equiv \varphi^n(x) \equiv \varphi(x, t^n)$. Also define the following backward difference operators

$$\begin{aligned} &\text{(a)} \quad \delta\varphi^n = \varphi^n - \varphi^{n-1} \\ &\text{(b)} \quad \delta^2\varphi^n = \varphi^n - 2\varphi^{n-1} + \varphi^{n-2} \\ &\text{(c)} \quad \delta^3\varphi^n = \varphi^n - 3\varphi^{n-1} + 3\varphi^{n-2} - \varphi^{n-3} \\ &\text{(d)} \quad \delta^4\varphi^n = \varphi^n - 4\varphi^{n-1} + 6\varphi^{n-2} - 4\varphi^{n-3} + \varphi^{n-4}. \end{aligned} \tag{2.15}$$

Then define $d_t\varphi^n = \delta\varphi^n/k$.

3. Description of the methods. In this section we will describe various methods for efficiently time-stepping the Galerkin spatial procedures. We will first present several multistep methods that will form our base schemes. We will next introduce terms that will make the spatial operators comparable to alternating direction variants of the base schemes. We then use the iterative stabilization ideas which were introduced in [7], [10] and used for multistep methods in [4], [10] to define very efficient alternating direction variants of the basic multistep schemes.

For various special choices of parameters, we define the following class of extrapolated coefficient, backward differentiation, multistep discrete time methods. Let $U : \{t^0, \cdots, t^N\} \to M_h$ be an approximate solution of (1.2). Assume that U^k are known for $k \leq n$. Given a desired global time truncation error of order k^μ, $\mu = 1, 2, 3$, we choose parameters $\alpha_i(\mu)$, $i = 1, 2$, and $\beta(\mu)$ and an extrapolation operator $E(\mu)$ for the coefficients to define a method for determining U^{n+1} which satisfies

$$\begin{aligned} &P_0 c^{n+1} d_t U^{n+1} + \beta L_h(t^{n+1}) U^{n+1} \\ &\quad = k^{-1} P_0 c^{n+1} \left[\alpha_1 \delta U^n + \alpha_2 \delta U^{n-1}\right] + \beta P_0 f\left(t^{n+1}, E(\mu) U^{n+1}\right), \end{aligned} \tag{3.1}$$

where P_0 is the L^2-projection onto M_h and $c^{n+1} = c(t^{n+1})$. Choices of the parameters and extrapolation operators for $\mu = 1, 2, 3$ are given in the following table.

TABLE 1
Backward differentiation multistep methods.

μ	$\beta(\mu)$	$\alpha_1(\mu)$	$\alpha_2(\mu)$	$E(\mu)U^{n+1}$
1	1	0	0	$U^{n+1} - \delta U^{n+1}$
2	2/3	1/3	0	$U^{n+1} - \delta^2 U^{n+1}$
3	6/11	7/11	−2/11	$U^{n+1} - \delta^3 U^{n+1}$

We note that by extrapolating the values of U^k in the nonlinear term f, we have produced a linear operator equation for U^{n+1} in terms of the previous known values of U^k, $k \leq n$. See [4] for a detailed analysis of the stability and accuracy of corresponding methods where a and c depend upon U^{n+1} and must also contain the extrapolation $E(\mu)U^{n+1}$ to achieve linear equations.

If the coefficients a and c in (3.1) were constants (with constant values $\overline{a}$ and $\overline{c}$) then typical alternating direction variants of (3.1) could be derived by perturbing (3.1) as follows. Let U^{n+1} satisfy (for $d = 2$)

$$\begin{aligned} &P_0 \overline{c} d_t U^{n+1} + \beta L_h(t^{n+1}) U^{n+1} + k P_0 \frac{\beta^2 \overline{a}^2}{\overline{c}} L_{1,h} L_{2,h} D(\mu) U^{n+1} \\ &\quad = k^{-1} P_0 \overline{c} \left[\alpha_1 \delta U^n + \alpha_2 \delta U^{n-1}\right] + \beta P_0 f\left(t^{n+1}, E(\mu) U^{n+1}\right), \end{aligned} \tag{3.2}$$

where the operators $D(\mu)U^{n+1}$ make the additional term "small" enough so as not to increase the order of error already present in the approximation. For example, for $\mu = 1$ or $\mu = 2$, the choice $D(\mu)U^{n+1} = \delta U^{n+1}$ will yield convergent schemes. For $\mu = 3$, we will use $D(3)U^{n+1} = \delta^2 U^{n+1}$.

Recall that $P_0 L_h = L_h = \overline{a}(L_{1,h} + L_{2,h})$. Then, after multiplication by k, (3.2) will factor to yield

$$\begin{aligned} &P_0 (\overline{c}^{1/2} + k\beta\overline{a}\, \overline{c}^{-1/2} L_{1,h}) \left(\overline{c}^{1/2} + k\beta\overline{a}\, \overline{c}^{-1/2} L_{2,h}\right) U^{n+1} \\ &\quad \equiv P_0 S_{1,h} S_{2,h} U^{n+1} = F^n, \end{aligned} \tag{3.3}$$

where $F^n = F^n(t^n, k, \overline{c}, \beta, \mu, \alpha_i, \delta U^{n-1}; i = 0, 1, 2)$. Since $\overline{a}$ and $\overline{c}$ are constants and thus $S_{1,h}$ and $S_{2,h}$ each depend only upon a single space variable, (3.3) can be solved as two successive one-dimensional problems.

We will use (3.2) and (3.3) as motivation for our base schemes. Let the constants $\overline{a}$ and $\overline{c}$ be the spatial mean values of a^o and c^o and let

$$\lambda = \beta^2 \overline{a}^2 \overline{c}^{-1}.$$

Then as our base alternating direction methods, we define U^{n+1} to satisfy

$$\begin{aligned}(3.4)\qquad & P_0 c^{n+1} d_t U^{n+1} + \beta L_h(t^{n+1})U^{n+1} + kP_0\lambda L_{1,h}L_{2,h}D(\mu)U^{n+1} \\ &\quad = k^{-1}P_0 c^{n+1}\left[\alpha_1\delta U^n + \alpha_2\delta U^{n-1}\right] + \beta P_0 f\left(t^{n+1}, E(\mu)U^{n+1}\right).\end{aligned}$$

We note that the operators in (3.4) do not factor and thus cannot be used as an alternating direction method in that form. Also the methods described in (3.4) require the solution of a different linear system of equations at each timestep. However, the operators in (3.4) are comparable to the factorable operators in (3.2) and (3.3) which are time independent. Thus using the ideas presented in [7], [10], we will use the matrices arising from methods like (3.3) as preconditioners in an iterative procedure for only approximating the solution of the base schemes (3.4). The resulting algorithm will be shown to be stable and very efficient.

4. Iterative stabilization procedures. In this section, we will present the linear equations arising from (3.4). We note that the coefficient matrices change with each timestep. Also there is no direct alternating direction factorization possible for the linear equations arising from (3.4). To avoid factorization of different matrices at each timestep for solution of the linear equations and to make use of the tensor product form of the basis to factor our problems, we will discuss a preconditioned iterative method for approximating the solution of the linear equations to within sufficient accuracy. We will use matrices similar to those arising from (3.3) as preconditioning matrices.

We will consider the case when $d = 2$ and $\Omega = [0,1]^2 \equiv I^2$. Extensions for $d \geq 3$ are straightforward. Techniques for generalizations to unions of rectangles can be found in [5].

We define two orderings on the nodes in Ω. The first is a global ordering which assigns one of the numbers $1, 2, \cdots, M$ to each node in Ω. The second is a tensor product ordering of the M nodes. Grid lines of the form $x = x_j$, $0 \leq x_j \leq 1$, are numbered $1, 2, \cdots, M_x$ while the grid lines of the form $y = y_j$, $0 \leq y_j \leq 1$, are numbered $1, 2, \cdots, M_y$. With each node i, we associate an x-grid line and a y-grid line. The tensor product index of the node i is the pair $(m(i),\ n(i))$ where $m(i)$ is the number of the x-grid line and $n(i)$ is the number of the y-grid line. We then denote the tensor product basis as $B_i(\vec{x}) = \varphi_{m(i)}(x)\psi_{n(i)}(y) = \varphi_m(x)\psi_n(y)$, $1 \leq i \leq M$, where $\{\varphi_m(x)\}_{m=1}^{M_x}$ and $\{\psi_n(y)\}_{n=1}^{M_y}$ are bases for the one-dimensional spaces S_h for $x \in I$ and $y \in I$, respectively.

Let U^ℓ from (3.4) be written as

$$(4.1)\qquad U^\ell = \sum_{i=1}^{M} \xi_i^\ell B_i(\vec{x}) = \sum_{m=1}^{M_x}\sum_{n=1}^{M_y} \xi_{mn}^\ell \varphi_m(x)\psi_n(y).$$

Using (4.1), (3.4) can be written as

$$\begin{aligned}(4.2)\qquad L^{n+1}\left\{\xi^{n+1} - \xi^n\right\} &= C^{n+1}\left\{\sum_{j=1}^{2}\alpha_j\delta\xi^{n+1-j}\right\} + k\left\{F_1^n(\xi) + F_2^n(\xi)\right\} \\ &\equiv F^n(\xi),\end{aligned}$$

where the matrices and vectors are of the form

$$
\begin{aligned}
&\text{(a)}\quad L^n = C^n + kA^n + k^2\lambda G,\\
&\text{(b)}\quad C^n = \left((c(t^n)B_j, B_i)\right),\\
&\text{(c)}\quad A^n = \left(\beta\left(a(t^n)\nabla B_j, \nabla B_i\right)\right),\\
&\text{(d)}\quad G = \left(\left(\frac{\partial^2}{\partial x \partial y}B_j, \frac{\partial^2}{\partial x \partial y}B_i\right)\right),\\
&\text{(e)}\quad F_1^n(\xi) = -A^{n+1}\xi^n + \beta\left(\left(f\left(t^{n+1}, \sum_{i=1}^{M}\xi_i^{n+1}B_i - \delta^{\mu}\sum_{i=1}^{M}\xi_i^{n+1}B_i\right), B_j\right)\right),\\
&\text{(f)}\quad F_2^n(\xi) = 0, \text{ if } \mu = 1 \text{ or } \mu = 2,\\
&\text{(g)}\quad F_2^n(\xi) = G\left[\xi^n - \xi^{n-1}\right], \text{ if } \mu = 3,
\end{aligned}
\tag{4.3}
$$

for $i, j = 1, 2, \cdots, M$.

Instead of solving (4.2) exactly, we will approximate its solution by using an iterative procedure which has been preconditioned by $\overline{L}^0$, the matrix with constant coefficients used in (3.3), for each timestep. Since the matrix $\overline{L}^0$ has constant coefficients, we can use the tensor product property of the basis to factor $\overline{L}^0$ into the product

$$
\overline{L}^0 = (C_x + kA_x) \otimes (C_y + kA_y), \tag{4.4}
$$

where

$$
\begin{aligned}
&\text{(a)}\quad C_x = \left(\overline{c}^{1/2}(\varphi_j(x), \varphi_i(x))\right),\\
&\text{(b)}\quad A_x = \left(\beta\overline{a}\,\overline{c}^{-1/2}(\varphi_j'(x), \varphi_i'(x))\right),\\
&\text{(c)}\quad C_y = \left(\overline{c}^{1/2}(\psi_n(y), \psi_m(y))\right),\\
&\text{(d)}\quad A_y = \left(\beta\overline{a}\,\overline{c}^{-1/2}(\psi_n'(y), \psi_m'(y))\right),
\end{aligned}
\tag{4.5}
$$

for $i, j = 1, \cdots, M_x$, $m, n = 1, \cdots, M_y$, and $\lambda = \beta^2\overline{a}^2\overline{c}^{-1}$ from (3.3). Thus, inverting $\overline{L}^0$ corresponds to solving two one-dimensional problems.

The preconditioning process eliminates the need for factoring new matrices at each timestep and reduces the problem to successive solution of one-dimensional problems, while the iterative procedure stabilizes the resulting problem. The stabilization process requires iteration only until a predetermined norm reduction is achieved.

Denote by

$$
V^{\ell} = \sum_{i=1}^{M}\theta_i^{\ell}B_i(\overrightarrow{x}) = \sum_{m=1}^{M_x}\sum_{n=1}^{M_y}\theta_{mn}^{\ell}\varphi_m(x)\psi_n(y) \tag{4.6}
$$

the approximation to U^{ℓ} produced by only approximately solving (4.2). An iterative procedure for obtaining the necessary V^{ℓ} starting values will be discussed elsewhere. We assume that such a starting procedure has been used to obtain sufficiently accurate (see (5.9)) starting values. Thus assume $V^0, \cdots, V^n$ have been determined. We will determine the M-dimensional vector θ^{n+1} (and thus V^{n+1}) using a preconditioned iterative method to approximate ξ^{n+1} from (4.2). As an initial guess for $\xi^{n+1} - \xi^n$ we will extrapolate from previously determined values. Specifically, for a particular method having time-truncation error k^{μ}, we will use as the initialization for our iterative procedure

$$
x_0 = (\theta^{n+1} - \theta^n) - \delta^{\mu+1}\theta^{n+1},
$$

where the mth backward difference operator is defined in (2.15) for $m = 1, \cdots, 4$. Since we are using previously determined θ^i in the matrix problem (4.2) to determine θ^{n+1}, our errors accumulate.

To estimate the cumulative error, we first consider the single step error. We define $\overline{\theta}^{n+1}$ to satisfy

$$L^{n+1}\left\{\overline{\theta}^{n+1} - \theta^n\right\} = F^n(\theta), \quad n \geq \mu. \tag{4.7}$$

We can use any preconditioned iterative method that yields norm reductions of the form

$$\left|\left|(L^{n+1})^{1/2}\left(\overline{\theta}^{n+1} - \theta^{n+1}\right)\right|\right|_e \leq \rho_n \left|\left|(L^{n+1})^{1/2}\left(\overline{\theta}^{n+1} - \theta^{n+1} + \delta^{\mu+1}\theta^{n+1}\right)\right|\right|_e, \tag{4.8}$$

where $0 < \rho_n < 1$ and the subscript e denotes the Euclidean norm of the vector. A specific iterative procedure for obtaining (4.8) is the preconditioned conjugate gradient method analyzed in [1], [2], [7], [10].

Let

$$\begin{aligned}
&\text{(a)} \quad ||\varphi||_{c^n}^2 \equiv (c(t^n)\varphi, \varphi), \\
&\text{(b)} \quad ||\varphi||_{a^n}^2 \equiv (a(t^n)\nabla\varphi, \nabla\varphi), \\
&\text{(c)} \quad |||\varphi|||_n \equiv ||\varphi||_{c^n} + k^{1/2}\,||\varphi||_{a^n}, \\
&\text{(d)} \quad |||\varphi||| \equiv ||\varphi|| + k^{1/2}\,||\varphi||_1, \\
&\text{(e)} \quad |||\varphi|||_{S^n} \equiv |||\varphi|||_n + k\lambda^{1/2}\left|\left|\frac{\partial^2\varphi}{\partial x\partial y}\right|\right|, \\
&\text{(f)} \quad |||\varphi|||_S \equiv |||\varphi||| + k\lambda^{1/2}\left|\left|\frac{\partial^2\varphi}{\partial x\partial y}\right|\right|
\end{aligned} \tag{4.9}$$

be special norms. Note that $||\cdot||_{c^n}$, $||\cdot||_{a^n}$, $|||\cdot|||_n$, and $|||\cdot|||_{S^n}$ are uniformly equivalent to $||\cdot||$, $||\nabla\cdot||$, $|||\cdot|||$, and $|||\cdot|||_S$, respectively. Then letting

$$\overline{V}^{\ell} = \sum_{i=1}^{M} \overline{\theta}_i^{\ell} B_i(\vec{x}) = \sum_{m=1}^{M_x}\sum_{n=1}^{M_y} \overline{\theta}_{mn}^{\ell}\varphi_m(x)\psi_n(y), \tag{4.10}$$

with $\overline{\theta}^{\ell}$ defined in (4.7), we see that $\overline{V}^{n+1}$ satisfies

(4.11)

$$\begin{aligned}
P_0 c^{n+1}\frac{\overline{V}^{n+1} - V^n}{k} &+ \beta L_h(t^{n+1})\overline{V}^{n+1} + kP_0\lambda L_{1,h}L_{2,h}\left[\overline{V}^{n+1} - V^{n+1} + D(\mu)V^{n+1}\right] \\
&= k^{-1}P_0 c^{n+1}\sum_{i=1}^{2}\alpha_i\delta V^{n+1-i} + \beta P_0 f\left(t^{n+1}, E(\mu)V^{n+1}\right)
\end{aligned}$$

on M_h. Also using (4.9), our single step error (4.8) satisfies

$$\left|\left|\left|\overline{V}^{n+1} - V^{n+1}\right|\right|\right|_{S^{n+1}} \leq \frac{\rho_n}{1-\rho_n}\,|||\delta^{\mu+1}V^{n+1}|||_{S^{n+1}}, \quad n \geq \mu. \tag{4.12}$$

We note that as in [1], [2], [7], [10], there is a Q, depending upon bounds for the coefficients, such that

$$\begin{aligned}
&\text{(a)} \quad \rho_n < 2Q^{\kappa}, \text{ with } 0 < Q < 1, \\
&\text{(b)} \quad \frac{\rho_n}{1-\rho_n} \equiv \rho_n' < 1, \text{ for } \kappa \text{ sufficiently large.}
\end{aligned} \tag{4.13}$$

5. A priori error estimates. In this section we develop a priori bounds for the errors $V^n - u^n$ for the procedure defined in (4.11) using the base schemes defined in (3.4). We will present details for the special case of (3.4) where $\mu = 3$. Similar techniques will treat the case where $\mu = 2$ and $\mu = 1$.

For $\mu = 3$, using the definitions of $L_{1,h}$, $L_{2,h}$, and L_h, the base approximation scheme (3.4) can be written as

(5.1)
$$(c^{n+1}d_tU^{n+1}, \chi) + \frac{6}{11}(a^{n+1}\nabla U^{n+1}, \nabla\chi) + k\lambda\left(\frac{\partial^2}{\partial x\partial y}\delta^2 U^{n+1}, \frac{\partial^2}{\partial x\partial y}\chi\right)$$
$$= \left(c^{n+1}\left[\frac{7}{11}d_tU^n - \frac{2}{11}d_tU^{n-1}\right], \chi\right) + \frac{6}{11}(f(t^{n+1}, U^{n+1} - \delta^3 U^{n+1}), \chi), \chi \in M_h.$$

We have chosen $D(3)$ in (3.4) to be $\delta^2 U^{n+1}$ in this case. We know from Lemma 2.4 that Z is a function in M_h that is sufficiently close to u to obtain optimal order spatial estimates. We next estimate how close Z and V are.

Let $\eta^n = u^n - Z^n$ and $E^n = V^n - Z^n$. From (1.2), (4.11), and (5.1), we obtain the following error equation:

$$(c^{n+1}d_tE^{n+1}, x) + \frac{6}{11}(a^{n+1}\nabla E^{n+1}, \nabla\chi) + k\lambda\left(\frac{\partial^2}{\partial x\partial y}\delta^2 E^{n+1}, \frac{\partial^2}{\partial x\partial y}\chi\right)$$
$$= \left(c^{n+1}\left[\frac{7}{11}d_tE^n - \frac{2}{11}d_tE^{n-1}\right], \chi\right)$$
$$+ \left(c^{n+1}\left[d_t\eta^{n+1} - \frac{7}{11}d_t\eta^n + \frac{2}{11}d_t\eta^{n-1}\right], \chi\right)$$
$$+ \left(c^{n+1}\left[\frac{6}{11}\frac{\partial u^{n+1}}{\partial t} - d_tu^{n+1} + \frac{7}{11}d_tu^n - \frac{2}{11}d_tu^{n-1}\right], \chi\right)$$
$$+ \frac{6}{11}\left(a^{n+1}\nabla\left(W_a^{n+1} - Z^{n+1}\right), \nabla\chi\right) - k\lambda\left(\frac{\partial^2}{\partial x\partial y}\delta^2 Z^{n+1}, \frac{\partial^2}{\partial x\partial y}\chi\right) \tag{5.2}$$
$$+ \frac{6}{11}\left([f(t^{n+1}, V^{n+1} - \delta^3 V^{n+1}) - f(t^{n+1}, u^{n+1})], \chi\right)$$
$$+ \left[k^{-1}\left(c^{n+1}(V^{n+1} - \overline{V}^{n+1}), \chi\right) + \frac{6}{11}\left(a^{n+1}\nabla(V^{n+1} - \overline{V}^{n+1}), \nabla\chi\right)\right.$$
$$\left. + k\lambda\left(\frac{\partial^2}{\partial x\partial y}(V^{n+1} - \overline{V}^{n+1}), \frac{\partial^2}{\partial x\partial y}\chi\right)\right]$$
$$\equiv \left(c^{n+1}\left[\frac{7}{11}d_tE^n - \frac{2}{11}d_tE^{n-1}\right], \chi\right) + T_1^{n+1}(\chi) + T_2^{n+1}(\chi)$$
$$+ T_3^{n+1}(\chi) + T_4^{n+1}(\chi) + T_5^{n+1}(\chi) + T_6^{n+1}(\chi), \quad \chi \in M_h.$$

Term T_1^{n+1} enters because we are comparing V to Z instead of directly to u. Term T_2^{n+1} measures how well the multistep scheme approximates $(\partial u/\partial t)$. Term T_3^{n+1} measures the difference between the different projections we have used and T_4^{n+1} arises from the artificial term we placed in our base scheme for comparability to an alternating direction scheme, T_5^{n+1} contains the contribution from the forcing function f. Finally, the single-step error made by using the iterative procedure to approximately solve the linear equations appears in T_6^{n+1}.

We will present a few lemmas which will help to separate the various parts of our analysis. First, we note that the parameters $\beta(\mu)$ and $\alpha_i(\mu)$, $i = 1, 2$, from Table 1 were chosen to ensure the following consistency results.

LEMMA 5.1. *For each $\mu = 1, 2, 3$, the choice of parameters $\beta(\mu)$ and $\alpha_i(\mu)$, $i = 1, 2$, given in Table 1 yields*

$$\left|\left|\beta(\mu)\frac{\partial u^{n+1}}{\partial t} - \left[d_t u^{n+1} - \sum_{i=1}^{2} \alpha_i(\mu) d_t u^{n+1-i}\right]\right|\right| \leq K(u)k^{\mu}. \tag{5.3}$$

For various results, since our norms depend upon n we will need the following trivial shift lemma.

LEMMA 5.2. *For the norms defined in (4.9), we have, for $\chi \in M_h$,*

$$\begin{aligned}
||\chi^m||_{c^n}^2 &\leq ||\chi^m||_{c^{n-1}}^2 + K\,||\chi^m||^2\, k,\\
||\chi^m||_{a^n}^2 &\leq ||\chi^m||_{a^{n-1}}^2 + K\,||\chi^m||_1^2\, k,\\
|||\chi^m|||_n^2 &\leq |||\chi^m|||_{n-1}^2 + K\,|||\chi^m|||^2\, k,\\
|||\chi^m|||_{S^n}^2 &\leq |||\chi^m|||_{S^{n-1}}^2 + K\,|||\chi^m|||_S^2\, k.
\end{aligned} \tag{5.4}$$

We next present a lemma which will provide the estimates for the basic stability of our method. We consider the cases $\mu \leq 3$.

LEMMA 5.3. *Assume that $\zeta^n \in M_h$ satisfies, for $m \geq 2$, $n = m, \cdots, \ell - 1$,*

$$\begin{aligned}
(c^{n+1} d_t \zeta^{n+1}, \chi) + \beta(\mu)\,(a^{n+1}\nabla\zeta^{n+1}, \nabla\chi) + k\lambda\left(\frac{\partial^2}{\partial x\partial y}\delta^2\zeta^{n+1}, \frac{\partial^2}{\partial x \partial y}\chi\right)&\\
= \left(c^{n+1}\sum_{i=1}^{2}\alpha_i(\mu) d_t\zeta^{n+1-i}, \chi\right) + (G^{n+1}, \chi)\,,\ \ \chi \in M_h&.
\end{aligned} \tag{5.5}$$

Then, if $\alpha \equiv \sum_{i=1}^{2} |\alpha_i| < 1$, there exist positive constants C_1 and τ_0 such that, if $k \leq \tau_0$,

$$\begin{aligned}
\sum_{n=m}^{\ell-1} k\left[|||d_t\zeta^{n+1}|||^2 + \left|\left|\frac{\partial^2}{\partial x \partial y}\delta^2\zeta^{n+1}\right|\right|^2\right] + ||\zeta^\ell||_1^2 + \left|\left|\frac{\partial^2}{\partial x\partial y}\delta\zeta^\ell\right|\right|^2 k&\\
\leq C_1\left[IC(m) + \sum_{n=m}^{\ell-1} ||\zeta^n||_1^2\, k + \left|\sum_{n=m}^{\ell-1} (G^{n+1}, d_t\zeta^{n+1})\, k\right|\right]&,
\end{aligned} \tag{5.6}$$

where $IC(m)$ refers to the initial values of the same quantities that are estimated for $t \leq t^m$,

$$IC(m) = K\left[||\zeta^m||_1^2 + k\left|\left|\frac{\partial^2}{\partial x \partial y}\delta\zeta^m\right|\right|^2 + k\left(||d_t\zeta^m||^2 + ||d_t\zeta^{m-1}||^2\right)\right].$$

Proof. Let $\chi = \delta\zeta^{n+1} = k d_t\zeta^{n+1}$ in (5.5) and sum it for $m \leq n \leq \ell - 1$ to obtain

$$\begin{aligned}
\sum_{n=m}^{\ell-1}\Bigg[k\,||d_t\zeta^{n+1}||_{c^{n+1}}^2 + \frac{\beta(\mu)}{2}\left\{k^2\,||d_t\zeta^{n+1}||_{a^{n+1}}^2 + \left[||\zeta^{n+1}||_{a^{n+1}}^2 - ||\zeta^n||_{a^{n+1}}^2\right]\right\}&\\
+ \frac{k\lambda}{2}\left\{\left|\left|\frac{\partial^2}{\partial x\partial y}\delta^2\zeta^{n+1}\right|\right|^2 + \left[\left|\left|\frac{\partial^2}{\partial x \partial y}\delta\zeta^{n+1}\right|\right|^2 - \left|\left|\frac{\partial^2}{\partial x \partial y}\delta\zeta^n\right|\right|^2\right]\right\}\Bigg]&\\
= \sum_{n=m}^{\ell-1}\left[\left(c^{n+1}\sum_{i=1}^{2}\alpha_i(\mu) d_t\zeta^{n+1-i}, d_t\zeta^{n+1}\right) k + (G^{n+1}, d_t\zeta^{n+1})\, k\right]&.
\end{aligned} \tag{5.7}$$

Note that by Cauchy's inequality, inequalities between the arithmetic and geometric means, and Lemma 5.2,

$$\begin{aligned}
&\sum_{n=m}^{\ell-1}\left(c^{n+1}\sum_{i=1}^{2}\alpha_i(\mu)d_t\zeta^{n+1-i}, d_t\zeta^{n+1}\right)k \\
&\le \sum_{n=m}^{\ell-1}\sum_{i=1}^{2}|\alpha_i|\,||d_t\zeta^{n+1-i}||_{c^{n+1}}\,||d_t\zeta^{n+1}||_{c^{n+1}}\,k \\
&\le \alpha\sum_{n=m-2}^{\ell-1}||d_t\zeta^{n+1}||_{c^{n+1}}^2\,k + K\sum_{n=m-2}^{\ell-1}||d_t\zeta^{n+1}||_{c^{n+1}}^2\,k^2.
\end{aligned}\tag{5.8}$$

Finally using Lemma 5.2 to obtain a telescoping sum, noting the definitions of the norms, and taking k sufficiently small, we can combine (5.7) and (5.8) to yield the desired result.

We next state the major result for the case $\mu = 3$.

THEOREM 5.1. *Let u and U satisfy* (1.2) *and* (5.1), *respectively. Let V be the iterative variant of U satisfying* (4.11), (4.12), *and* (5.2) *with $\rho_n' \le k$. Assume that a start-up procedure is used which satisfies*

$$\begin{aligned}
&||E^3||_1^2 + k\left[\sum_{n=2}^{3}||d_tE^n||^2 + \left|\left|\frac{\partial^2}{\partial x\partial y}\delta E^3\right|\right|^2\right] \\
&+ k^2\left[||d_tE^1||^2 + k\sum_{n=1}^{3}||d_tE^n||_1^2 + \sum_{n=1}^{2}\left|\left|\frac{\partial^2}{\partial x\partial y}\delta E^n\right|\right|^2\right] \le Kk^6.
\end{aligned}\tag{5.9}$$

Then there exist constants $C_2(u)$, depending upon various norms of u, and τ_0 such that, if $k \le \tau_0$,

$$\sup_n ||u^n - V^n|| \le C_2(u)\left[h^{r+1} + k^3\right].\tag{5.10}$$

Proof. Letting $\chi = \delta E^{n+1} = kd_tE^{n+1}$ in (5.2) and applying Lemma 5.3 with $m = 3$ and (5.9), we obtain

$$\begin{aligned}
&||E^\ell||_1^2 + \left|\left|\frac{\partial^2}{\partial x\partial y}\delta E^\ell\right|\right|^2 k + \sum_{n=3}^{\ell-1}|||d_tE^{n+1}|||^2\,k \\
&\le C\left[k^6 + \sum_{n=3}^{\ell-1}||E^n||_1^2\,k + \left|\sum_{n=3}^{\ell-1}\sum_{i=1}^{6}T_i^{n+1}(\delta E^{n+1})\right|\right].
\end{aligned}\tag{5.11}$$

Using Lemma 2.4, we see that

$$\begin{aligned}
\sum_{n=3}^{\ell-1}\left|T_1^{n+1}\left(kd_tE^{n+1}\right)\right| &\le \sum_{n=3}^{\ell-1}\sum_{i=0}^{2}||d_t\eta^{n+1-i}||_{c^{n+1}}\,||d_tE^{n+1}||_{c^{n+1}}\,k \\
&\le Ch^{2r+2} + \epsilon_1\sum_{n=3}^{\ell-1}|||d_tE^{n+1}|||^2\,k.
\end{aligned}\tag{5.12}$$

Similarly, Lemma 5.1 yields

$$\sum_{n=3}^{\ell-1}\left|T_2^{n+1}\left(kd_tE^{n+1}\right)\right| \le Ck^6 + \epsilon_1\sum_{n=3}^{\ell-1}|||d_tE^{n+1}|||^2\,k.\tag{5.13}$$

We then let $\xi^n = W_a^n - Z^n$ and use Lemma 2.3 and Lemma 5.2 to see that

$$\begin{aligned}
&\left|\sum_{n=3}^{\ell-1} T_3^{n+1}\left(\delta E^{n+1}\right)\right| \\
&\quad = \left|\sum_{n=3}^{\ell-1} \frac{6}{11}\left[\left(a^{n+1}\nabla\xi^{n+1}, \nabla E^{n+1}\right) - \left(a^n\nabla\xi^n, \nabla E^n\right)\right]\right. \\
&\qquad \left. - \sum_{n=3}^{\ell-1} \frac{6}{11}\left(a^{n+1}\nabla d_t\xi^{n+1}, \nabla E^n\right)k - \sum_{n=3}^{\ell-1} \frac{6}{11}\left(\left(a^{n+1} - a^n\right)\nabla\xi^n, \nabla E^n\right)\right| \\
&\quad \le \epsilon_2 \left\|E^\ell\right\|_1^2 + \left\|E^3\right\|_1^2 + C\left\{h^{2r+2} + \sum_{n=3}^{\ell-1}\left\|E^n\right\|_1^2 k\right\}.
\end{aligned} \tag{5.14}$$

The smoothness of u, projection properties of Z given in (2.7), and summation by parts in time then yield

$$\begin{aligned}
\left|\sum_{n=3}^{\ell-1} T_4^{n+1}\left(\delta E^{n+1}\right)\right| &= \left|k\lambda\sum_{n=3}^{\ell-1}\left(\frac{\partial^2}{\partial x\partial y}\delta^2\left(Z^{n+1} - u^{n+1} + u^{n+1}\right), \frac{\partial^2}{\partial x\partial y}\delta E^{n+1}\right)\right| \\
&= \left|-k\lambda\sum_{n=3}^{\ell-1}\left(\frac{\partial^3}{\partial x^2\partial y}\delta^3 u^{n+1}, \frac{\partial}{\partial y}E^n\right) + k\lambda\left(\frac{\partial^3}{\partial x^2\partial y}\delta^2 u^\ell, \frac{\partial}{\partial y}E^\ell\right)\right. \\
&\qquad \left. - k\lambda\left(\frac{\partial^3}{\partial x^2\partial y}\delta^2 u^3, \frac{\partial}{\partial y}E^3\right)\right| \\
&\le \epsilon_2\left\|E^\ell\right\|_1^2 + \left\|E^3\right\|_1^2 + C\left(\|u\|_{H^3(J;H^3)}\right)k^6 + C\sum_{n=3}^{\ell-1}\left\|E^n\right\|_1^2 k.
\end{aligned} \tag{5.15}$$

We easily see that Lemma 2.4 and smoothness of Z yield

$$\begin{aligned}
\left|\sum_{n=3}^{\ell-1} T_5^{n+1}\left(kd_tE^{n+1}\right)\right| &\le \epsilon_1\sum_{n=3}^{\ell-1}\left|\left|\left|d_tE^{n+1}\right|\right|\right|^2 k \\
&\quad + C\left\{h^{2r+2} + k^6 + \sum_{n=3}^{\ell-1}\left\|E^{n+1}\right\|_1^2 k\right\}.
\end{aligned} \tag{5.16}$$

Next, using (4.9), (4.12), and (4.13), we see that

$$\begin{aligned}
\left|\sum_{n=3}^{\ell-1} T_6^{n+1}\left(kd_tE^{n+1}\right)\right| &\le \sum_{n=3}^{\ell-1}\left|\left|\left|V^{n+1} - \overline{V}^{n+1}\right|\right|\right|_{S^{n+1}}\left|\left|\left|d_tE^{n+1}\right|\right|\right|_{S^{n+1}} \\
&\le \sum_{n=3}^{\ell-1}\rho_n'\left|\left|\left|\delta^4V^{n+1}\right|\right|\right|_{S^{n+1}}\left|\left|\left|d_tE^{n+1}\right|\right|\right|_{S^{n+1}} \\
&\le \sum_{n=3}^{\ell-1}C\rho_n'k\left\{\sum_{i=0}^{3}\left|\left|\left|d_tE^{n+1-i}\right|\right|\right|_{S^{n+1-i}} + k^3\right\}\left|\left|\left|d_tE^{n+1}\right|\right|\right|_{S^{n+1}} \\
&\le Ck^6 + C\sum_{n=0}^{\ell-1}\left\{\left|\left|\left|d_tE^{n+1}\right|\right|\right|^2 k^2 + \left|\left|\left|\frac{\partial^2}{\partial x\partial y}\delta E^{n+1}\right|\right|\right|^2 k^2\right\},
\end{aligned} \tag{5.17}$$

if we iterate sufficiently often that

$$\rho_n' \leq k. \tag{5.18}$$

We next combine (5.11)–(5.17), use (5.9), and choose ϵ_1 and ϵ_2 sufficiently small to obtain

$$\begin{aligned} &||E^\ell||_1^2 + \left|\left|\frac{\partial^2}{\partial x \partial y}\delta E^\ell\right|\right|^2 k + \sum_{n=3}^{\ell-1} |||d_t E^{n+1}|||^2 k \\ &\quad \leq C\left[h^{2r+2} + k^6 + \sum_{n=3}^{\ell-1}\left\{||E^{n+1}||_1^2 + \left|\left|\frac{\partial^2}{\partial x \partial y}\delta E^{n+1}\right|\right|^2 k\right\} k\right]. \end{aligned} \tag{5.19}$$

The desired result follows from (5.19) by use of a discrete version of Gronwall's lemma and the triangle inequality.

6. Computational considerations. In this section we will consider some rough estimates of the computational complexity of the methods presented here. We will see that the preconditioned iterative methods using an alternating-direction preconditioner allow us to obtain quasi-optimal order work estimates. These methods are thus very efficient computationally.

We will give estimates for $d = 2$. Similar results hold for $d \geq 3$. The process of setting up and factoring L^n requires $O(M^{3/2})$ operations, where $M = \dim M_h$. The solution of (4.2) for $\mu \leq 3$, given the factorization requires $O(M \log M)$ operations. Such bounds have been shown to be minimal. If we conjecture the validity of the above estimates for our problem and refactor L^n and solve (4.2) at each time step, the total amount of work done in the case $\mu \leq 3$ is

$$O\left(N\left\{M^{\frac{3}{2}} + M\log M\right\}\right) = O\left(NM^{\frac{3}{2}}\right), \tag{6.1}$$

where N is the total number of time steps ($N \approx k^{-1}$). Note that the work of factorization dominates the estimate.

Using the preconditioned iterative procedure presented here, only the preconditioner $\overline{L}^0$ must be factored. Since $\overline{L}^0$ can be factored into matrices corresponding to one-dimensional problems, the decomposition of $\overline{L}^0$ requires $O(M_x + M_y)$ operations, which must be done only once. Then the solution process requires only $O(M_x M_y) = O(M)$ operations. The total amount of work done using κ iterations of the iterative procedure at each timestep is then

$$O\left(M_x + M_y + N\kappa M_x M_y\right) = O\left(MN\log N\right), \tag{6.2}$$

since $\kappa = O(\log N)$ for getting (5.18). Since the number of unknowns at each time level is $M = M_x M_y$, the estimate (6.2) is a quasi-optimal order work estimate. Similarly, in three spatial dimensions, the total work is $O(N \log N M_x M_y M_z) = O(MN \log N)$, which is again of quasi-optimal order.

We note that it is computationally wasteful to iterate sufficiently many times to achieve the pessimistic bounds for the norm reductions given in the theorem. Instead, in most applicable iteration methods, it is possible to monitor the norm reduction actually produced at each timestep of the iteration and stop iterating when sufficient norm reduction is achieved. Additional stopping criteria can be imposed in this monitoring process. See [7] for a discussion of stopping criteria for related time-stepping methods.

REFERENCES

[1] O. AXELSSON, *On preconditioning and convergence acceleration in sparse matrix problems*, CERN European Organization for Nuclear Research, Geneva, 1974.

[2] ———, *On the computational complexity of some matrix iterative algorithms*, Report 74.06, Department of Computer Science, Chalmers University of Technology, Göteborg, Sweden, 1974.

[3] J.H. BRAMBLE, *Multistep methods for quasilinear parabolic equations*, Proceedings Second International Conference on Computational Methods in Nonlinear Mechanics, Austin, Texas, March 26-29, 1979.

[4] J.H. BRAMBLE AND P.H. SAMMON, *Efficient higher order multistep methods for parabolic problems: Part I*, to appear.

[5] J.E. DENDY, JR. AND G. FAIRWEATHER, *Alternating-direction Galerkin methods for parabolic and hyperbolic problems on rectangular polygons*, SIAM J. Numer. Anal., 12 (1975), pp. 144-163.

[6] J. DOUGLAS, JR. AND T. DUPONT, *Alternating-direction Galerkin methods on rectangles*, Proceedings Symposium on Numerical Solution of Partial Differential Equations, II, B. Hubbard, ed., Academic Press, New York, 1971, pp. 133-214.

[7] 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., 16 (1979), pp. 503-522.

[8] J. DOUGLAS, JR., T. DUPONT, AND M.F. WHEELER, *A Galerkin procedure for approximating the flux on the boundary for elliptic and parabolic boundary value problems*, RAIRO Anal. Numer., 8 (1974), pp. 47-59.

[9] T. DUPONT, G. FAIRWEATHER, AND J.P. JOHNSON, *Three-level Galerkin methods for parabolic equations*, SIAM J. Numer. Anal., 11 (1974), pp. 392-410.

[10] R.E. EWING, *Efficient multistep procedures for nonlinear parabolic problems with nonlinear Neumann boundary conditions*, Math. Research Center Report 1982, University of Wisconsin, Madison, 1979, and Trans. 25th Conf. Army Math., ARO Report 80-1, and Calcolo 19, (1982), pp. 231-252.

[11] ———, *Alternating-direction Galerkin methods for parabolic, hyperbolic, and Sobolev partial differential equations*, Proceedings Special Year in Numerical Analysis, University of Maryland, College Park, Lecture Notes 20, 1981, pp. 123-150.

[12] L.J. HAYES, *Galerkin alternating-direction methods for nonrectangular regions using patch approximations*, SIAM J. Numer. Anal., 18 (1981), pp. 627-643.