SIAM J. NUMER. ANAL.
Vol. 19, No. 5, October 1982

0036-1429/82/1905-0001 $01.00/0

NUMERICAL METHODS FOR CONVECTION-DOMINATED DIFFUSION PROBLEMS BASED ON COMBINING THE METHOD OF CHARACTERISTICS WITH FINITE ELEMENT OR FINITE DIFFERENCE PROCEDURES*

JIM DOUGLAS, JR.† AND THOMAS F. RUSSELL‡

Abstract. Finite element and finite difference methods are combined with the method of characteristics to treat a parabolic problem of the form $cu_t + bu_x - (au_x)_x = f$. Optimal order error estimates in L^2 and $W^{1,2}$ are derived for the finite element procedure. Various error estimates are presented for a variety of finite difference methods. The estimates show that, for convection-dominated problems ($b \gg a$), these schemes have much smaller time-truncation errors than those of standard methods. Extensions to n-space variables and time-dependent or nonlinear coefficients are indicated, along with applications of the concepts to certain problems described by systems of differential equations.

1. Introduction. In many diffusion processes arising in physical problems, convection essentially dominates diffusion, and it is natural to seek numerical methods for such problems that reflect their almost hyperbolic nature. We shall consider combining the method of characteristics with finite element or finite difference techniques to treat the model problem given by

$$\text{(1.1)}\quad \begin{aligned} &\text{(a)}\quad c(x)\frac{\partial u}{\partial t}+b(x)\frac{\partial u}{\partial x}-\frac{\partial}{\partial x}\left(a(x)\frac{\partial u}{\partial x}\right)=f(x,t), \qquad x\in\mathbb{R},\quad t>0,\\ &\text{(b)}\quad u(x,0)=u_0(x), \qquad x\in\mathbb{R}; \end{aligned}$$

in the last section we shall indicate a number of extensions and applications of our concepts.

Let

$$\psi(x)=[c(x)^2+b(x)^2]^{1/2}, \tag{1.2}$$

and let the characteristic direction associated with the operator cu_t+bu_x be denoted by $\tau=\tau(x)$, where

$$\frac{\partial}{\partial\tau(x)}=\frac{c(x)}{\psi(x)}\frac{\partial}{\partial t}+\frac{b(x)}{\psi(x)}\frac{\partial}{\partial x}. \tag{1.3}$$

Then, equation (1.1a) can be put in the form

$$\psi(x)\frac{\partial u}{\partial\tau(x)}-\frac{\partial}{\partial x}\left(a(x)\frac{\partial u}{\partial x}\right)=f(x,t), \qquad x\in\mathbb{R},\quad t>0. \tag{1.4}$$

In each of the procedures to be treated below we shall consider a time step $\Delta t>0$ and approximate the solution at times $t^n=n\,\Delta t$. The characteristic derivative will be approximated basically in the following manner: Let

$$\bar{x}=x-b(x)\,\Delta t/c(x), \tag{1.5}$$

and note that

$$\psi(x)\frac{\partial u}{\partial\tau}\approx\psi(x)\frac{u(x,t^n)-u(\bar{x},t^{n-1})}{[(x-\bar{x})^2+(\Delta t)^2]^{1/2}}=c(x)\frac{u(x,t^n)-u(\bar{x},t^{n-1})}{\Delta t}. \tag{1.6}$$

* Received by the editors March 9, 1981, and in final revised form January 4, 1982.

† Department of Mathematics, University of Chicago, Chicago, Illinois 60637.
‡ Marathon Oil Company, Denver Research Center, Littleton, Colorado 80160.

Let u_h denote any of the approximate solutions to be defined below. In the case of a Galerkin approximate solution, the interpretation of $u_h(\bar{x}, t^{n-1})$ is simple; it is just the evaluation of u_h at the point $(\bar{x}, t^{n-1})$. However, if u_h is being determined by a finite difference procedure, $u_h(\bar{x}, t^{n-1})$ must be evaluated by an interpolation of the values $u_h(x_i, t^{n-1})$ of the approximate solution at the grid points. Two different interpolations will be treated. If piecewise-linear interpolation is employed, then an error estimate of the form

$$\max_{0\leqq t^n\leqq T} \|(u-u_h)(t^n)\|_{0,\infty} = O(\Delta t+\min(h, h^2/\Delta t)) \tag{1.7}$$

will be derived. This scheme will be the appropriate choice when convection is distinctly more important than diffusion. If a local quadratic interpolation is used and the time step is constrained by the asymptotic relation $\Delta t = O(h^2)$, then the error will be shown to satisfy the inequality

$$\max_{0\leqq t^n\leqq T} \|(u-u_h)(t^n)\|_{0,2} = O(h^2+\Delta t). \tag{1.8}$$

If convection and diffusion are roughly equivalent in importance, there would be advantages to this second method. (In the above and in what follows, $\|\cdot\|_{k,p}$ is the norm in the Sobolev space $W^{k,p}(\mathbb{R})$; usually the subscript $p=2$ will be suppressed.)

The principal gains from these procedures appear in time truncation. Approximation of $\partial u/\partial t$ by standard backward differencing leads to errors of the form $K\|\partial^2 u/\partial t^2\|\,\Delta t$ in suitable norms, while our methods will be shown to yield $K\|\partial^2 u/\partial \tau^2\|\,\Delta t$. In problems with significant convection, the solution changes much less rapidly in the characteristic τ direction than in the t direction. Thus, our schemes will permit the use of larger time steps, with corresponding improvements in efficiency, at no cost in accuracy. We shall see that there is no stability limitation on the size of Δt.

An outline of the remainder of the paper is as follows. In § 2 a Galerkin method will be formulated and analyzed. Optimal order error estimates in $L^2(\mathbb{R})$ and $W^{1,2}(\mathbb{R})$ will be obtained. A critical part of the argument is the proof that the norm of the difference $\eta(x)-\eta(\bar{x})$ in the space $W^{-1,2}(\mathbb{R})$ is bounded by a constant multiple of $\Delta t\|\eta\|_0$. In § 3 the finite difference procedure using linear interpolation is set up and analyzed, and the method using quadratic interpolation is treated in § 4. Finally, several extensions and applications are indicated in § 5. Throughout, the symbols K and ε will denote, respectively, a generic constant and a generic small positive constant.

2. A Galerkin method. Throughout this section we assume that the coefficients $b(x)$ and $c(x)$ are bounded, that c is bounded below by a positive constant, and that

$$\left|\frac{b(x)}{c(x)}\right| + \left|\frac{d}{dx}\left(\frac{b(x)}{c(x)}\right)\right| \leqq K. \tag{2.1}$$

In particular, this implies that $|x-\bar{x}|\leqq K\,\Delta t$. We also assume that the solution u of (1.1) satisfies

$$\begin{aligned}
&\text{(a)}\quad u\in L^\infty(0, T; W^{q,2}(\mathbb{R})),\\
(2.2)\quad &\text{(b)}\quad \frac{\partial u}{\partial t}\in L^2(0, T; W^{q-1+\theta,2}(\mathbb{R})), \qquad \theta=1 \quad \text{if } q=2 \quad \text{and} \quad \theta=0 \quad \text{if } q>2,\\
&\text{(c)}\quad \frac{\partial^2 u}{\partial t^2}\in L^2(0, T; L^2(\mathbb{R})),
\end{aligned}$$

for some $q \geqq 2$. Let $A: W^{1,2}(\mathbb{R}) \times W^{1,2}(\mathbb{R}) \to \mathbb{R}$ be the bilinear form

$$A(w, z) = (aw', z'),$$

where the pairing is the inner product on $L^2(\mathbb{R})$. Then, (1.1) can be written in the equivalent form (at least, for reasonable $a(x)$ such that $a(x) \geqq a_0 > 0$)

$$(2.3)\quad \begin{aligned} &\text{(a)} \quad \left(\psi \frac{\partial u}{\partial \tau}, v\right) + A(u, v) = (f, v), \qquad v \in W^{1,2}(\mathbb{R}), \quad t > 0, \\ &\text{(b)} \quad A(u(0) - u_0, v) = 0, \qquad v \in W^{1,2}(\mathbb{R}), \end{aligned}$$

and it is this form that will be discretized below.

Let $\mathcal{M}_h$ be a subspace of $W^{1,2}(\mathbb{R}) \cap W^{1,\infty}(\mathbb{R})$ such that, for $v \in W^{s,2}(\mathbb{R})$ with $1 \leqq s \leqq q$,

$$(2.4)\quad \inf_{\chi \in \mathcal{M}_h} [\|v - \chi\|_0 + h\|v - \chi\|_1] \leqq K\|v\|_s h^s.$$

Note that $\mathcal{M}_h$ is necessarily infinite-dimensional, since we are working on the whole line; practically, we can assume that the support of u_0 is compact, that the portion of the line on which we need to know u is bounded, and that u is very small outside that set. Then, $\mathcal{M}_h$ can be taken to be finite-dimensional.

Our Galerkin procedure is the determination of the map $u_h: \{t^0, t^1, \cdots\} \to \mathcal{M}_h$ satisfying the relations

$$(2.5)\quad \begin{aligned} &\text{(a)} \quad \left(c \frac{u_h^n - \bar{u}_h^{n-1}}{\Delta t}, v\right) + A(u_h^n, v) = (f^n, v), \qquad v \in \mathcal{M}_h, \quad n \geqq 1, \\ &\text{(b)} \quad A(u_h^0 - u_0, v) = 0, \qquad v \in \mathcal{M}_h, \end{aligned}$$

where

$$(2.6)\quad u_h^n = u_h(t^n), \qquad \bar{u}_h^{n-1}(x) = u_h^{n-1}(\bar{x}) = u_h^{n-1}(x - b(x)\,\Delta t/c(x)).$$

It is obvious that (2.5) determines $\{u_h^n\}$ uniquely in terms of the data u_0 and f. Thus, we can turn to the analysis of the convergence of the method, and it is convenient [6] to introduce the elliptic projection of the solution.

Let $w_h: [0, T] \to \mathcal{M}_h$ be given by the relations

$$(2.7)\quad A(u - w_h, v) = 0, \qquad v \in \mathcal{M}_h, \quad 0 \leqq t \leqq T.$$

Let

$$(2.8)\quad \eta = u - w_h, \quad \xi = u_h - w_h, \quad \zeta = u - u_h.$$

It is well known that, for $p = 2$ or ∞ and $1 \leqq s \leqq q$,

$$(2.9)\quad \|\eta\|_{L^p(0,T;L^2(\mathbb{R}))} + h\|\eta\|_{L^p(0,T;W^{1,2}(\mathbb{R}))} \leqq K\|u\|_{L^p(0,T;W^{s,2}(\mathbb{R}))} h^s.$$

Since the form $A(\cdot, \cdot)$ is independent of time, it also follows easily that, for $q \geqq 3$ and $1 \leqq s \leqq q$,

$$(2.10)\quad \left\|\frac{\partial \eta}{\partial t}\right\|_{L^2(0,T;W^{-1,2}(\mathbb{R}))} + h\left\|\frac{\partial \eta}{\partial t}\right\|_{L^2(0,T;L^2(\mathbb{R}))} \leqq K\left\|\frac{\partial u}{\partial t}\right\|_{L^2(0,T;W^{s-1,2}(\mathbb{R}))} h^s;$$

in the case $q = 2$ there is no gain of a factor h in the $W^{-1,2}$ estimate.

A calculation shows that, for $v \in \mathcal{M}_h$,

$$\left(c\frac{\xi^n-\bar{\xi}^{n-1}}{\Delta t}, v\right)+A(\xi^n, v)=\left(\psi\frac{\partial u^n}{\partial \tau}-c\frac{u^n-\bar{u}^{n-1}}{\Delta t}, v\right)+\left(c\frac{\eta^n-\bar{\eta}^{n-1}}{\Delta t}, v\right). \tag{2.11}$$

In order to derive an $L^2(\mathbb{R})$ error estimate, choose the test function $v=\xi^n$. For the first term on the right-hand side of (2.11), we shall bound

$$\left\|\psi\frac{\partial u^n}{\partial \tau}-c\frac{u^n-\bar{u}^{n-1}}{\Delta t}\right\|$$

through the representation below involving an integral in the parameter τ along the tangent to the characteristic from $(\bar{x}, t^{n-1})$ to (x, t^n). Denote the coordinates of the point on the segment by $(x(\tau), t(\tau))$. The standard backward difference quotient error equation is given by

$$\frac{\partial u^n}{\partial t}-\frac{u^n-u^{n-1}}{\Delta t}=\frac{1}{\Delta t}\int_{t^{n-1}}^{t^n}(t-t^{n-1})\frac{\partial^2 u}{\partial t^2}\,dt;$$

analogously, along the tangent to the characteristic

$$\psi\frac{\partial u^n}{\partial \tau}-c\frac{u^n-\bar{u}^{n-1}}{\Delta t}=\frac{c}{\Delta t}\int_{(\bar{x},t^{n-1})}^{(x,t^n)}\sqrt{(x(\tau)-\bar{x})^2+(t(\tau)-t^{n-1})^2}\,\frac{\partial^2 u}{\partial \tau^2}\,d\tau.$$

Taking the $L^2(\mathbb{R})$-norm of this error term we obtain

$$\begin{aligned}\left\|\psi\frac{\partial u^n}{\partial \tau}-c\frac{u^n-\bar{u}^{n-1}}{\Delta t}\right\|^2 &\leqq \int_R\left(\frac{c}{\Delta t}\right)^2\left(\frac{\psi}{c}\Delta t\right)^2\left|\int_{(\bar{x},t^{n-1})}^{(x,t^n)}\frac{\partial^2 u}{\partial \tau^2}\,d\tau\right|^2 dx\\ &\leqq \Delta t\left\|\frac{\psi^3}{c}\right\|_{0,\infty}\int_{\mathbb{R}}\int_{(\bar{x},t^{n-1})}^{(x,t^n)}\left|\frac{\partial^2 u}{\partial \tau^2}\right|^2 d\tau\,dx\\ &\leqq \Delta t\left\|\frac{\psi^4}{c^2}\right\|_{0,\infty}\int_R\int_{t^{n-1}}^{t^n}\left|\frac{\partial^2 u}{\partial \tau^2}\left(\frac{t^n-t}{\Delta t}\bar{x}+\frac{t-t^{n-1}}{\Delta t}x, t\right)\right|^2 dt\,dx.\end{aligned}$$

To relate this to a standard norm of $\partial^2 u/\partial \tau^2$, consider the transformation

$$S:(x, t)\mapsto(z, t)=\left(\frac{t^n-t}{\Delta t}\bar{x}+\frac{t-t^{n-1}}{\Delta t}x, t\right)=(\theta(t)\bar{x}+(1-\theta(t))x, t).$$

The Jacobian of this map is given by

$$DS=\begin{pmatrix}1-\theta\,\Delta t\left(\frac{b}{c}\right)'(x) & \frac{b(x)}{c(x)}\\ 0 & 1\end{pmatrix};$$

by (2.1), S is invertible for sufficiently small Δt and its determinant is $1+O(\Delta t)$. For any fixed t, S obviously maps $\mathbb{R}\times\{t\}$ onto itself, so the same is true for $\mathbb{R}\times[t^{n-1}, t^n]$. It follows that

$$\left\|\psi\frac{\partial u^n}{\partial \tau}-c\frac{u^n-\bar{u}^{n-1}}{\Delta t}\right\|^2\leqq 2\left\|\frac{\psi^4}{c^2}\right\|_{0,\infty}\left\|\frac{\partial^2 u}{\partial \tau^2}\right\|^2_{L^2(\mathbb{R}\times[t^{n-1},t^n])}\Delta t,$$

and the first term on the right-hand side of (2.11) is bounded by

$$K\left\|\frac{\partial^2 u}{\partial \tau^2}\right\|^2_{L^2(\mathbb{R}\times[t^{n-1},t^n])}\Delta t+\|\xi^n\|^2. \tag{2.12}$$

In (2.12), for $(x, t) \in \mathbb{R} \times [t^{n-1}, t^n]$, $\tau(x, t)$ is not necessarily the same direction as $\tau(x)$; it is $\tau(x')$, where x' is the point such that the characteristic segment from (x', t^n) passes through (x, t). This difference could be reduced by a modification of the method, in which the position $\bar{x}$ would be determined by a characteristic polygon corresponding to a partition of the time interval $[t^{n-1}, t^n]$.

Write $\eta^n - \bar{\eta}^{n-1}$ as the sum $(\eta^n - \eta^{n-1}) + (\eta^{n-1} - \bar{\eta}^{n-1})$. Then,

$$\left(c\frac{\eta^n - \eta^{n-1}}{\Delta t}, \xi^n\right) \leqq K\frac{1}{\Delta t}\int_{t^{n-1}}^{t^n} \left\|\frac{\partial \eta}{\partial t}\right\|_{-1} d\alpha \cdot \|\xi^n\|_1 \leqq \varepsilon\|\xi^n\|_1^2 + \frac{K}{\Delta t}\left\|\frac{\partial \eta}{\partial t}\right\|^2_{L^2(t^{n-1}, t^n; W^{-1,2})}. \tag{2.13}$$

Finally, we consider

$$\left(c\frac{\eta^{n-1} - \bar{\eta}^{n-1}}{\Delta t}, \xi^n\right) \leqq K\|\xi^n\|_1 \left\|\frac{\eta^{n-1} - \bar{\eta}^{n-1}}{\Delta t}\right\|_{-1}. \tag{2.14}$$

LEMMA 1. *If $\eta \in L^2(\mathbb{R})$ and $\bar{\eta}(x) = \eta(x - g(x)\,\Delta t)$, where g and g' are bounded, then*

$$\|\eta - \bar{\eta}\|_{-1} \leqq K\|\eta\|\,\Delta t.$$

Proof. Let $z = F(x) = x - g(x)\,\Delta t$. Then F is invertible for Δt sufficiently small, and F' and $(F^{-1})'$ are both of the form $1 + O(\Delta t)$. Hence,

$$\begin{aligned}
\|\eta - \bar{\eta}\|_{-1} &= \sup_{\varphi \in W^{1,2}} \left(\|\varphi\|_1^{-1} \int_{\mathbb{R}} [\eta(x) - \eta(x - g(x)\,\Delta t)]\varphi(x)\,dx\right) \\
&= \sup_{\varphi \in W^{1,2}} \left(\|\varphi\|_1^{-1} \left[\int_{\mathbb{R}} \eta(x)\varphi(x)\,dx - \int_{\mathbb{R}} \eta(z)\varphi(F^{-1}(z))(1 + O(\Delta t))\,dz\right]\right) \\
&\leqq \sup_{\varphi \in W^{1,2}} \left(\|\varphi\|_1^{-1} \int_{\mathbb{R}} \eta(x)[\varphi(x) - \varphi(F^{-1}(x))]\,dx\right) \\
&\quad + K\,\Delta t \sup_{\varphi \in W^{1,2}} \left(\|\varphi\|^{-1} \int_{\mathbb{R}} \eta(x)\varphi(F^{-1}(x))\,dx\right).
\end{aligned} \tag{2.15}$$

Let $G(x) = x - F^{-1}(x)$; then $|G(x)| \leqq K\,\Delta t$, and

$$\begin{aligned}
\|\varphi(x) - \varphi(F^{-1}(x))\|^2 &\leqq \int_{\mathbb{R}} \left(\int_{F^{-1}(x)}^{x} \left|\frac{d\varphi}{dx}\right|\right)^2 dx \\
&\leqq K(\Delta t)^2 \int_{\mathbb{R}} \int_0^1 \left|\frac{d\varphi}{dx}(x - G(x)y)\right|^2 dy\,dx \leqq K(\Delta t)^2\|\varphi\|_1^2,
\end{aligned} \tag{2.16}$$

where the last step uses the change of variable $\tilde{x} = x - G(x)y$, which induces a factor of $1 + O(\Delta t)$. A similar change of variable demonstrates that

$$\|\varphi \circ F^{-1}\|^2 = \|\varphi\|^2(1 + \gamma K\,\Delta t), \qquad |\gamma| \leqq 1, \tag{2.17}$$

where K is the constant of (2.1); the same is true for $\varphi \circ F$. Combining (2.15)–(2.17), we obtain the lemma. Note that the argument is valid for $x \in \mathbb{R}^k$, $k \geqq 1$.

An application of Lemma 1 to (2.14) shows that

$$\left(c\frac{\eta^{n-1} - \bar{\eta}^{n-1}}{\Delta t}, \xi^n\right) \leqq \varepsilon\|\xi^n\|_1^2 + K\|\eta^{n-1}\|^2, \tag{2.18}$$

and this completes the treatment of the right-hand side.

The left-hand side is boundable below by

$$\left(c\frac{\xi^n-\bar{\xi}^{n-1}}{\Delta t},\xi^n\right)+A(\xi^n,\xi^n)$$

$$\geqq\frac{1}{2\Delta t}[(c\xi^n,\xi^n)-(c\bar{\xi}^{n-1},\bar{\xi}^{n-1})]+A(\xi^n,\xi^n) \tag{2.19}$$

$$=\frac{1}{2\Delta t}[(c\xi^n,\xi^n)-(c\xi^{n-1},\xi^{n-1})(1+\gamma^nK\,\Delta t)]+A(\xi^n,\xi^n),\qquad |\gamma^n|\leqq 1,$$

where (2.17) has been used. The inequalities (2.12), (2.13), (2.18), and (2.19) can be combined with (2.11) to give the recursion relation

$$\frac{1}{2\,\Delta t}[(c\xi^n,\xi^n)-(c\xi^{n-1},\xi^{n-1})]+\frac{a_0}{2}\|\xi^n\|_1^2$$

$$\leqq K\|\xi^n\|^2+K\,\Delta t\left\|\frac{\partial^2u}{\partial\tau^2}\right\|^2_{L^2(t^{n-1},t^n;L^2)}+\frac{K}{\Delta t}\left\|\frac{\partial\eta}{\partial t}\right\|^2_{L^2(t^{n-1},t^n;W^{-1,2})}+K\|\eta^{n-1}\|^2+K\|\xi^{n-1}\|^2. \tag{2.20}$$

If (2.20) is multiplied by $2\,\Delta t$ and summed in time and if it is noted that (2.5b) and (2.7) imply that $\xi^0=0$, then it follows that

$$\max_{0\leqq t^n\leqq T}\|\xi^n\|+\left(\sum_{n=1}^{T/\Delta t}\|\xi^n\|_1^2\,\Delta t\right)^{1/2}$$

$$\leqq K\left\|\frac{\partial^2u}{\partial\tau^2}\right\|_{L^2(0,T;L^2)}\Delta t+K\left\|\frac{\partial\eta}{\partial t}\right\|_{L^2(0,T;W^{-1,2})}+K\|\eta\|_{L^\infty(0,T;L^2)}. \tag{2.21}$$

Recall that the error $\zeta=u-u_h=\eta-\xi$; consequently, the inequalities (2.9), (2.10), and (2.21) together imply the following theorem:

THEOREM 2. *Let the solution u of* (1.1) *satisfy the requirement* (2.2) *for some* $q\geqq 2$, *and let the Galerkin solution* u_h *be defined by* (2.5). *Then,*

$$\max_{0\leqq t^n\leqq T}\|(u-u_h)(t^n)\|\leqq K\left[\|u\|_{L^\infty(0,T;W^{q,2}(\mathbb{R}))}+\left\|\frac{\partial u}{\partial t}\right\|_{L^2(0,T;W^{q-1+\theta,2}(\mathbb{R}))}\right]h^q$$

$$+K\left\|\frac{\partial^2u}{\partial\tau^2}\right\|_{L^2(0,T;L^2(\mathbb{R}))}\Delta t, \tag{2.22}$$

where $\theta=1$ *for* $q=2$ *and* $\theta=0$ *for* $q\geqq 3$.

An optimal order estimate for $u-u_h$ in $W^{1,2}(\mathbb{R})$ can be derived in a similar fashion, starting from the test function $v=(\xi^n-\xi^{n-1})/\Delta t$ in (2.11). The left-hand side satisfies the inequality

$$\left(c\frac{\xi^n-\bar{\xi}^{n-1}}{\Delta t},\frac{\xi^n-\xi^{n-1}}{\Delta t}\right)+A\left(\xi^n,\frac{\xi^n-\xi^{n-1}}{\Delta t}\right)$$

$$\geqq\left(c\frac{\xi^n-\xi^{n-1}}{\Delta t},\frac{\xi^n-\xi^{n-1}}{\Delta t}\right)+\frac{1}{2\,\Delta t}[A(\xi^n,\xi^n)-A(\xi^{n-1},\xi^{n-1})] \tag{2.23}$$

$$-K\|\xi^{n-1}\|_1^2-\varepsilon\left\|\frac{\xi^n-\xi^{n-1}}{\Delta t}\right\|^2,$$

since

$$\left|\left(c\frac{\xi^{n-1}-\bar{\xi}^{n-1}}{\Delta t},\frac{\xi^n-\xi^{n-1}}{\Delta t}\right)\right|\leqq K\left\|\frac{\partial \xi^{n-1}}{\partial x}\right\|\left\|\frac{\xi^n-\xi^{n-1}}{\Delta t}\right\|. \tag{2.24}$$

The right-hand side terms can be estimated as below. First,

$$\left(\psi\frac{\partial u^n}{\partial \tau}-c\frac{u^n-\bar{u}^{n-1}}{\Delta t},\frac{\xi^n-\xi^{n-1}}{\Delta t}\right)\leqq K\left\|\frac{\partial^2 u}{\partial \tau^2}\right\|^2_{L^2(t^{n-1},t^n;L^2)}\Delta t+\varepsilon\left\|\frac{\xi^n-\xi^{n-1}}{\Delta t}\right\|^2. \tag{2.25}$$

Second,

$$\left(c\frac{\eta^n-\bar{\eta}^{n-1}}{\Delta t},\frac{\xi^n-\xi^{n-1}}{\Delta t}\right)\leqq \frac{K}{\Delta t}\left\|\frac{\partial \eta}{\partial t}\right\|^2_{L^2(t^{n-1},t^n:L^2)}+K\|\eta^{n-1}\|_1^2+\varepsilon\left\|\frac{\xi^n-\xi^{n-1}}{\Delta t}\right\|^2. \tag{2.26}$$

Note that no use of Lemma 1 is needed for this estimate. The above inequalities lead to the bound

$$\max_{0\leqq t^n\leqq T}\|\xi^n\|_1\leqq K\left\|\frac{\partial^2 u}{\partial \tau^2}\right\|_{L^2(0,T;L^2)}\Delta t+K\left[\|\eta\|_{L^\infty(0,T;W^{1,2})}+\left\|\frac{\partial \eta}{\partial t}\right\|_{L^2(0,T;L^2)}\right], \tag{2.27}$$

from which it follows that

(2.28)

$$\max_{0\leqq t^n\leqq T}\|(u-u_h)(t^n)\|_1\leqq K\left[\|u\|_{L^\infty(0,T;W^{q,2})}+\left\|\frac{\partial u}{\partial t}\right\|_{L^2(0,T;W^{q-1,2})}\right]h^{q-1}+K\left\|\frac{\partial^2 u}{\partial \tau^2}\right\|_{L^2(0,T;L^2)}\Delta t.$$

3. Some finite difference methods based on linear interpolation. Let $x_i=ih$, $t^n=n\,\Delta t$, and $z(x_i,t^n)=z_i^n$. Let $\bar{x}_i=x_i-b_i\,\Delta t/c_i$. Denote the centered, weighted second difference in x of a grid function z by

$$\delta_{\bar{x}}(a\delta_x z)_i=h^{-2}[a_{i+1/2}(z_{i+1}-z_i)-a_{i-1/2}(z_i-z_{i-1})],$$

where $a_{i+1/2}=a(x_{i+1/2})=a((i+1/2)h)$. Let w_i^n denote the grid values of the appropriate solution. Let $\bar{w}_i^{n-1}=w^{n-1}(\bar{x}_i)$, where $w^{n-1}(x)$ is an extension of the grid values to a function; for the moment leave the definition of w^{n-1} unspecified. Continue letting $\psi(x)=(b(x)^2+c(x)^2)^{1/2}$. Then, the basic finite difference method is given by the equations

$$\begin{aligned}&\text{(a)}\quad c_i\frac{w_i^n-\bar{w}_i^{n-1}}{\Delta t}-\delta_{\bar{x}}(a\delta_x w^n)_i=f_i^n, \qquad -\infty<i<\infty,\quad n\geqq 1,\\ &\text{(b)}\quad w_i^0=u_0(x_i), \qquad -\infty<i<\infty.\end{aligned} \tag{3.1}$$

Again it is clear that there exists a unique solution of (3.1) as soon as an evaluation is specified for $w^{n-1}(x)$. A convergence argument similar to that given in the previous section can be constructed for the finite difference procedure and such an argument will be employed in the next section; here, we shall use a maximum principle argument.

Note first that, with $\bar{u}^{n-1}(x)=u(\bar{x},t^{n-1})$,

$$c_i\frac{u_i^n-\bar{u}_i^{n-1}}{\Delta t}-\delta_{\bar{x}}(a\delta_x u^n)_i=f_i^n+e_i^n, \tag{3.2}$$

where a calculation shows that

$$e_i^n=\frac{c_i^2+b_i^2}{2c_i}\frac{\partial^2 u^*}{\partial \tau^2}\Delta t+O(\|u^n\|_{3,\infty}h)\quad\text{or}\quad O(\|u^n\|_{4,\infty}h^2) \tag{3.3}$$

and $\partial^2 u^*/\partial\tau^2$ is some evaluation of the second tangential derivative along the characteristic segment between (x_i, t^n) and $(\bar{x}_i, t^{n-1})$. In the convection-dominated case this second derivative is relatively much smaller than $\partial^2 u/\partial x^2$ or $\partial^2 u/\partial t^2$. Thus, if $\zeta = u - w$,

$$(3.4)\qquad \begin{aligned} &\text{(a)}\quad c_i\frac{\zeta_i^n - \bar{\zeta}_i^{n-1}}{\Delta t} - \delta_{\bar{x}}(a\delta_x\zeta^n)_i = e_i^n, \qquad -\infty < i < \infty, \quad n \geqq 1,\\ &\text{(b)}\quad \zeta_i^0 = 0, \qquad -\infty < i < \infty; \end{aligned}$$

however, it is not necessarily true that $\zeta^0(x) \equiv 0$, because of the possibility of using some form of interpolation to define $w^0(x)$. It follows trivially from a maximum principle argument that

$$(3.5)\qquad \max_i |\zeta_i^n| \leqq \max_i |\bar{\zeta}_i^{n-1}| + K \max_i |e_i^n|\,\Delta t.$$

To this point the argument has been independent of the definition of $w^{n-1}(x)$; however, in order to relate $\max|\bar{\zeta}_i^{n-1}|$ to $\max|\zeta_i^{n-1}|$, it is necessary to select the evaluation rule. The choice to be made in this section is piecewise-linear interpolation, the operator for which we shall denote by I_1. Then,

$$(3.6)\qquad \max_i |\bar{\zeta}_i^{n-1}| \leqq \max_i |\zeta_i^{n-1}| + \max_i |u(\bar{x}_i, t^{n-1}) - I_1 u(\cdot, t^{n-1})(\bar{x}_i)|.$$

Set

$$(3.7)\qquad \begin{aligned} &\text{(a)}\qquad j(i) = \{j : |\bar{x}_i - x_j| = \min_k |\bar{x}_i - x_k|\},\\ &\text{(b)}\qquad h_i^* = |\bar{x}_i - x_{j(i)}| \leqq \min(h/2, K\,\Delta t). \end{aligned}$$

Then it is an easy consequence of the Peano kernel theorem that

$$(3.8)\qquad |u(\bar{x}_i, t^{n-1}) - (I_1 u(\cdot, t^{n-1}))(\bar{x}_i)| \leqq K\|u^{n-1}\|_{2,\infty} h_i^* h.$$

Thus,

$$(3.9)\qquad \max_i |\zeta_i^n| \leqq \max_i |\zeta_i^{n-1}| + K\left[\max_i |e_i^n| + \|u^{n-1}\|_{2,\infty} \min\left(h, \frac{h^2}{\Delta t}\right)\right]\Delta t.$$

If $w^0(x)$ is given by $I_1\{w_i^0\} = I_1\{u_0(x_i)\}$, then

$$(3.10)\qquad |\bar{\zeta}_i^0| \leqq K\|u_0\|_{2,\infty} \min(h^2, h\,\Delta t).$$

Thus, it follows from (3.9) and (3.10) that

$$(3.11)\qquad \begin{aligned} \max_i |\zeta_i^n| &\leqq K[\max_{j,m} |e_j^m| + \|u\|_{L^\infty(0,T;W^{2,\infty})} \min(h, h^2/\Delta t)]\\ &\leqq K\left[\sup_{\mathbb{R}\times[0,T]} \left|\frac{\partial^2 u}{\partial\tau(x)^2}\right| \Delta t + \|u\|_{L^\infty(0,T;W^{3,\infty})} h\right] \end{aligned}$$

as

$$(3.12)\qquad \|u\|_{L^\infty(0,T;W^{3,\infty})} h \geqq \|u\|_{L^\infty(0,T;W^{2,\infty})} \min(h, h^2/\Delta t).$$

The result above can be described in the following theorem.

THEOREM 3. *Let the solution of* (1.1) *belong to* $W^{2,\infty}(\mathbb{R}\times[0,T]) \cap L^\infty(0,T; W^{3,\infty}(\mathbb{R}))$, *and let the approximate solution* w_i^n *be defined by the finite difference method* (3.1), *where* $\bar{w}_i^{n-1} = w^{n-1}(\bar{x}_i) = w^{n-1}(x_i - b_i\,\Delta t/c_i) = (I_1\{w_j^{n-1}\})(\bar{x}_i)$, *the piecewise-linear interpolant of* $\{w_j^{n-1}\}$. *Then the error* $\zeta = u - w$ *satisfies the bound* (3.11).

Later, an extension of (3.1) to several space variables will be discussed. In the case of a single space variable, it is easy to generalize the method to time-dependent, variable-spacing grids, and we shall indicate this flexibility here. Let $t^n = n\,\Delta t$ as above and let

$$\delta^n = \{\cdots, x_{-2}^n, x_{-1}^n, x_0^n, x_1^n, x_2^n, \cdots\},\ x_i^n - x_{i-1}^n = h_i^n > 0, \quad x_j^n \to \pm\infty \quad \text{as } j \to \pm\infty, \tag{3.13}$$

be the grid to be employed at time t^n. Let the approximation to $(au_x)_x$ be given by $(u_i^n = u(x_i^n, t^n),\ a_{i+1/2}^n = a((x_i^n + x_{i+1}^n)/2))$,

$$\begin{aligned}\delta_{\bar{x}}(a\delta_x u)_i^n &= \frac{2}{h_i^n + h_{i+1}^n}\left[a_{i+1/2}^n \frac{u_{i+1}^n - u_i^n}{h_{i+1}^n} - a_{i-1/2}^n \frac{u_i^n - u_{i-1}^n}{h_i^n}\right] \\ &= (au_x)_x(x_i^n, t^n) + O(\|u^n\|_{W^{3,\infty}([x_{i-1}^n, x_{i+1}^n])}(h_i^n + h_{i+1}^n)),\end{aligned} \tag{3.14}$$

with the constant depending on $a(x)$ but not on any properties of δ^n. Again interpret w^{n-1} as the piecewise-linear interpolant of $\{w_j^{n-1}\}$, now over the grid δ^{n-1}. Let

$$\bar{x}_i^n = x_i^n - b(x_i^n)\,\Delta t / c(x_i^n). \tag{3.15}$$

Then the natural generalization of (3.1) is the equation

$$\begin{aligned}&\text{(a)} \qquad c(x_i^n)\frac{w_i^n - w^{n-1}(\bar{x}_i^n)}{\Delta t} - \delta_{\bar{x}}(a\delta_x w)_i^n = f(x_i^n, t^n), \qquad n \geqq 1 \quad \forall i, \\ &\text{(b)} \qquad w_i^0 = u_0(x_i^0) \quad \forall i.\end{aligned} \tag{3.16}$$

Then, if $\bar{x}_i^n \in [x_{j-1}^{n-1}, x_j^{n-1}] = J_i^n$,

$$|u(\bar{x}_i^n, t^{n-1}) - (I_1 u(\cdot, t^{n-1}))(\bar{x}_i^n)| \leqq K\|u^{n-1}\|_{W^{2,\infty}(J_i^n)} \min(|J_i^n|^2, |J_i^n|\,\Delta t), \tag{3.17}$$

where $|J_i^n|$ is the length of J_i^n. Since

$$\max_i |e_i^n| \leqq K\left[\sup_{\mathbb{R}\times[t^{n-1}, t^n]} \left|\frac{\partial^2 u}{\partial\tau(x)^2}\right| \Delta t + \sup_i \|u^n\|_{W^{3,\infty}([x_{i-1}^n, x_{i+1}^n])}(h_i^n + h_{i+1}^n)\right], \tag{3.18}$$

the somewhat awkward, but reasonably precise, inequality

$$\begin{aligned}\max_i |\zeta_i^n| \leqq \max_i |\zeta_i^{n-1}| + K\Bigg[&\left\|\frac{\partial^2 u}{\partial\tau^2}\right\|_{L^\infty(R\times[t^{n-1}, t^n])} \Delta t \\ &+ \sup_i \|u^{n-1}\|_{W^{2,\infty}(J_i^n)} \min(|J_i^n|, |J_i^n|^2/\Delta t) \\ &+ \sup_i \|u^n\|_{W^{3,\infty}([x_{i-1}^n, x_{i+1}^n])}(h_i^n + h_{i+1}^n)\Bigg]\Delta t\end{aligned} \tag{3.19}$$

can be derived in exactly the same manner as (3.9); consequently,

$$\begin{aligned}\max_{0\leqq t^n \leqq T} \max_i |\zeta_i^n| \leqq K\Bigg[&\left\|\frac{\partial^2 u}{\partial\tau^2}\right\|_{L^\infty(R\times[0,T])} \Delta t \\ &+ \sum_{n=0}^{T/\Delta t - 1} \sup_i \|u^{n-1}\|_{W^{2,\infty}(J_i^n)} \min(|J_i^n|, |J_i^n|^2/\Delta t)\cdot\Delta t \\ &+ \sum_{n=1}^{T/\Delta t} \sup_i \|u^n\|_{W^{3,\infty}([x_{i-1}^n, x_{i+1}^n])}(h_i^n + h_{i+1}^n)\,\Delta t\Bigg].\end{aligned} \tag{3.20}$$

It should be emphasized that no relation was assumed between the grids at successive time steps in the derivation of the error bound (3.20). It is clear that a reasonably nearly optimal error is achieved by selecting δ^n to have close spacing where the solution has large higher derivatives in space and sparse spacing where it is nearly linear in x. Practically, a selection of δ^n, given δ^{n-1} and $\{w_i^{n-1}\}$, can be based on the following concepts:

i) Let the initial selection of δ^n be given by

$$x_i^n = x_i^{n-1} + b(x_i^{n-1})\,\Delta t/c(x_i^{n-1}); \tag{3.21}$$

reorder the grid points where necessary. Also, maintain some small minimum spacing to avoid ridiculous work.

ii) Compute $\{w_i^n\}$ using the current selection of δ^n. Then, use finite differences to approximate the second and third derivatives of the solution, and use these approximations to check that the second and third error terms in (3.19) are sufficiently small. If not, refine the grid where required. Also, coarsen the grid where it can be done. If refinement was required, repeat this step.

At points where the rule (3.21) is retained, the error term of the form involving $|J_i^n|$ and $|J_i^n|^2/\Delta t$ can be modified. Note that $\bar{x}_i^n$ is almost equal to x_i^{n-1}. In fact,

$$x_i^{n-1} - \bar{x}_i^n = \left(\frac{b(x_i^n)}{c(x_i^n)} - \frac{b(x_i^{n-1})}{c(x_i^{n-1})}\right)\Delta t = O((\Delta t)^2), \tag{3.22}$$

by (2.1). Thus, (3.17) can be improved to read

$$|u(\bar{x}_i^n, t^{n-1}) - I_1 u(t^{n-1})\bar{x}_i^n)| \leqq K\|u^{n-1}\|_{W^{2,\infty}(J_i^n)}|J_i^n|(\Delta t)^2, \tag{3.23}$$

with the consequence that the second term on the right-hand side of (3.19) can be replaced by the expression

$$\begin{aligned} &\sup\{\|u^{n-1}\|_{W^{2,\infty}(J_i^n)}|J_i^n|\,\Delta t\colon i \text{ such that } x_i^n = x_i^{n-1} + b(x_i^{n-1})\,\Delta t/c(x_i^{n-1})\} \\ &+\sup\{\|u^{n-1}\|_{W^{2,\infty}(J_i^n)}\min(|J_i^n|, |J_i^n|^2/\Delta t)\colon x_i^n \neq x_i^{n-1} + b(x_i^{n-1})\,\Delta t/c(x_i^{n-1})\}. \end{aligned} \tag{3.24}$$

This observation is important in that (3.24) measures the error in the interpolation arising in the transport term, while the third term on the right-hand side of (3.19) measures the error in the diffusion term and will have a smaller multiplicative constant than that for (3.24) when transport seriously dominates diffusion. Effectively, an extra factor of Δt has been introduced into a large portion of the error in transport.

The initial-boundary value problem can easily be treated by a simple variant of the method. Consider the case of the half-line $x > 0$, and assume that u is specified for $x = 0$:

$$u(0, t) = g(t), \qquad t \geqq 0. \tag{3.25}$$

Let $x_0^n = 0$ and set $w_0^n = g(t^n)$, $n \geqq 0$. If $\bar{x}_i^n \geqq 0$ for $i \geqq 1$, then (3.1a) suffices. If for some $i \geqq 1$, $\bar{x}_i^n < 0$, then the replacement of ψu_τ must be modified for such an i. First, there exists k_i^n, $0 < k_i^n < \Delta t$, such that

$$x_i^n - b(x_i^n)k_i^n/c(x_i^n) = 0. \tag{3.26}$$

Then, rewrite (3.1a) as

$$c(x_i^n)\frac{w_i^n - g(t^n - k_i^n)}{k_i^n} - \delta_{\bar{x}}(a\delta_x w)_i^n = f_i^n. \tag{3.27}$$

The procedure is now well defined, and the error analysis is essentially unchanged, so that the obvious analogues of (3.20) and (3.24) hold.

The treatment of a Neumann condition is slightly more complicated. Let

$$\frac{\partial u}{\partial x}(0,t)=g(t), \qquad t \geqq 0, \tag{3.28}$$

and assume that $b(x)>0$ in the neighborhood of $x=0$. Assume that h_1^n and h_2^n are such that $\bar{x}_1^n<0$ and $\bar{x}_2^n \geqq 0$, so that only the equation for $i=1$ must be treated specially. Apply (3.28) by assigning the condition

$$w_0^n = w_1^n - h_1^n g(t^n) \tag{3.29}$$

and assign the value of the approximate solution at $(0, t^n - k_1^n)$ as the linear interpolate

$$\tilde{w}_1^{n-1} = \left(1-\frac{k_1^n}{\Delta t}\right)(w_1^n - h_1^n g(t^n)) + \frac{k_1^n}{\Delta t} w_0^{n-1}. \tag{3.30}$$

After a bit of manipulation on the relation

$$c(x_1^n)\frac{w_1^n-\tilde{w}_1^{n-1}}{k_1^n} - \delta_{\bar{x}}(a\delta_x w)_1^n = f_1^n$$

using (3.26), it follows that

$$c(x_1^n)\frac{w_1^n - w_0^{n-1}}{\Delta t} - \delta_{\bar{x}}(a\delta_x w)_1^n = f_1^n - b(x_1^n)\left(1-\frac{k_1^n}{\Delta t}\right) g(t^n). \tag{3.31}$$

Again the maximum principle argument can be carried through, and an analogous error estimate results.

4. A finite difference method based on quadratic interpolation. The procedures discussed in the previous section treat the case of strong domination of diffusion by convection. This section is intended to recover second order accuracy in the spatial discretization when the convection and diffusion are roughly equivalent. Let us return to uniform intervals, $h_i^n = h$, and modify the definition of $w^{n-1}(x)$. We anticipate an error of size $O(h^2+\Delta t)$; consequently, we assume that

$$\Delta t = O(h^2) \quad \text{as } h \to 0. \tag{4.1}$$

Thus, the point $\bar{x}_i$ falls between x_{i-1} and x_{i+1} and, in fact converges to x_i as h tends to zero, by (2.1), so that the values w_{i-1}^{n-1}, w_i^{n-1}, and w_{i+1}^{n-1} can be used to define $w^{n-1}(\bar{x}_i)$:

$$\begin{aligned} w^{n-1}(\bar{x}_i) &= (I_2\{w_j^{n-1}\})(\bar{x}_i) \\ &= \tfrac{1}{2}\alpha_i^2(w_{i+1}^{n-1}+w_{i-1}^{n-1}) + (1-\alpha_i^2)w_i^{n-1} + \tfrac{1}{2}\alpha_i(w_{i+1}^{n-1}-w_{i-1}^{n-1}), \end{aligned} \tag{4.2}$$

where

$$\alpha_i = -b_i\,\Delta t / c_i h. \tag{4.3}$$

Then, it follows simply from the Peano kernel theorem that

$$|((1-I_2)u^{n-1})(\bar{x}_i)| \leqq K\|u^{n-1}\|_{W^{3,2}([x_{i-1},x_{i+1}])} h^{3/2}\,\Delta t. \tag{4.4}$$

We shall employ the l^2-norm in this section, where

$$\langle \alpha, \beta\rangle = \sum_i \alpha_i\beta_i h, \qquad \|\alpha\| = \|\alpha\|_{l^2} = \langle \alpha, \alpha\rangle^{1/2}. \tag{4.5}$$

We shall also use the following notation:

$$\|\alpha\|_{\tilde{l}^2}^2 = \sum_i \max\{|\alpha(x)|^2 : |x - x_i| \leqq Kh\} \cdot h, \tag{4.6}$$

where the constant K is that of (2.1). Then, the truncation terms in (3.4a) satisfy the bound

$$\|e^n\|^2 \leqq K\left\{\|u^n\|_4^2 h^4 + \left\|\frac{\partial^2 u}{\partial \tau^2}\right\|^2_{L^2(t^{n-1},t^n;\tilde{l}^2)} \Delta t\right\}. \tag{4.7}$$

Take the inner product of the relation (3.4a) against the test function ζ_i^n. Then,

$$\frac{1}{2\,\Delta t}\{\langle c\zeta^n, \zeta^n\rangle - \langle c\bar{\zeta}^{n-1}, \bar{\zeta}^{n-1}\rangle\} + \langle a\delta_x\zeta^n, \delta_x\zeta^n\rangle \leqq \|e^n\|\|\zeta^n\| \leqq \tfrac{1}{2}\|e^n\|^2 + \tfrac{1}{2}\|\zeta^n\|^2. \tag{4.8}$$

Next, we relate $\langle c\bar{\zeta}^{n-1}, \bar{\zeta}^{n-1}\rangle$ to $\langle c\zeta^{n-1}, \zeta^{n-1}\rangle$. Note that

$$\bar{\zeta}_i^{n-1} = (I_2\{\zeta_i^{n-1}\})(\bar{x}_i) + ((1-I_2)u^{n-1})(\bar{x}_i);$$

thus, since $(\Delta t)^2 h^2 = O(\Delta t)$,

$$\begin{aligned}
&\langle c\bar{\zeta}^{n-1}, \bar{\zeta}^{n-1}\rangle \\
&= \sum c_i((I_2\zeta^{n-1})(\bar{x}_i))^2 h + O(\|\zeta^{n-1}\|\|u^{n-1}\|_3 h^2\,\Delta t + \|u^{n-1}\|_2^2 h^4 (\Delta t)^2) \\
&= \sum c_i\left\{(\zeta_i^{n-1})^2 + O(\Delta t)((\zeta_{i-1}^{n-1})^2 + (\zeta_i^{n-1})^2 + (\zeta_{i+1}^{n-1})^2) + \frac{b_i\,\Delta t}{2c_i h}(\zeta_i^{n-1}\zeta_{i+1}^{n-1} - \zeta_i^{n-1}\zeta_{i-1}^{n-1})\right\} h \\
&\quad + O(\|\zeta^{n-1}\|^2\,\Delta t + \|u^{n-1}\|_3^2 h^4\,\Delta t) \\
&= \langle c\zeta^{n-1}, \zeta^{n-1}\rangle + \tfrac{1}{2}\sum (b_{i-1} - b_i)\zeta_i^{n-1}\zeta_{i-1}^{n-1}\,\Delta t + O(\|\zeta^{n-1}\|^2\Delta t + \|u^{n-1}\|_3^2 h^4\,\Delta t) \\
&= \langle c\zeta^{n-1}, \zeta^{n-1}\rangle + O(\|\zeta^{n-1}\|^2\,\Delta t + \|u^{n-1}\|_3^2 h^4\,\Delta t).
\end{aligned} \tag{4.9}$$

Then, (4.8) and (4.9) imply that

$$\begin{aligned}
&\frac{1}{2\,\Delta t}[\langle c\zeta^n, \zeta^n\rangle - \langle c\zeta^{n-1}, \zeta^{n-1}\rangle] + \langle a\delta_x\zeta^n, \delta_x\zeta^n\rangle \\
&\qquad \leqq K(\|\zeta^n\|^2 + \|\zeta^{n-1}\|^2 + K\left((\|u^n\|_4^2 + \|u^{n-1}\|_3^2)h^4 + \left\|\frac{\partial^2 u}{\partial \tau^2}\right\|^2_{L^2(t^{n-1},t^n;\tilde{l}^2)} \Delta t\right),
\end{aligned} \tag{4.10}$$

from which it follows that

$$\max_{0\leqq t^n\leqq T} \|\zeta^n\|_{l^2} + \|\zeta\|_{l^2(0,T;w^{1,2}(\mathbb{R}))} \leqq K\left(\|u\|_{L^\infty(0,T;W^{4,2}(\mathbb{R}))}h^2 + \left\|\frac{\partial^2 u}{\partial \tau^2}\right\|_{L^2(0,T;\tilde{l}^2)} \Delta t\right). \tag{4.11}$$

THEOREM 4. *Let the solution u of* (1.1) *be in the space* $L^\infty(0, T; W^{4,2}(\mathbb{R}))$ *and assume that the derivative* $\partial^2 u/\partial\tau(x)^2 \in L^2(0, T; \tilde{l}^2)$. *Let w be the solution of* (3.1), *where* $w^{n-1}(x)$ *is given by* (4.2). *Then, the error* $\zeta = u - w$ *satisfies the inequality* (4.11).

5. Extensions and applications. The easiest and most obvious extension of the methods and results of the finite element procedure (2.5) and the finite difference procedures associated with (3.1) is to allow the x-variable to be n-dimensional. Nothing in the finite element development had a dependence on the dimension of the spatial domain. In particular, the changes of variable used in the Galerkin proofs can be extended to the multi-dimensional case by simple topological arguments. For the finite difference methods, a uniform grid is required for the quadratic interpolation procedure to give second order accuracy in h, because of the $\nabla\cdot(a\nabla u)$-term. With linear interpolation it is sufficient that the mesh can be a product of not necessarily uniform, one-dimensional meshes. It is also clear that the coefficients a, b, and c can depend explicitly on the time without affecting the convergence results. The algebraic equations arising at each time step are symmetric and can be treated by conjugate gradient methods.

Nonlinear problems can also be treated by the methods. Consider, for example, the problem given by

$$(5.1)\qquad \begin{aligned} &\text{(a)} \qquad c(x,u)\frac{\partial u}{\partial t}+b(x,u)\frac{\partial u}{\partial x}-\frac{\partial}{\partial x}\left(a\frac{\partial u}{\partial x}\right)=f, \qquad x\in\mathbb{R},\quad t>0,\\ &\text{(b)} \qquad u(x,0)=u_0(x), \qquad x\in\mathbb{R}, \end{aligned}$$

where a and f will be taken as functions of x and t alone to emphasize the more important aspects of the methods. A finite difference method generalizing (3.1) can be formulated as below. Note that, if we use piecewise-linear interpolation as in § 3, we anticipate convergence with an error that is $O(h+\Delta t)$. Thus, we should be able to evaluate coefficients with w_i^{n-1} replacing u. So, let

$$(5.2)\qquad X_i^n=x_i-b(x_i,w_i^{n-1})\,\Delta t/c(x_i,w_i^{n-1}),$$

and take the finite difference procedure to be

$$(5.3)\qquad c(x_i,w_i^{n-1})\frac{w_i^n-w^{n-1}(X_i^n)}{\Delta t}-\delta_{\bar{x}}(a\delta_x w)_i^n=f_i^n, \qquad n\geqq 1,\quad \forall i.$$

An easy calculation shows that

$$(5.4)\qquad \left(\psi\frac{\partial u}{\partial \tau}\right)_i^n=c(x_i,u_i^n)\frac{u_i^n-u(\bar{x}_i^n,t^{n-1})}{\Delta t}+O(\Delta t),$$

where

$$(5.5)\qquad \bar{x}_i^n=x_i-b(x_i,u_i^n)\,\Delta t/c(x_i,u_i^n).$$

Then, it follows that the error $\zeta=u-w$ satisfies the relation (with explicit dependence on x suppressed in the notation and $e_i^n=O(h+\Delta t)$)

$$(5.6)\qquad \begin{aligned} &c(w_i^{n-1})\frac{\zeta_i^n-\zeta^{n-1}(\bar{X}_i^n)}{\Delta t}-\delta_{\bar{x}}(a\delta_x\zeta)_i^n\\ &\quad=e_i^n+[c(w_i^{n-1})-c(u_i^n)]\frac{u_i^n-u^{n-1}(\bar{X}_i^n)}{\Delta t}+c(u_i^n)\frac{u^{n-1}(\bar{x}_i^n)-u^{n-1}(\bar{X}_i^n)}{\Delta t}. \end{aligned}$$

If $\gamma=b/c$, then

$$(5.7)\qquad \bar{x}_i^n-\bar{X}_i^n=\Delta t\left(\frac{\partial\gamma}{\partial u}\frac{\partial u}{\partial t}\Delta t+\frac{\partial\gamma}{\partial u}\zeta_i^{n-1}\right)=O((\Delta t)^2+\max_j|\zeta_j^{n-1}|\,\Delta t).$$

A few more lines of calculation will lead to the conclusion that

$$(5.8)\qquad \max_i|\zeta_i^n|=O(h+\Delta t),$$

as in the linear case.

Similarly, a finite element extension of (2.5) can be set up as follows. Let

$$\bar{x}^n = x - \frac{b(x, u_h^{n-1})}{c(x, u_h^{n-1})}\Delta t \quad \text{and} \quad \bar{u}_h^{n-1}(x) = u_h^{n-1}(\bar{x}^n),$$

and define u_h^n by

$$\left(c(x, u_h^{n-1})\frac{u_h^n - \bar{u}_h^{n-1}}{\Delta t}, v\right) + A(u_h^n, v) = (f^n, v), \qquad v \in \mathcal{M}_h. \tag{5.9}$$

For the error analysis, take w_h as before, and set

$$\frac{\partial}{\partial \tau(x, u)} = \frac{c(x, u)}{\psi(x, u)}\frac{\partial}{\partial t} + \frac{b(x, u)}{\psi(x, u)}\frac{\partial}{\partial x}, \quad \tilde{x}(t) = x - \frac{b(x, u(x, t))}{c(x, u(x, t))}\Delta t, \quad \tilde{u}^{n-1}(x) = u^{n-1}(\tilde{x}(t^n)).$$

A calculation gives the analogue of (2.11),

$$\begin{aligned} &\left(c(u_h^{n-1})\frac{\xi^n - \bar{\xi}^{n-1}}{\Delta t}, v\right) + A(\xi^n, v) \\ &\quad = \left(\psi(u^n)\frac{\partial u^n}{\partial \tau} - c(u^n)\frac{u^n - \tilde{u}^{n-1}}{\Delta t}, v\right) + \left(c(u^n)\frac{\bar{u}^{n-1} - \tilde{u}^{n-1}}{\Delta t}, v\right) \\ &\qquad + \left([c(u^n) - c(u_h^{n-1})]\frac{u^n - \bar{u}^{n-1}}{\Delta t}, v\right) + \left(c(u_h^{n-1})\frac{\eta^n - \bar{\eta}^{n-1}}{\Delta t}, v\right). \end{aligned} \tag{5.10}$$

The first and last terms on the right-hand side of (5.10) (as well as the left-hand side) can be treated as before. The second term is handled by a calculation like (5.7), so that

$$\left\|\frac{\bar{u}^{n-1} - \tilde{u}^{n-1}}{\Delta t}\right\|^2 \leqq K((\Delta t)^2 + \|\eta^{n-1}\|^2 + \|\xi^{n-1}\|^2). \tag{5.11}$$

The third term, after bounding $(u^n - \bar{u}^{n-1})/\Delta t$ in L^∞, leads to

$$\left\|[c(u^n) - c(u_h^{n-1})]\frac{u^n - \bar{u}^{n-1}}{\Delta t}\right\|^2 \leqq K((\Delta t)^2 + \|\eta^{n-1}\|^2 + \|\xi^{n-1}\|^2). \tag{5.12}$$

Combining these estimates, we obtain the same results as in the linear problem.

The authors have considered [1], [4] the applications of the techniques of this paper to the nonlinear system

$$\begin{aligned} &\text{(a)} \qquad \phi(x)\frac{\partial s}{\partial t} + v \cdot \nabla s - \frac{\partial}{\partial x_i}\left(D_{ij}(v)\frac{\partial s}{\partial x_j}\right) = f_1(x, t, s), \\ &\text{(b)} \qquad \nabla \cdot v = -\nabla \cdot \left(\frac{k(x)}{\mu(s)}\nabla p\right) = f_2(x, t), \end{aligned} \tag{5.13}$$

which describes a model for the miscible displacement of one fluid by another in a porous medium. A closely related procedure has been analyzed by Pironneau [3] for the Navier–Stokes equations.

Numerical experiments have been conducted with our Galerkin procedure for a constant-coefficient version of (1.1) [2] and for the miscible displacement system (5.13) in two space dimensions [5]. As our theory predicts, much larger time steps were feasible than with standard methods. In [2], time-truncation error was reduced to the order of spatial error by values of Δt two to three orders of magnitude larger than those for standard time stepping. In [5], time steps were one order of magnitude larger than those of previous studies in the petroleum engineering literature, and the usual numerical difficulties were absent or sharply reduced. For details, see [2], [5].

REFERENCES

[1] J. DOUGLAS, JR., *Finite difference methods for two-phase, incompressible flow in porous media*, to appear.

[2] R. E. EWING AND T. F. RUSSELL, *Multistep Galerkin methods along characteristics for convection-diffusion problems*, Advances in Computer Methods for Partial Differential Equations—IV, R. Vichnevetsky and R. S. Stepleman, eds., IMACS, Rutgers Univ., New Brunswick, NJ, 1981, pp. 28–36.

[3] O. PIRONNEAU, *On the transport-diffusion algorithm and its application to the Navier–Stokes equations*, to appear.

[4] T. F. RUSSELL, *An incompletely iterated characteristic finite element method for a miscible displacement problem*, Thesis, Univ. of Chicago, Chicago, IL, June 1980.

[5] ———, *Finite elements with characteristics for two-component incompressible miscible displacements*, Rep. SPE 1–500, Society of Petroleum Engineers, Dallas, TX, 1982.

[6] M. F. WHEELER, *A priori L^2-error estimates for Galerkin approximations to parabolic partial differential equations*, this Journal, 10 (1973), pp. 723–759.