

Generalized Finite-Difference Schemes

By Blair Swartz* and Burton Wendroff**

Abstract. Finite-difference schemes for initial boundary-value problems for partial differential equations lead to systems of equations which must be solved at each time step. Other methods also lead to systems of equations. We call a method a generalized finite-difference scheme if the matrix of coefficients of the system is sparse. Galerkin's method, using a local basis, provides unconditionally stable, implicit generalized finite-difference schemes for a large class of linear and nonlinear problems. The equations can be generated by computer program. The schemes will, in general, be not more efficient than standard finite-difference schemes when such standard stable schemes exist. We exhibit a generalized finite-difference scheme for Burgers' equation and solve it with a step function for initial data. ■

1. Well-Posed Problems and Semibounded Operators. We consider a system of partial differential equations of the following form:

$$(1.1) \quad \partial u / \partial t = Au + f,$$

where $u = (u_1(x, t), \dots, u_m(x, t))$, $f = (f_1(x, t), \dots, f_m(x, t))$, and A is a matrix of partial differential operators in $x = (x_1, \dots, x_n)$,

$$\begin{aligned} A &= A(x, t, D) = \sum a_i D^i, \\ D^i &= (\partial / \partial x_1)^{i_1} \cdots (\partial / \partial x_n)^{i_n}, \\ a_i(x, t) &= a_{i_1 \dots i_n}(x, t) = \text{matrix}. \end{aligned}$$

Equation (1.1) is assumed to hold in some n -dimensional region Ω with boundary $\partial\Omega$.

An initial condition is imposed in Ω ,

$$(1.2) \quad u(x, 0) = u_0(x), \quad x \in \Omega.$$

The boundary conditions are that there is a collection \mathfrak{B} of operators $B(x, D)$ such that

$$(1.3) \quad B(x, D)u = 0, \quad x \in \partial\Omega, B \in \mathfrak{B}.$$

We assume the operator A satisfies the following condition: *There exists a scalar product \langle, \rangle such that for all sufficiently smooth functions $\phi(x)$ which satisfy (1.3),*

$$(1.4) \quad 2 \operatorname{Re} \langle A\phi, \phi \rangle \leq C \langle \phi, \phi \rangle, \quad 0 < t \leq T,$$

where C is a constant independent of ϕ . An operator A satisfying (1.4) is called semi-bounded. A general theory of such operators is given by Kreiss [6], Lions [7], and

Received December 8, 1967, revised May 27, 1968.

* Work performed under the auspices of the U. S. Atomic Energy Commission, Contract No. W-7405-Eng. 36.

** Partly supported by N. S. F. Contract No. GP-7641, and partly supported by U.S.A.E.C. Contract No. W-7405-Eng. 36.

in the case $C = 0$ by Lumer and Phillips [8] (see also Yosida [15]). If the operator A is independent of t , and if $f = 0$, this condition is essentially necessary and sufficient for the solution operator to form a semigroup on a suitable space; therefore (1.4) is not really a restriction for physically interesting problems.

It now follows from (1.1) that

$$\langle u_t, u \rangle + \langle u, u_t \rangle = (d/dt)\langle u, u \rangle \leq C\langle u, u \rangle + 2|\langle f, u \rangle|,$$

or

$$\frac{d}{dt} \|u\|^2 \leq (C + 1)\|u\|^2 + \|f\|^2.$$

Therefore,

$$(1.5) \quad \|u(t)\|^2 \leq e^{(C+1)t}\|u(0)\|^2 + \int_0^t e^{(C+1)(t-t')}\|f(t')\|^2 dt'.$$

In other words, we have an energy inequality for the solution $u(t)$, which means that our problem is well posed.

A trivial example of a semibounded operator is the scalar operator

$$A = -\partial/\partial x + \sigma, \quad x > 0,$$

with boundary condition $u(0, t) = 0$, for if we put $\langle u, v \rangle = \int_0^a u\bar{v}dx$, then

$$\begin{aligned} \langle Au, u \rangle &= -\int_0^a u_x \bar{u} dx + \int_0^a \sigma u \bar{u} dx \\ &= -|u(a)|^2 + \int_0^a u \bar{u}_x dx + \int_0^a \sigma u \bar{u} dx \\ &= -\overline{\langle Au, u \rangle} - |u(a)|^2 + 2 \int_0^a \operatorname{Re} \sigma u \bar{u} dx. \end{aligned}$$

Thus,

$$2 \operatorname{Re} \langle Au, u \rangle \leq 2 \max |\sigma| \langle u, u \rangle.$$

The classical linear equations of mathematical physics, with suitable boundary conditions, are all based on semibounded operators. In a later section we discuss an extension of this notion to nonlinear equations, in particular Burgers' equation.

2. An Approximating System of Ordinary Differential Equations. Let $\phi_i(x)$, $i = 1, 2, \dots, N$, be linearly independent functions (vector-valued) satisfying (1.3). In the applications each ϕ_i will depend on N , but this dependence is suppressed for notational simplicity.

We seek a function $v(x, t)$ of the form

$$(2.1) \quad v(x, t) = \sum c_i(t) \phi_i(x).$$

This function will be the solution of the following system of ordinary differential equations

$$(2.2) \quad \langle v_t - Av - f, \phi_i \rangle = 0, \quad i = 1, 2, \dots, N.$$

Multiplying the i th equation by \bar{c}_i and summing we have

$$(2.3) \quad \langle v_t - Av - f, v \rangle = 0.$$

It follows immediately that $v(x, t)$ satisfies the energy inequality (1.5).

The initial condition on $v(x, t)$ may be taken to be

$$(2.4) \quad \langle v, \phi_i \rangle = \langle u(x, 0), \phi_i \rangle, \quad i = 1, 2, \dots, N.$$

In that case

$$\langle v(0), v(0) \rangle = \langle u(0), v(0) \rangle \leq \|u(0)\| \|v(0)\|,$$

or

$$(2.5) \quad \|v(0)\| \leq \|u(0)\|.$$

It follows from the energy inequality that $\|v(t)\|$ is uniformly bounded for $0 < t \leq T$, independently of the choice of ϕ_i .

To discuss the convergence of $v(x, t)$ to $u(x, t)$, let $w(x, t)$ be any function of the form $w(x, t) = \sum w_i(t)\phi_i(x)$ and let $Lu = \partial u/\partial t - Au - f$. Then

$$\|u - v\| \leq \|u - w\| + \|w - v\|.$$

Now, let $e = v - w$. Then, $\langle Lv - Lw, e \rangle = -\langle Lw, e \rangle$, and

$$\frac{d}{dt} \|e\|^2 \leq C\|e\|^2 + 2|\langle Lw, e \rangle|.$$

So $e(t)$ satisfies (1.5) with u replaced by e and f by Lw . Therefore we have convergence in the norm, as $N \rightarrow \infty$, if for each N there exists a function $w(x, t) = \sum w_i(t)\phi_i(x)$ such that $\|w - v\|(0) \rightarrow 0$ and

$$\|u - w\| \rightarrow 0, \quad \|Lw\| \rightarrow 0$$

uniformly in t . This means that $u(x, t)$ must be well approximated by linear combinations of the ϕ_i .

It is also possible to obtain an *a posteriori* error estimate. Let $\delta = u - v$. Since $Lu = 0$, $\langle Lu - Lv, \delta \rangle = -\langle Lv, \delta \rangle$, therefore

$$(2.6) \quad \|\delta(t)\|^2 \leq e^{(C+1)t} \|\delta(0)\|^2 + \int_0^t e^{(C+1)(t-t')} \|Lv\|^2 dt'.$$

The existence of this error estimate makes this method particularly attractive; however, we do not pursue this any further in this paper. The system (2.2) is discussed in Lions [7].

3. An Approximating Discrete System. The differential equations obtained in the preceding section can only be solved numerically; therefore one might as well start out by replacing the time derivative by a difference quotient. We do this as follows: Let $v^n(x)$ be a function of the form

$$(3.0) \quad v^n(x) = \sum c_i^n \phi_i(x).$$

For each n let

$$(3.1) \quad \left\langle v^{n+1} - v^n - A(t^{n+1/2}) \frac{\Delta t}{2} (v^{n+1} + v^n) - f\Delta t, \phi_i \right\rangle = 0, \quad i = 1, 2, \dots, N.$$

A different system is considered by Lions [7]. We take initially

$$(3.2) \quad \langle v^0, \phi_i \rangle = \langle u_0, \phi_i \rangle.$$

The matrix $\langle \phi_i, \phi_j \rangle$ is positive definite, so (3.2) has a unique solution $\{c_1^0, \dots, c_N^0\}$. Suppose the c_i^n have been found. Then (3.1) has a unique solution if and only if the system

$$\langle v^{n+1} - A(t^{n+1/2}) \Delta t v^{n+1}/2, \phi_i \rangle = 0$$

has only the trivial solution. Multiplying by \bar{c}_i^{n+1} and summing we have

$$\langle v^{n+1}, v^{n+1} \rangle = \langle A \Delta t v^{n+1}/2, v^{n+1} \rangle.$$

Then

$$\begin{aligned} \langle v^{n+1}, v^{n+1} \rangle &= \operatorname{Re} \langle A \Delta t v^{n+1}/2, v^{n+1} \rangle \\ &\leq (\Delta t C/4) \langle v^{n+1}, v^{n+1} \rangle. \end{aligned}$$

Thus, if

$$(3.3) \quad C \Delta t < 4,$$

(3.1) has a unique solution.

We obtain an energy inequality from (3.1) by multiplying by $(\bar{c}_i^{n+1} + \bar{c}_i^n)$ and summing over i . This leads to

$$\|v^{n+1}\|^2 - \|v^n\|^2 \leq C \frac{\Delta t}{4} \|v^{n+1} + v^n\|^2 + \Delta t |\langle f, v^{n+1} + v^n \rangle|.$$

Using the Schwarz inequality, the triangle inequality, and dividing by $\|v^{n+1}\| + \|v^n\|$ we get

$$(3.4) \quad \|v^{n+1}\| \leq \frac{1 + C \Delta t/4}{1 - C \Delta t/4} \|v^n\| + \frac{\Delta t}{1 - C \Delta t/4} \|f\|.$$

This implies that

$$(3.5) \quad \|v^n\| \leq K_1(\Delta t, t) \|v^0\| + \left(\max_{0 < t \leq T} \|f\| \right) K_2(\Delta t, t)$$

where

$$\lim_{\Delta t \rightarrow 0} K_1(\Delta t, t) = e^{Ct/2}, \quad \lim_{\Delta t \rightarrow 0} K_2(\Delta t, t) = \int_0^t e^{C(t-t')/2} dt'.$$

Let $w^n(x) = \sum w_i^n \phi_i(x)$. Let

$$\bar{L}w = \frac{w^{n+1} - w^n}{\Delta t} - \frac{1}{2} A(w^{n+1} + w^n) - f.$$

Then, similarly as in Section 2, if as $N \rightarrow \infty$, $\Delta t \rightarrow 0$,

$$(3.6) \quad \|u(n\Delta t) - w^n\| \rightarrow 0, \quad \|v^0 - w^0\| \rightarrow 0,$$

and

$$(3.7) \quad \|\bar{L}w\| \rightarrow 0$$

uniformly in n , $n\Delta t < T$, for some $w^n(x)$, we have convergence.

We have exploited the semiboundedness of the differential operator A to obtain

a convergent discrete scheme. This has also been done by Kreiss [6] in a different way.

4. Approximating Systems of Higher Order in Time. The system (3.1) has been shown to be convergent as $N \rightarrow \infty$, $t \rightarrow 0$ if (3.6) and (3.7) hold. We might, however, keep N fixed and think of (3.1) as an approximation to (2.2). In that case there are approximations of higher order in Δt that might be considered. Suppose that $f \equiv 0$ and A is independent of t . Then there is no loss of generality in assuming that $C = 0$, in which case we have $\|v^{n+1}\| \leq \|v^n\|$.

Now, let

$$(4.1) \quad P = \{\langle \phi_j, \phi_i \rangle\},$$

$$(4.2) \quad Q = \{\langle A\phi_j, \phi_i \rangle\}.$$

Then (2.2) becomes

$$(4.3) \quad dc/dt = P^{-1}Qc \equiv Sc,$$

where $c(t) = (c_1(t), \dots, c_N(t))$. Let

$$\langle c, d \rangle = \langle \sum c_i \phi_i(x), \sum d_i \phi_i(x) \rangle$$

and let $\|S\|$ be the induced norm. From (3.1) we have

$$c^{n+1} = (I - \Delta t S/2)^{-1}(I + \Delta t S/2)c^n.$$

Since

$$\|c^{n+1}\| = \|v^{n+1}\| \leq \|v^n\| = \|c^n\|$$

we have shown that

$$\|(I - \Delta t S/2)^{-1}(I + \Delta t S/2)\| \leq 1.$$

Let $r(z)$ be a diagonal element of the Padé table for e^z . Then $\|r(S\Delta t)\| \leq 1$ (see Thomeé [14]) and $c^{n+1} = r(S\Delta t)c^n$ will be a stable (in N and Δt) discretization of (4.3). By choosing the right $r(z)$, the truncation error (for fixed N), will be $O(\Delta t)^l$ for any l .

5. Nonlinear Problems. Much of the previous discussion can be extended to nonlinear problems. Suppose the problem is

$$(5.1) \quad u_t = A(u) + f,$$

where $A(u)$ is a nonlinear operator. Suppose further that

$$(5.2) \quad 2 \operatorname{Re} \langle A(u) - A(\bar{u}), u - \bar{u} \rangle \leq C \langle u - \bar{u}, u - \bar{u} \rangle.$$

Then (5.1) is well posed, for if u, \bar{u} are two solutions corresponding to different initial data, and if $\delta = u - \bar{u}$, then

$$\frac{d}{dt} \|\delta\|^2 \leq C \|\delta\|^2,$$

so that we again have an energy inequality.

An example of an operator satisfying (5.2) is

$$A(u) = \frac{\partial}{\partial x} f(u_x), \quad f'(y) \leq 0, \quad u(0) = u(1) = 0.$$

We set up a nonlinear approximating discrete system as follows:

$$(5.3) \quad \langle v^{n+1} - v^n - \Delta t A((v^{n+1} + v^n)/2) - \Delta t f, \phi_i \rangle = 0, \quad i = 1, 2, \dots, N.$$

Let $c = (c_1^{n+1}, \dots, c_N^{n+1})$, and let $P = \{\langle \phi_j, \phi_i \rangle\}$. Then (5.3) can be rewritten as

$$(5.4) \quad c + G(c) = 2b$$

where

$$b_i = c_i^n + \Delta t \sum_j (P^{-1})_{ij} \langle f, \phi_j \rangle, \\ (G(c))_i = c_i - 2\Delta t \sum_j (P^{-1})_{ij} \left\langle A\left(\frac{v^{n+1} + v^n}{2}\right), \phi_j \right\rangle.$$

For any vectors c, \bar{c} in \mathbb{C}^N , let $v(x) = \sum c_i \phi_i(x)$, $\bar{v}(x) = \sum \bar{c}_i \phi_i(x)$, and let $(c, \bar{c}) = \langle v, \bar{v} \rangle$. Then $(c, \bar{c}) = \sum_i (Pc)_i \bar{c}_i$, so

$$(G(c) - G(\bar{c}), c - \bar{c}) = \langle v - \bar{v}, v - \bar{v} \rangle - 4\Delta t \left\langle A\left(\frac{v + \bar{v}}{2}\right) - A\left(\frac{\bar{v} + v^n}{2}\right), \frac{v - \bar{v}}{2} \right\rangle.$$

Therefore,

$$\operatorname{Re} (G(c) - G(\bar{c}), c - \bar{c}) \geq (1 - C\Delta t/2) \langle v - \bar{v}, v - \bar{v} \rangle \geq 0$$

if $C\Delta t < 2$.

Thus, $G(c)$ is a *monotone* operator in the sense of Minty [9]. Minty proves, for Hilbert spaces, not just \mathbb{C}^N , that (5.4) has a unique solution if $G(c)$ is monotone.

Of course (5.4) will have to be solved by iteration. In the numerical example considered in the next section a form of nonlinear successive substitution was used. We conjecture that there is some "natural" iteration which is sure to be convergent.

A nonlinear problem which does not quite fit the hypothesis of semiboundedness is provided by Burgers' equation:

$$(5.5) \quad A(u) = \epsilon u_{xx} - uu_x, \quad f = 0, \quad \epsilon > 0$$

with boundary conditions $u(0) = u(1) = 0$. It is shown in Lions [7] that if

$$\langle u, v \rangle = \int_0^1 u(x)v(x)dx,$$

then $\langle A(u) - A(\bar{u}), u - \bar{u} \rangle \leq C(u) \langle u - \bar{u}, u - \bar{u} \rangle$. In this case the constant $C(u)$ depends on one of the functions. However, we can still make a statement about convergence; we now require in addition to (3.6), (3.7), that $C(w^n) \leq \text{const}$ for $n\Delta t \leq T, N \rightarrow \infty$.

There are some practical questions concerning the use of (5.3). Suppose that $\langle u, v \rangle = \int u(x)v(x)dx$. Then, except in very simple cases, the numbers $\langle A(v), \phi_i \rangle$ should be generated by quadrature formulas. That is to say, if $g(x) = A(v)(x)$ then

$$(5.6) \quad \langle A(v), \phi_i \rangle \cong \sum \alpha_j g(x_j) \phi_i(x_j)$$

with suitable weights $\{\alpha_j\}$, and nodes x_j . In this way the equations (5.4) can be formed by a computer program. Done in this way, however, the method is very

inefficient, for $\langle A(v), \phi_i \rangle$ must be recomputed at each iteration. Even in the linear case $\langle Av, \phi_i \rangle$ must be recomputed at each time step if the coefficients of the differential equation depend on t (but see above (5.4)).

One possibility for speeding things up is to replace $A(v)(x)$ by a linear combination of the ϕ_i , by interpolation. Probably, however, some derivatives should be removed by integration by parts. For example, for Burgers' equation

$$(5.7) \quad \int_0^1 (u^2)_x \phi_i dx = - \int_0^1 u^2 \frac{\partial \phi_i}{\partial x} dx .$$

Now replace $u^2(x)$ by $\sum d_j \phi_j(x)$, so that $\int \phi_j (\partial \phi_i / \partial x) dx$ need be computed only once. This seems to work in this case (see next section). In general, of course, we would be perturbing the operator A , and there is no guarantee that the problem is stable with respect to such a perturbation.

6. Local Bases and Generalized Finite-Difference Schemes. We now observe that in many cases an inner product under which the operator A is semibounded is given by an integral over the set Ω . Therefore, if each ϕ_i vanishes outside some small set Ω_i , the matrices derived from (3.1) are sparse. The analogy with finite-difference schemes becomes more direct if there are, say, N of the x_i in Ω such that $\phi_j(x_i) = \delta_{ij}$; for then the c_i^n in (3.0) are $v^n(x_i)$ and, through (3.1), become approximations to $u(x_i, n \Delta t)$. Similarly, other unknowns might be the derivatives of v in various directions. In other cases (like that of the splines to be discussed), although $\phi_j(x_i) \neq \delta_{ij}$, v may still be regarded as giving a smooth interpolation of a function which approximates $u(x, n \Delta t)$ at the x_i . The ϕ_i to be considered here, then, are functions whose support in Ω is small and which satisfy the boundary conditions (1.3). The matrix problem then generated by (3.1) we shall call a *generalized finite-difference scheme*. The set of functions $\{\phi_i\}$ is called a local basis.

We now restrict our attention to one space dimension, reserving a comment on higher dimensional problems for the end.

As an example of (3.1) we consider the heat equation using piecewise linear approximations. That is, we consider

$$(6.1) \quad \begin{aligned} u_t &= u_{xx}, & (x, t) \in (0, 1) \times (0, T] ; \\ u(0, t) &= u(1, t) = 0, & u(x, 0) = u_0(x) . \end{aligned}$$

The operator $A = \partial^2 / \partial x^2$ is semibounded with respect to the inner product $\langle f, g \rangle = \int_0^1 f \bar{g} dx$, with constant $C = 0$. Let $h = 1/N$, and let ϕ_i be the local basis for piecewise linear interpolation of a function at $x_i = i/N, i = 1, 2, \dots, N - 1$:

$$(6.2) \quad \begin{aligned} \phi_i(x) &= (x - x_{i-1})/h, & x_{i-1} \leq x \leq x_i, \\ &= (x_{i+1} - x)/h, & x_i \leq x \leq x_{i+1}, \\ &= 0, & \text{otherwise .} \end{aligned}$$

Since $\phi_i(x_j) = \delta_{ij}$ the coefficients c_i^n of (3.0) may be regarded as $v_i^n = v^n(x_i)$. We make a weak extension of $\partial^2 / \partial x^2$ to continuous, piecewise differentiable functions v by setting

$$\langle \partial^2 v / \partial x^2, w \rangle \equiv - \int_0^1 v_x \bar{w}_x + v_x \bar{w} \Big|_0^1 .$$

(Although the piecewise linear functions are not in the domain of $A = \partial^2/\partial x^2$ we have selected them for our example because the inner product, as extended, is quite easy to calculate and yields by (3.1) a matrix problem which bears a very strong resemblance to a quite reasonable finite-difference scheme.) Applying (3.1) we obtain equations which are (with the notation $v_{i,t} = (v_i^{n+1} - v_i^n)/\Delta t$ and $v_{xx}^n = (v_{i+1}^n - 2v_i^n + v_{i-1}^n)/h^2$)

$$(6.3) \quad \left(\frac{1}{6}\right)(v_{i+1,t} + 4v_{i,t} + v_{i-1,t}) = \left(\frac{1}{2}\right)(v_{xx}^{n+1} + v_{xx}^n).$$

(Next to the boundaries, (6.3) is to be regarded as holding with $v_0^n = v_N^n = 0$, all n .)

If $\Delta t = \lambda h^2$ this is equivalent to (5) on page 189 of Richtmyer and Morton [10] with $\theta = 1/2 - 1/6\lambda$; thus is quite close to their accurate and stable scheme (6) where $\theta = 1/2 - 1/12\lambda$. The initial condition (3.2) amounts to integral least squares approximation of the initial data by piecewise linear functions.

Because $\partial^2/\partial x^2$ remains (weakly) semibounded on the piecewise C^1 functions, the proof that $\|v^{n+1}\| \leq \|v^n\|$ of Section 3 holds. Our convergence analysis breaks down, however, for $\|\partial^2 v/\partial x^2\|$ remains undefined.

Suppose, now, the boundary conditions in (6.1) are of the form

$$(6.4) \quad u_x(0, t) = a_0 u(0, t), \quad u_x(1, t) = -a_1 u(1, t); \quad a_0, a_1 \geq 0,$$

(for which $A = \partial^2/\partial x^2$ remains semibounded as before). A local basis, ψ_i , for the space of piecewise linear functions satisfying (6.4) is given by $\psi_i = \phi_i$, $2 \leq i \leq N - 2$; together with the linear combination, ψ_1 , of ϕ_0 and ϕ_1 which satisfies (6.4) and $\psi_1(x_1) = 1$, plus a similar linear combination, ψ_{N-1} , of ϕ_{N-1} and ϕ_N . It is seen that $v(0)$ and $v(1)$ do not enter the resulting matrix problem; from the difference equation viewpoint the "boundary" equations have already been substituted in the "interior" equations. However, one can regard the resulting matrix equation as (6.3) at each interior mesh point together with two additional equations for $v_0^n = v(0)$ and $v_N^n = v(1)$ of the form

$$\begin{aligned} (v_1^n - v_0^n)/h &= [a_0 + O(h)]v_1^n; \\ (v_N^n - v_{N-1}^n)/h &= [-a_1 + O(h)]v_{N-1}^n. \end{aligned}$$

Other schemes, probably more accurate at the boundaries, could be found by placing the boundary $h/2$ from its interior neighbor or by using local bases for piecewise quadratic (or higher degree) interpolation near each boundary. The support of the latter local basis functions spreads farther than h into the interior; thus, from the difference scheme view, they let more interior mesh points interact with the boundary data.

We should emphasize that we do not necessarily intend that the user of this method explicitly calculate the coefficients of the matrix problems, as we have done in (6.3). Rather, we expect the coefficients to be obtained by numerical quadrature. As we have indicated in Section 5, this will be very inefficient unless the coefficients of the differential operator are independent of t , or unless one is willing to replace those coefficients by functions of the form $\sum \alpha_i(t)\beta_i(x)$, in which case the quadratures need be done only once.

Before leaving the piecewise linear approximations for higher order ones we wish to consider another example: that of Burgers' equation

$$u_t + uu_x - \epsilon u_{xx} = 0.$$

It was indicated in Section 5 that the space operator is semibounded in the sense of (5.2) under the inner product $\langle f, g \rangle = \int_a^b f \bar{g} dx$ with appropriate boundary conditions. Applying the previously discussed piecewise linear technique we arrive at the matrix problem, (in the notation preceding (6.3) together with $v_\xi = (v_{i+1} - v_{i-1})/2h$),

$$\begin{aligned}
 (6.5) \quad 0 = & (v_{i+1,t} + 4v_{i,t} + v_{i-1,t})/6 \\
 & + \frac{1}{2}[v_\xi^n(v_{i+1}^n + v_i^n + v_{i-1}^n)/3 - \epsilon v_{x\xi}^n] \\
 & + \frac{1}{2}[v_\xi^{n+1}(v_{i+1}^{n+1} + v_i^{n+1} + v_{i-1}^{n+1})/3 - \epsilon v_{x\xi}^{n+1}].
 \end{aligned}$$

As initial data we used

$$\begin{aligned}
 u_0(x) &= 1, \quad 0 \leq x < 1/2, \\
 &= 0, \quad 1/2 \leq x \leq 1;
 \end{aligned}$$

and took $v(x)$ at $t = 0$ to be the integral least squares piecewise linear approximation to $u_0(x)$. (It might be noted that this type of approximation exhibits symmetric Gibbs phenomenon of about 13% of the jump at each corner of the discontinuity.) Cole [4] exhibits the exact solution of Burgers' equation with such discontinuous data for the pure initial-value problem; it is a wave traveling to the right at velocity 1/2 which has the shape (asymptotically, as t becomes large) of $[1 - \tanh(x/4\epsilon)]/2$. We considered, however, the boundary-value problem, fixing $u(0, t) \equiv 1$ and $u(1, t) \equiv 0$. Rather than subtract off, say, $1 - x$ to make the boundary data homogeneous by putting a source in the equation, we simply formed (6.5) by applying (3.1) with

$$v(x) = \sum_{i=0}^N c_i^n \phi_i(x), \quad c_0^n \equiv 1, \quad c_N^n \equiv 0;$$

taking inner products only with the ϕ_i of (6.2) with $i = 1, \dots, N - 1$.

The matrix problem was solved at each time step by making a guess of v_i^{n+1} using (6.5) with the average of v_i replaced by $v_{i,t}$ and all space differences taken at time $n \Delta t$, then applying the following nonlinear successive displacement technique: regarding (6.5) as a system of equations $g_i(v^{n+1}) = 0$; replace v_i^{n+1} by $v_i^{n+1} - g_i(v^{n+1})/[\partial g_i/\partial v_i^{n+1}(v^{n+1})]$, and sweep from right to left ($i = N - 1, N - 2, \dots, 1$). Convergence of this process to the solution of (6.5) generally occurred in 6 to 8 sweeps (convergence was judged complete when (6.5) was satisfied with a maximum error of 10^{-4}). Sweeping from left to right took perhaps 3 times as many sweeps to attain comparable convergence.

Some of the results are indicated in Fig. 1. Fig. 1a is a graph of the initial data with $h = 1/100$. The half-width of the final traveling wave of this amplitude may be estimated from Cole's solution as 9ϵ to 14ϵ (depending on whether one can discriminate between 1% or 0.1% of the wave height and 0, respectively). Fig. 1b indicates the final form (taken at $t = 1/2$) of the traveling wave calculated using (6.5) with $\epsilon = 3 \times 10^{-3}$, $\Delta t = h = 1/100$. Fig. 1c shows the result under similar conditions with $\epsilon = 7.5 \times 10^{-3}$. Experiments with other parameters indicate (as might be expected) that, when the wave width is less than about $4h$ or $5h$, permanent oscillations are found in the shape of the numerical solution in back of the crest.

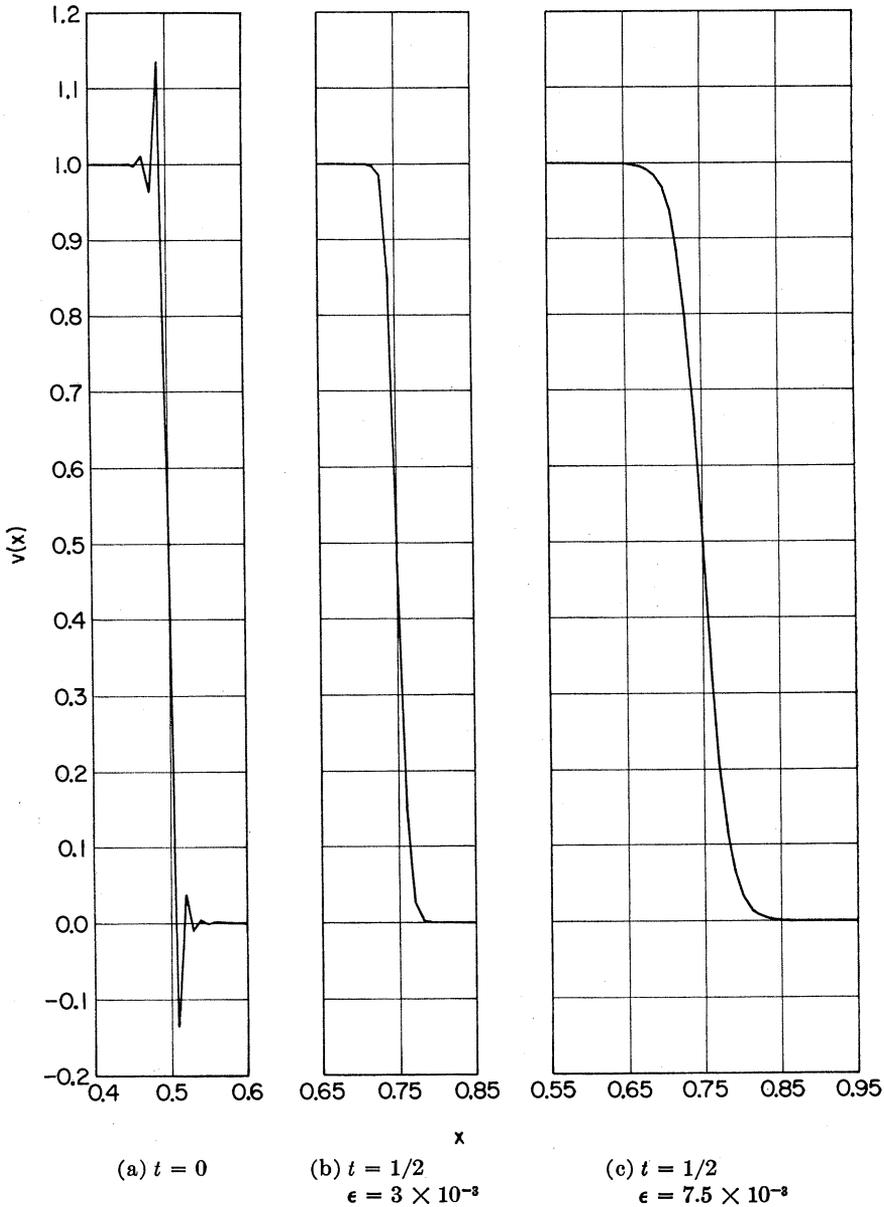


FIGURE 1. The significant part of some piecewise linear approximate solutions to Burgers' equation using (6.5) f with $\Delta t = h = 1/100$.

Anticipating that with a more complex set of basis functions or a more involved nonlinear term, one would really like to compute the inner products once and for all (rather than recomputing those involving the nonlinear terms at each time step), we tried expressing uu_x in Burgers' equation as $\frac{1}{2}(u^2)_x$, interpolating u^2 as $\sum u_i^2 \phi_i$, and then using (5.7). The resulting matrix problem is remarkably like (6.5); being obtained from it by replacing $(v_{i+1} + v_i + v_{i-1})/3$ by $(v_{i+1} + v_{i-1})/2$. The numerical results are essentially the same.

We now turn to the use of local bases which, if applied to interpolate a smooth function of one space variable, would yield more accurate approximations than piecewise linear interpolations do.

The first example is the local basis for local two-point Hermite interpolation of degree $2k + 1$ (see, for example Ciarlet, Schultz, and Varga [3]). The coefficients may be regarded as $v(x_i), v'(x_i), \dots, v^{(k)}(x_i)$. The local basis is two mesh intervals wide. Each element, ϕ_{ij} , is a polynomial of degree $2k + 1$ between the $x_i = i/N$; is in $C^k[0, 1]$; is identically zero outside $[x_{i-1}, x_{i+1}]$; and satisfies

$$\phi_{ij}^{(l)}(x_{i-1}) = \phi_{ij}^{(l)}(x_{i+1}) = 0, \quad \phi_{ij}^{(l)}(x_i) = \delta_{jl}, \quad l, j = 0, \dots, k; i = 0, \dots, N.$$

For a given mesh size, $h = 1/N$, there are about $(k + 1)N$ elements in the basis; and the $(k + 1)N \times (k + 1)N$ matrices which result from integral inner products have band widths of $4k + 3$. Error bounds for the interpolation of $f \in C^{2k+2}[0, 1]$ are of the form

$$\|f^{(l)} - \sum_{i,j} f^{(j)}(x_i)\phi_{ij}^{(l)}\|_{\infty} \leq O(h^{2k+2-l}), \quad 0 \leq l \leq 2k + 1.$$

The second example of such a local basis is the basis for polynomial spline interpolation of degree $m = 2k$ or $2k + 1$ (see, e. g., Schoenberg [11, p. 71], [12], and de Boor [5]). There are roughly N basis functions. Each element ϕ_i vanishes outside an interval of width $(m + 1)h$, is strictly positive inside that interval, and is in $C^{m-1}[0, 1]$. The roughly $N \times N$ matrices which result from integral inner products also have a band width of $2m + 1$. Error bounds for some forms of spline interpolation of $f \in C^{m+1}[0, 1]$ have been shown to be of the same order as those for the corresponding Hermite interpolation: $\|f^{(l)} - \sum c_i \phi_i^{(l)}\|_{\infty} \leq O(h^{m+1-l}), 0 \leq l \leq m$ (see, e. g., Swartz [13]).

Now, it is clear that we can satisfy any homogeneous boundary condition by only including those basis functions which are identically zero at the boundary; it is just as clear that we would be leaving out too many basis functions. In other words, we would be satisfying extra boundary conditions which would, in general, reduce the accuracy of our solution. So, we must find linear combinations of basis functions which satisfy the given boundary conditions, but, roughly speaking, not any others. We do this as follows for the scalar case: suppose the differential operator, A , is of order σ ; and suppose the boundary condition at $x = 0$ is

$$(6.6) \quad [Bu](0) \equiv \left[\left(\sum_{l \in \mathcal{L}} b_l D^l \right) u \right](0) = 0,$$

where \mathcal{L} is a nonempty subset of $\{0, 1, \dots, \sigma - 1\}$ (this restriction on the order of B is not essential); $b_l \neq 0, l \in \mathcal{L}$.

Consider first a local Hermite basis of degree $2k + 1$ not less than $2\sigma + 1$ (so that A may be applied without difficulty). If \mathcal{L} consists of a single integer, l , simply drop ϕ_{0l} from the collection of ϕ_{ij} . Otherwise, let \mathcal{L}' be \mathcal{L} without its largest integer, l' . Define

$$\begin{aligned} \psi_{0,l} &= \phi_{0,l'} - b_{l'}\phi_{0,l}/b_l, \quad l \text{ in } \mathcal{L}', \\ &= \phi_{0,l}, \quad l \text{ in } \{0, 1, \dots, \sigma - 1\} - \mathcal{L}; \end{aligned}$$

then the $\sigma - 1$ $\psi_{0,l}$ satisfy (6.6) and have their support in $[0, h]$. Similarly, $\sigma - 1$ ψ_{Nl} may be defined using the boundary condition at $x = 1$. These $(2\sigma - 2)$ func-

tions, together with the ϕ_{ij} , $j = 0, \dots, k$; $i = 1, N - 1$, and the remaining ϕ_{0j} , ϕ_{Nj} , $\sigma \leq j \leq k$, are used as a basis.

A similar sort of trick works for the splines of degree $m \geq \sigma + 1$. For, by the work above, it suffices to show that there are splines ψ_l whose support lies in $[0, x_{m+1}]$ and which satisfy $\psi_l^{(j)}(0) = \delta_{jl}$, $0 \leq j, l \leq m$. Now, if

$$\begin{aligned} (t)_+^m &\equiv t^m, & t \geq 0, \\ &\equiv 0, & \text{otherwise;} \end{aligned}$$

then the functions $(l + 1 - x)_+^m$, $0 \leq l \leq m$ are splines with the required limited support. Since the determinant of the matrix

$$\{M_{kl}\} = \{(l + 1 - x)_+^m\}^{(k)}(0), \quad 0 \leq k, l \leq m$$

is proportional to a Vandermonde determinant, the required ψ_l exist.

Turning to the more general case, we now suppose u is a M -dimensional vector; and that there are K boundary conditions at $x = 0$:

$$(6.7) \quad \left[\left(\sum_{\lambda=0}^L b^{\kappa\lambda} D^\lambda \right) u \right] (0) = 0, \quad \kappa = 1, \dots, K;$$

$$b^{\kappa\lambda} = \{b_{\mu\nu}^{\kappa\lambda}\}, \quad 0 \leq \lambda \leq L, 1 \leq \mu, \nu \leq M.$$

Here L would be $\sigma - 1$ in general. Let ϕ_ι be the set of $L + 1$ linearly independent scalar functions satisfying (as before)

$$(D^\lambda \phi_\iota)(0) = \delta_{\iota\lambda}, \quad 0 \leq \iota, \lambda \leq L;$$

and define M unit M -vectors

$$e_\mu = \{\delta_{\nu\mu}\}, \quad 1 \leq \mu \leq M.$$

We wish to find as many linear combinations of the $(L + 1)M$ vectors, $\phi_\iota e_\mu$, as we can which satisfy (6.7) and are linearly independent. Thus, if $\psi = \sum a_{\iota\mu} \phi_\iota e_\mu$, the numbers $a_{\iota\mu}$ should satisfy the K vector equations

$$0 = \sum_{\lambda, \mu, \iota} b^{\kappa\lambda} a_{\iota\mu} [D^\lambda \phi_\iota](0) e_\mu, \quad 1 \leq \kappa \leq K.$$

Hence we are looking for all linearly independent solutions to the KM homogeneous linear equations

$$0 = \sum_{\lambda=0}^L \sum_{\nu=1}^M b_{\mu\nu}^{\kappa\lambda} a_{\lambda\nu}, \quad 1 \leq \kappa \leq K, 1 \leq \mu \leq M,$$

in the $(L + 1)M$ unknowns, $a_{\lambda\nu}$. Presumably the solutions would be found numerically (rather than analytically as we did below (6.6) where $K = M = 1$).

Local bases can also be constructed in more than one space dimension, with an integral scalar product. It is not clear how to construct them so that they do not satisfy extra restrictive boundary conditions. It is hoped that the method of splitting (see Richtmyer and Morton [10, pp. 216–217]) will provide a way of reducing multi-dimensional problems to sequences of one dimensional problems.

Los Alamos Scientific Laboratory
Los Alamos, New Mexico 87544

University of Denver
Denver, Colorado 80210

1. F. E. BROWDER, "Nonlinear equations of evolution," *Ann. of Math.*, v. 80, 1964, pp. 485-523. MR 30 #4167.
2. F. E. BROWDER, "Non-linear initial value problems," *Ann. of Math.*, v. 82, 1965, pp. 51-87. MR 34 #7923.
3. P. G. CIARLET, M. H. SCHULTZ & R. S. VARGA, "Numerical methods of high-order accuracy for nonlinear boundary value problems," *Numer. Math.*, v. 9, 1967, pp. 394-430.
4. J. D. COLE, "On a quasi-linear parabolic equation occurring in aerodynamics," *Quart. Appl. Math.*, v. 9, 1951, pp. 225-236. MR 13, 178.
5. C. DE BOOR, *The Method of Projections as Applied to the Numerical Solution of Two Point Boundary Value Problems using Cubic Splines*, Doctoral thesis, Univ. of Michigan, Ann Arbor, Mich., 1966.
6. H. O. KREISS, "Über implizite Differenzmethoden für partielle Differentialgleichungen," *Numer. Math.*, v. 5, 1963, pp. 24-47. MR 27 #5376.
7. J. L. LIONS, *Equations Differentielles Operationnelles, et Problèmes aux Limites*, Die Grundlehren der math. Wissenschaften, Bd. 111, Springer-Verlag, Berlin, 1961. MR 27 #3935.
8. G. LUMER & R. S. PHILLIPS, "Dissipative operators in a Banach space," *Pacific J. Math.*, v. 11, 1961, pp. 679-698. MR 24 #A2248.
9. G. J. MINTY, "Monotone (nonlinear) operators in Hilbert space," *Duke Math. J.*, v. 29, 1962, pp. 341-346. MR 29 #6319.
10. R. D. RICHTMYER & K. W. MORTON, *Difference Methods for Initial Value Problems*, 2nd ed., Interscience, New York, 1967. MR 36 #3515.
11. I. J. SCHOENBERG, "Contributions to the problem of approximation of equidistant data by analytic functions," Part A, *Quart. Appl. Math.*, v. 4, 1946, pp. 45-99. MR 7, 487.
12. I. J. SCHOENBERG, "On spline functions," in *Inequalities*, O. Shisha (Editor), Academic Press, New York, 1967, pp. 255-291.
13. B. SWARTZ, " $O(h^{2n+2-l})$ bounds on some spline interpolation errors," *Bull. Amer. Math. Soc.*, v. 74, 1968, pp. 1072-1078.
14. V. THOMÉÉ, "Generally unconditionally stable difference operators," *SIAM J. Numer. Anal.*, v. 4, 1967, pp. 55-69. MR 35 #3916.
15. K. YOSIDA, *Functional Analysis*, Die Grundlehren der math. Wissenschaften, Bd. 123, Springer-Verlag, Berlin, 1965. MR 31 #5054.