ON THE NUMERICAL INTEGRATION OF QUASI-LINEAR PARABOLIC DIFFERENTIAL EQUATIONS

JIM DOUGLAS, JR.

1. Introduction. The following differential equation will be considered in the region $0 \leqq x \leqq 1,\ t \geqq 0$:

$$\frac{\partial^2 u}{\partial x^2} = F(x, t, u)\frac{\partial u}{\partial t} + G(x, t, u), \qquad F \geqq m > 0. \tag{1.1}$$

Physical phenomena leading to equations of this type include heat conduction problems in which the thermal diffusivity depends on both position and temperature, certain diffusion problems, and the flow of compressible single-phase fluids through porous media.

The simplest example of (1.1) is the classical heat flow equation

$$\frac{\partial^2 u}{\partial x^2} = K\frac{\partial u}{\partial t}, \qquad K \text{ constant},$$

which describes the flow of a fluid of constant compressibility in a linear reservoir as well as the conduction of heat in a bar insulated except possibly at its ends. If either of these problems is considered in an annular region in which radial symmetry exists, then the equation

$$\frac{\partial^2 u}{\partial x^2} = Ke^{2x}\frac{\partial u}{\partial t}$$

applies, where x is to be interpreted as the logarithm of the radius.

A somewhat more complex example is furnished by the linear flow of an ideal gas. In this case,

$$\frac{\partial^2 u}{\partial x^2} = \frac{K}{2u^{1/2}}\frac{\partial u}{\partial t},$$

where $u = p^2$, p being the pressure. The effect of treating real gases rather than ideal is that the coefficient of u_t becomes more involved.

Diffusion problems involving chemical reactions often may be analyzed by studying equations of the type

$$\frac{\partial^2 u}{\partial x^2} = K\frac{\partial u}{\partial t} + g(x, t, u).$$

As each of the examples cited are special cases of the general equa-

Received November 12, 1954. Presented to the American Mathematical Society December 27, 1954.

tion (1.1), it is of considerable practical interest to obtain methods for its solution. The form of the functions F and G or the boundary conditions arising in most engineering work usually prevents the solution of (1.1) in terms of the known functions of mathematical physics.

Numerous previous papers have been published dealing with various numerical integration schemes using finite difference techniques for (1.1) or some simpler parabolic equation. Until quite recently apparently no proofs were offered to show that the solution of the difference equation converged to that of the differential equation; however, F. John [4] has presented an extensive study of initial value problems for certain quasi-linear equations in the half-plane $-\infty < x < \infty$, $t > 0$, and several papers [3, 5, 6, 8] have been devoted to the study of the heat flow equation.

In each of these articles the difference equation used was of explicit type; that is, the function $u(x, t+\Delta t)$ can be expressed in the form

$$u(x, t+\Delta t) = \sum_{k=-N}^{N} c_k(x, t)\, u(x+k\Delta x, t) . \tag{1.2}$$

It is well known that such methods require the use of quite small time steps for both numerical stability and adequate convergence.

Certain implicit schemes requiring the solution of simultaneous linear equations have been proposed [2, 7, 9] that allow the computer to use larger increments in the time direction; these methods have been shown to be numerically stable, though no attempt has been made to demonstrate convergence of the methods.

It is the purpose of this report to prove convergence for one such method. The method of proof is based on the procedure of Rothe [10] in his paper on the existence of solutions of (1.1) when $F(x, t, u) = F(x, t)$. The existence of a sufficiently smooth solution of (1.1) satisfying the initial and boundary conditions will be assumed; sufficient conditions for this solution to exist can be obtained by extensions of Rothe's method.

2. Difference equation. Let the initial and boundary conditions associated with (1.1) be

$$\begin{cases} u(x, 0) = f(x) , & 0 \leqq x \leqq 1 \\ u(0, t) = g_1(t) , & t > 0 \\ u(1, t) = g_2(t) , & t > 0 . \end{cases} \tag{2.1}$$

Assume that a solution of (1.1), (2.1) exists in the closed region R: $0 \leqq x \leqq 1$, $0 \leqq t \leqq T$ such that $\partial^4 u/\partial x^4$ and $\partial^2 u/\partial t^2$ exist and are bounded in R. Moreover, assume that F and G are boundedly differentiable with respect to u.

Consider a lattice imposed on $0 \leqq x \leqq 1$, $0 \leqq t \leqq T$ with grid points $x_i = i\Delta x$, $i = 0, 1, \cdots, N$, $t_n = n\Delta t$, $n = 0, 1, \cdots, [T/\Delta t]$, with $\Delta x = 1/N$. Denote $Z(x_i, t_n)$ by Z_{in}.

The following difference equation will be studied as an approximation to (1.1), (2.1):

$$(2.2)\qquad \begin{cases} w_{i0} = u_{i0} = f(x_i) \\ \Delta^2 w_{i,n+1} = F(x_i, t_{n+1}, w_{in}) \dfrac{w_{i,n+1} - w_{in}}{\Delta t} + G(x_i, t_{n+1}, w_{in}), \quad n \geqq 0 \\ w_{0n} = u_{0n},\ w_{Nn} = u_{Nn}, \end{cases}$$

where

$$(2.3)\qquad \Delta^2 w_{in} = (w_{i+1,n} - 2w_{in} + w_{i-1,n})/(\Delta x)^2.$$

3. Truncation error equation. The truncation error at a mesh point is

$$(3.1)\qquad v_{in} = u_{in} - w_{in}.$$

In order to bound v_{in} as a function of Δt, it is necessary to develop its difference equation. It is easy to see that

$$(3.2)\qquad \begin{cases} \Delta^2 u_{i,n+1} = \dfrac{\partial^2 u_{i,n+1}}{\partial x^2} + \dfrac{1}{12} \overline{\dfrac{\partial^4 u}{\partial x^4}} (\Delta x)^2 \\ \dfrac{u_{i,n+1} - u_{in}}{\Delta t} = \dfrac{\partial u_{i,n+1}}{\partial t} - \dfrac{1}{2} \overline{\dfrac{\partial^2 u}{\partial t^2}} \Delta t \\ F(x_i, t_{n+1}, u_{in}) = F(x_i, t_{n+1}, u_{i,n+1}) - (u_{i,n+1} - u_{in}) \overline{\dfrac{\partial F}{\partial u}} \\ G(x_i, t_{n+1}, u_{in}) = G(x_i, t_{n+1}, u_{i,n+1}) - (u_{i,n+1} - u_{in}) \overline{\dfrac{\partial G}{\partial u}}, \end{cases}$$

where the barred derivatives are evaluated at intermediate argument values as called for by the mean value theorem. Substituting (3.2) into (1.1),

$$(3.3)\qquad \begin{aligned} \Delta^2 u_{i,n+1} = {} & F(x_i, t_{n+1}, u_{in}) \frac{u_{i,n+1} - u_{in}}{\Delta t} + G(x_i, t_{n+1}, u_{in}) \\ & + \left\{ \frac{1}{12} \overline{\frac{\partial^4 u}{\partial x^4}} (\Delta x)^2 + \frac{1}{2} \overline{F(x_i, t_{n+1}, u_{in}) \frac{\partial^2 u}{\partial t^2}} \Delta t \right. \\ & + \overline{\frac{\partial F}{\partial u}} \frac{(u_{i,n+1} - u_{in})^2}{\Delta t} + \frac{1}{2} \overline{\frac{\partial F}{\partial u}}\, \overline{\frac{\partial^2 u}{\partial t^2}} (u_{i,n+1} - u_{in}) \Delta t \\ & \left. + \overline{\frac{\partial G}{\partial u}} (u_{i,n+1} - u_{i,n}) \right\}. \end{aligned}$$

The assumptions on u above require the boundedness of all the derivatives appearing inside the bracket along with the ratio $(u_{i,n+1}-u_{in})/\Delta t$ in the region $0\leqq x\leqq 1$, $0\leqq t\leqq T$. Hence, in this region,

$$\Delta^2 u_{i,n+1}=F(x_i,\ t_{n+1},\ u_{in})\frac{u_{i,n+1}-u_{in}}{\Delta t}+G(x_i,\ t_{n+1},\ u_{in})+g_{in}\ , \tag{3.4}$$

with

$$g_{in}=O((\Delta x)^2+\Delta t)\ . \tag{3.5}$$

If (2.2) is subtracted from (3.4),

$$\begin{aligned}\Delta^2 v_{i,n+1}&=F(x_i,\ t_{n+1},\ u_{in})\frac{u_{i,n+1}-u_{in}}{\Delta t}-F(x_i,t_{n+1},w_{in})\frac{w_{i,n+1}-w_{in}}{\Delta t}\\&\quad+G(x_i,\ t_{n+1},\ u_{in})-G(x_i,\ t_{n+1},\ w_{in})+g_{in}\\&=F(x_i,\ t_{n+1},\ w_{in})\frac{v_{i,n+1}-v_{i,n}}{\Delta t}+\frac{\overline{\partial F}}{\partial u}\frac{u_{i,n+1}-u_{in}}{\Delta t}v_{in}+\frac{\overline{\partial G}}{\partial u}v_{in}+g_{in}\ ,\end{aligned} \tag{3.6}$$

as

$$F(x_i,\ t_{n+1},\ u_{in})=F(x_i,\ t_{n+1},\ w_{in})+\frac{\overline{\partial F}}{\partial u}(u_{in}-w_{in})\ .$$

Thus,

$$\begin{aligned}&\Delta^2 v_{i,n+1}-\frac{1}{\Delta t}F(x_i,\ t_{n+1},\ w_{in})v_{i,n+1}\\&\quad=-\frac{1}{\Delta t}F(x_i,\ t_{n+1},\ w_{in})\{1+h_{in}\Delta t\}v_{in}+g_{in}\ ,\end{aligned} \tag{3.7}$$

where

$$h_{in}=\left(\frac{\overline{\partial F}}{\partial u}\ \frac{u_{i,n+1}-u_{in}}{\Delta t}+\frac{\overline{\partial G}}{\partial u}\right)\Big/\ F(x_i,\ t_{n+1},\ w_{in})=O(1). \tag{3.8}$$

4. Solution of truncation error equation. The following lemma on ordinary difference equations will be important in the treatment of the truncation error equation (3.7).

LEMMA. *If*

$$\begin{cases}\Delta^2 y_i-\rho_i y_i=g_i\ , & i=1,\ 2,\ \cdots,\ N-1\\ y_0=y_N=0\ ,\end{cases} \tag{4.1}$$

and $\rho_i>0$, *then*

$$\max_i\left|y_i\right|\leqq\max_i\left|\frac{g_i}{\rho_i}\right|\ . \tag{4.2}$$

Proof. Three cases will be considered. First, let $g_i \geqq 0$. Then, y has no positive maximum; for $\Delta^2 y \leqq 0$ at the maximum and

$$y_{\max} = \frac{1}{\rho}(\Delta^2 y - g) \leqq 0 .$$

Consequently, $y_i \leqq 0$. At a negative minimum, $\Delta^2 y \geqq 0$ and

$$y_{\min} = \frac{1}{\rho}(\Delta^2 y - g) \geqq -\frac{g}{\rho} .$$

Hence,

$$\left| y \right| \leqq \max \left| \frac{g}{\rho} \right| ,$$

as was desired.

Next, let $g_i \leqq 0$. A similar argument shows that $y_i \geqq 0$ and

$$y_{\max} \leqq \frac{g}{\rho} .$$

The argument is completed by decomposing g into a sum $g^+ + g^-$, where $g^+ \geqq 0$, $g^- \leqq 0$, and $g^+ g^- = 0$.

Let

$$\begin{cases} \Delta^2 y_i^+ - \rho_i y_i^+ = g_i^+ \\ \Delta^2 y_i^- - \rho_i y_i^- = g_i^- . \end{cases}$$

Then, $0 \leqq y_i^- \leqq \max | g_i/\rho_i |$, $0 \leqq -y_i^+ \leqq \max | g_i/\rho_i |$ and $y = y^+ + y^-$. As y^+ and y^- have opposite signs,

$$| y_i | = | y_i^+ + y_i^- | \leqq \max(-y_i^+, y_i^-) \leqq \max | g_j/\rho_j | .$$

Now, v_{in} vanishes for $n=0$ and for $i=0$ or N, as $w_{in} = u_{in}$ initially and on the boundary. Hence, the lemma may be applied to (3.7) to obtain

$$\text{(4.3)} \qquad \begin{cases} \max\limits_i \left| v_{i,0} \right| = 0 \\ \max\limits_i \left| v_{i,n+1} \right| \leqq (1 + A\Delta t) \max\limits_i \left| v_{in} \right| + \max\limits_i \left| \dfrac{g_{in}}{F(x_i,\ t_{n+1},\ w_{in})} \right| \Delta t , \end{cases}$$

where $A = \max\limits_{i,n} | h_{in} |$. As $F(x_i,\ t_{n+1},\ w_{in}) > m > 0$ by assumption,

$$\text{(4.4)} \qquad \max_i | v_{i,n+1} | \leqq (1 + A\Delta t) \max_i | v_{in} | + B[(\Delta t)^2 + (\Delta t)(\Delta x)^2] .$$

Let

$$\Delta x = O((\Delta t)^{\alpha}), \qquad \alpha > 0. \tag{4.5}$$

Then, for some $C > 0$,

$$\max_i |v_{i,n+1}| \leqq (1 + C\Delta t) \max_i |v_{in}| + C(\Delta t)^{\beta}, \tag{4.6}$$

where

$$\beta = \min(2, 1 + 2\alpha). \tag{4.7}$$

LEMMA. *If* $\varepsilon_0 = 0$ and $\varepsilon_{n+1} = (1 + C\Delta t)\varepsilon_n + C(\Delta t)^{\beta}$, *then*

$$\varepsilon_m \leqq C(\Delta t)^{\beta} m(1 + C\Delta t)^m. \tag{4.8}$$

Proof. The demonstration is by induction. Equation (4.8) holds for $m=0$, as $\varepsilon_0 = 0$. Also,

$$\varepsilon_{m+1} \leqq C(\Delta t)^{\beta} m(1 + C\Delta t)^{m+1} + C(\Delta t)^{\beta} < C(\Delta t)^{\beta}(m+1)(1 + C\Delta t)^{m+1},$$

as $(1 + C\Delta t)^{n+1} > 1$.

Thus,

$$\max_i |v_{im}| \leqq C(\Delta t)^{\beta} m(1 + C\Delta t)^m. \tag{4.9}$$

Now,

$$(1 + C\Delta t)^m = \left(1 + \frac{Ct_m}{m}\right)^m < e^{Ct_m}. \tag{4.10}$$

Thus, if $\gamma = \min(1, 2\alpha)$,

$$\max_i |v_{im}| \leqq Ct_m e^{Ct_m} (\Delta t)^{\gamma}. \tag{4.11}$$

The above results may be collected into the following convergence theorem.

THEOREM. *If the differential equation* (1.1) *with boundary conditions* (2.1) *possesses a solution* $u(x, t)$ *in the region* R: $0 \leqq x \leqq 1$, $0 \leqq t \leqq T$ *such that* $\partial^4 u/\partial x^4$ *and* $\partial^2 u/\partial t^2$ *exist and are bounded in this region and the coefficients* $F(x, t, u)$ *and* $G(x, t, u)$ *have bounded first derivatives with respect to* u *in the region, then the solution* w_{in} *of* (2.2) *converges to* $u(x_i, t_n)$ *in such a manner that, if* $\Delta x = A(\Delta t)^{\alpha}$ *and* $\gamma = \min(1, 2\alpha)$, *the truncation error at a point* (x_i, t_n) *in the region is less than*

$$Ct_n e^{Ct_n} (\Delta t)^{\gamma}.$$

The constant $C > 0$ *depends only on* A, *the lower bound* m *of* $F(x, t, u)$ *in* R, *and the upper bounds on* F, F_u, G_u, u_t, u_{tt}, *and* u_{xxxx} *in* R.

Note particularly the relation (4.5). In the explicit difference methods, the ratio of Δt to $(\Delta x)^2$ is usually bounded from above; in this case with $\alpha=1/2$, the ratio is bounded from below. Consequently, the number of time steps necessary to complete the numerical solution may be reduced materially.

5. Optimum choice of α. The following question may be asked: if the ratio

$$\frac{\Delta x}{(\Delta t)^{\alpha}}=\lambda \tag{5.1}$$

is considered fixed for all α, what choice of α leads to least total work to obtain the numerical solution out to a given time T with the truncation error held less than a preassigned $\varepsilon>0$ throughout the region $0\leqq x\leqq 1$, $0\leqq t\leqq T$? As the total work is the product of the number of time steps required and the number of calculations to complete one time step, it is necessary to determine the work for each step.

First, a set of Jocobi equations $A\vec{x}=\vec{y}$, where $a_{ij}=0$ if $|i-j|>1$ and $a_{ij}=1$ for $|i-j|=1$, requires $6N$ arithmetic operations (N being the number of equations) to complete the solution for $\vec{x}$ [1, p. 82]. It will be assumed that the evaluation of $F(x, t, u)$ and $G(x, t, u)$ require the same number of operations regardless of the values of x, t, and u; this certainly is the case if they are represented by polynomials. Then, the evaluation of the coefficients in the equations is some fixed multiple m of the number of equations. Thus, the work per time step is $(m+6)/\Delta x$. The number of time steps is $T/\Delta t$. Hence, the total work is

$$W=\frac{(m+6)T}{(\Delta t)(\Delta x)}=\frac{(m+6)T}{\lambda(\Delta t)^{1+\alpha}}\,. \tag{5.2}$$

The truncation error is, by (4.11), bounded throughout the region by

$$CTe^{CT}(\Delta t)^{\gamma}, \qquad \gamma=\min(1, 2\alpha). \tag{5.3}$$

Now, for fixed λ, C may be found independent of α, as $C=\max(A, B(1+\lambda^2))$. If (5.3) is required to be less than $\varepsilon>0$,

$$(\Delta t)^{\gamma}\leqq\frac{\varepsilon}{CTe^{CT}}=\delta\,, \tag{5.4}$$

or

$$\Delta t\leqq\delta^{1/\gamma}. \tag{5.5}$$

Hence,

$$W\geqq\frac{(m+6)T}{\lambda\delta^{(1+\alpha)/\gamma}}\,. \tag{5.6}$$

As $\delta < 1$ to be of any practical interest, W is minimized by minimizing the exponent $(1+\alpha)/\gamma$ of δ. For $0 < \alpha \leqq 1/2$, $(1+\alpha)/\gamma = (1+\alpha)/2\alpha$; thus $\alpha = 1/2$ gives the smallest value in this range. For $\alpha \geqq 1/2$, $(1+\alpha)/\gamma = 1+\alpha$, which is minimized again by $\alpha = 1/2$. Thus,

$$W_{\min} \geqq \frac{(m+6)T}{\lambda \delta^{3/2}}. \tag{5.7}$$

THEOREM. *The choice of α leading to the least calculation to complete the numerical solution by* (2.2) *is $\alpha = 1/2$. For this choice, the truncation error is bounded by $CTe^{CT}\Delta t$ for some $C > 0$.*

There remains the problem of determining the best λ in the ratio $\Delta x/(\Delta t)^{1/2} = \lambda$.

REFERENCES

1. G. H. Bruce, D. W. Peaceman, H. H. Rachford and J. D. Rice, *Calculations of unsteady-state gas flow through porous media*, Trans. Amer. Inst. of Mining and Metallurgical Engineers, **198** (1953), 79–91.
2. J. Crank and P. Nicolson, *A practical method for numerical evaluation of solutions of partial differential equations of the heat conduction type*, Proc. Cambridge Philos. Soc., **43** (1947), 50–67.
3. F. B. Hildebrand, *On the convergence of numerical solutions of the heat flow equation*, J. Math. Phys., **31** (1952), 35–41.
4. F. John, *On integration of parabolic equations by difference methods*, Comm. Pure App. Math., **5** (1952), 155–211.
5. M. L. Juncosa and D. Young, *On the convergence of a solution of a difference equation to a solution of the equation of diffusion*, Proc. Amer. Math. Soc., **5** (1954), 168–174.
6. ———, *On the order of convergence of solutions of a difference equation to a solution of the diffusion equation*, J. Soc. Indust. Appl. Math., **1** (1953), 111–135.
7. P. Laasonen, *Ueber eine Methode zur Lösung der Wärmeleitungsgleichung*, Acta Math., **81** (1949), 309–317.
8. W. Leutert, *On the convergence of approximate solutions of the heat equation to the exact solution*, Proc. Amer. Math. Soc., **2** (1951), 433–439.
9. G. G. O'Brien, M. A. Hyman and S. Kaplan, *A study of the numerical solution of partial differential equations*, J. Math. Phys., **29** (1951), 223–251.
10. E. Rothe, *Zweidimensionale parabolische Randwertaufgaben als Grenzfall eindimensionaler Randwertaufgaben*, Math. Ann., **102** (1930), 650–670.

HUMBLE OIL AND REFINING CO., HOUSTON, TEXAS