ENERGY INEQUALITIES FOR THE SOLUTION OF DIFFERENTIAL EQUATIONS(1)

BY

MILTON LEES

1. Introduction. In this paper, we employ energy inequality methods to investigate properties of certain finite difference approximations to mixed initial-boundary value problems for second order quasi-linear hyperbolic equations of the form

$$\frac{\partial^2 u}{\partial t^2} - a(x, t)\frac{\partial^2 u}{\partial x^2} = F\left(x, t, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial t}\right). \tag{1.1}$$

We shall consider finite difference approximations to equation (1.1) using as principal part the *implicit* finite difference operator L defined by

$$L\phi(x, t) \equiv \phi_{t\bar{t}}(x, t) - a(x, t)\phi_{x\bar{x}}(x, t) \tag{1.2}$$

where the barred subscripts denote backward difference quotients and the unbarred subscripts denote forward difference quotients.

The fundamental problem concerning such finite difference approximations to (1.1) is to show that their solutions tend with diminishing mesh size to the solution of (1.1). Actually, it is sufficient to prove that the difference equations are stable since the convergence of a finite difference scheme can be derived from its stability(2) in a way which, by now, has become standard (e.g. see Douglas [2], F. John [6] and Lax and Richtmyer [7]).

The stability of the difference schemes considered in this paper are deduced from an energy inequality satisfied by the difference operator L. This energy inequality is a discrete analogue of the well-known energy inequality of Friedrichs and Lewy [5] for second order hyperbolic equations. The energy inequality for L states that any function ϕ together with its first order difference quotients can be estimated in the mean square along time lines in terms of $L\phi$. It is the fact that the first difference quotients of ϕ can be estimated in terms of $L\phi$ that enables us to treat difference approximations to differential equations which have nonconstant coefficients. It also enables us to treat certain quasi-linear equations.

Received by the editors June 20, 1958 and, in revised form, December 16, 1958.

(1) This research was supported in part by the United States Air Force through the Air Force Office of the Scientific Research and Development Command under contract No. AF 18(600)-1117.

(2) One also needs to show that the difference approximation is consistent with the differential equation (e.g. see [6; 7]).

Energy methods have been employed previously to deduce the stability of two *explicit* difference approximations to initial value problems for hyperbolic equations by Courant, Friedrichs and Lewy [1] and Friedrichs [4]. The former authors derived a stability condition for an explicit difference approximation to linear second order hyperbolic equations. Friedrichs gave a procedure for the construction of a stable explicit difference scheme for symmetric hyperbolic systems. These explicit difference schemes are conditionally stable, i.e. for stability, the mesh ratio must not exceed a certain bound which depends on the position of the characteristics of the differential equation relative to the mesh of the underlying lattice. The difference schemes considered in this paper are unconditionally stable, i.e., there is no restriction on the size of the mesh ratio. This is to be expected with implicit difference schemes.

Douglas [2] and Lax and Richtmyer [7] have established conditions for the stability of a wide class of difference approximations to hyperbolic equations amenable to Fourier Analysis or separation of variables technique. Their results do not apply to equations of the form (1.1).

In the final section of this paper, we consider semi-discrete approximations to the equation (1.1), i.e., difference approximations in which only the derivatives with respect to x are replaced by difference quotients. In addition to establishing the stability and convergence of such semi-discrete approximation schemes, we show that in many interesting cases it is possible to derive an explicit error estimate. Semi-discrete approximations of this type have been investigated by Douglas [3] for parabolic equations.

2. **Preliminaries.** Let Ω denote the rectangular region $0<x<1$, $0<t\leqq t_0$, and let $\bar{\Omega}$ denote its closure. The set $B=\bar{\Omega}-\Omega$ is called the boundary of Ω. We decompose B into the three segments B^0 ($0\leqq x\leqq 1, t=0$), B^1 ($x=0, 0<t\leqq t_0$) and B^2 ($x=1$, $0<t\leqq t_0$). Note that B is not the set-theoretical boundary of Ω.

Let M and N be positive integers, and denote by $\bar{D}$ a lattice with mesh (h, k) fitted over $\bar{\Omega}$, i.e., $\bar{D}$ consists of the points of intersection of the coordinate lines

$$x = nh, \qquad n = 0, 1, \cdots, N,$$

$$t = mk, \qquad m = 0, 1, \cdots, M,$$

where $h=N^{-1}$ and $k=t_0M^{-1}$. The quantity $\lambda=kh^{-1}$ is called the mesh ratio of $\bar{D}$. The mth row of the lattice $\bar{D}$ is defined to be the set

$$\bar{R}(mk) = \{(x, t) \mid (x, t) \in \bar{D} \text{ and } t = mk\}.$$

Put

$$R_0(mk) = \bar{R}(mk) - \{(x, t) \mid (x, t) \in \bar{D}, x \neq 0 \text{ and } x \neq 1\}.$$

The interior of $\bar{D}$ is the set D defined as follows:

$$D = \bigcup_{m=2}^{M} R_0(mk).$$

Let $\partial D = \overline{D} - D$ and $\partial^i D = B^i \cap \partial D$, $(i=0, 1, 2)$.

For functions $\phi(x, t)$ defined on the lattice we employ the following notation for their forward and backward difference quotients.

$$\begin{aligned}
\phi_x(x, t) &= h^{-1}[\phi(x + h, t) - \phi(x, t)], \\
\phi_{\bar{x}}(x, t) &= h^{-1}[\phi(x, t) - \phi(x - h, t)] = \phi_x(x - h, t), \\
\phi_t(x, t) &= k^{-1}[\phi(x, t + k) - \phi(x, t)], \\
\phi_{\bar{t}}(x, t) &= k^{-1}[\phi(x, t) - \phi(x, t - k)] = \phi_t(x, t - k).
\end{aligned}$$

Difference quotients of order higher than the first are formed by repeated application of the above formulas, for example,

$$\begin{aligned}
\phi_{x\bar{x}}(x, t) &= [\phi_x(x, t)]_{\bar{x}} = [\phi_{\bar{x}}(x, t)]_x \\
&= h^{-2}[\phi(x + h, t) - 2\phi(x, t) + \phi(x - h, t)], \\
\phi_{\bar{t}\bar{t}}(x, t) &= [\phi_{\bar{t}}(x, t)]_{\bar{t}} \\
&= k^{-2}[\phi(x, t) - 2\phi(x, t - k) + \phi(x, t - 2k)].
\end{aligned}$$

We shall not use subscripts to denote partial derivatives so that no confusion between partial derivatives and partial difference quotients can arise.

We introduce a hyperbolic differential operator M defined as follows:

$$Mu \equiv \frac{\partial^2 u}{\partial t^2} - a(x, t)\frac{\partial^2 u}{\partial x^2}, \tag{2.1}$$

where $a(x, t)$ satisfies the following two conditions. There exist constants μ_i $(i=0, 1, 2)$ such that

$$0 < \mu_1 \leqq a(x, t) \leqq \mu_2 \qquad (x, t) \in \overline{\Omega}, \tag{2.2}$$

$$|a(x, t) - a(x', t')| \leqq \mu[|x - x'| + |t - t'|], (x, t), (x', t') \in \overline{\Omega}. \tag{2.3}$$

As an approximation to M we take the implicit finite difference operator L defined by

$$L\phi(x, t) \equiv \phi_{\bar{t}\bar{t}}(x, t) - a(x, t)\phi_{x\bar{x}}(x, t). \tag{2.4}$$

L is called an implicit finite difference operator for the following reason: in the equation $L\phi = 0$ in D the values of ϕ on $R_0(mk)$ are defined implicitly in terms of its values on $R_0[(m-1)k]$, $R_0[(m-2)k]$ and $\overline{R}(mk) \cap \partial D$. Thus, if ϕ is prescribed on ∂D, the solution of the equation $L\phi = 0$ in D requires the inversion of $M-2$ systems of $N-1$ algebraic equations in $N-1$ unknowns. It is readily verified that the matrix of each of these systems is of the tri-diagonal type; a matrix $A = (a_{ij})$ is tri-diagonal if $a_{ij} = 0$ for $|i-j| > 1$. It follows from (2.2) that the matrices associated with the operator L have dominant main diagonal, and according to a theorem of Taussky [8] are non-singular. Hence, for arbitrary ϕ on ∂D, the equation $L\phi=\psi$ has a unique solution ϕ on D for any ψ defined on $\overline{D}$.

It is readily verified by Taylor's theorem that the difference operator L is consistent [6; 7] with the differential operator M, i.e., for any twice continuously differentiable function u on $\overline{\Omega}$, we have

$$|Lu - Mu| \to 0 \qquad \text{as } h, k \to 0 \tag{2.5}$$

at each point of Ω.

3. **The energy inequality.** If u is any twice continuously differentiable function defined on $\overline{\Omega}$ which vanishes along $x=0$ and $x=1$, then the energy inequality of Friedrichs and Lewy [5] states that there exists a constant C depending only on μ_i $(i=0, 1, 2)$ and T such that

$$\begin{aligned} &\int_0^1 |u(x,t)|^2 dx + \int_0^1 \left|\frac{\partial u}{\partial x}(x,t)\right|^2 dx + \int_0^1 \left|\frac{\partial u}{\partial t}(x,t)\right|^2 dx \\ &\quad \leqq c\left[\int_0^1 \left|\frac{\partial u}{\partial x}(x,0)\right|^2 dx + \int_0^1 \left|\frac{\partial u}{\partial t}(x,0)\right|^2 dx + \int_0^t \int_0^1 |Mu|^2 dx dt\right]. \end{aligned} \tag{3.1}$$

In this section, we prove that the difference operator L satisfies an analogous inequality.

Before proving the energy inequality for L, we give two lemmas.

LEMMA 1. *Let $\omega(t)$ and $\rho(t)$ be non-negative functions defined on the discrete set $\Lambda=\{2k, 3k, \cdots, Mk\}$, $(k>0)$. If $C\geqq 0$, $\rho(t)$ is nondecreasing and*

$$\omega(t) \leqq \rho(t) + Ck\sum_{s=2k}^{t-k} \omega(s) \tag{3.2}$$

then

$$\omega(t) \leqq \rho(t) \exp[C(t-2k)].$$

Proof. Let t_1 be an arbitrary point of Λ; $t_1\neq 2k$. Let $\eta(t)$ be that function on Λ defined by the formula

$$\omega(t) = \eta(t)\exp[C(t-2k)],$$

and set

$$\eta(t_2) = \max_{2k\leqq t\leqq t_1} \eta(t).$$

Then by (3.2)

$$\eta(t_2)\exp[C(t_2-2k)] \leqq \rho(t_2) + Ck\eta(t_2)\sum_{s=2k}^{t_2-k} \exp[C(s-2k)].$$

Comparing areas, we see that

$$k \sum_{s=2k}^{t_2-k} \exp\,[C(s-2k)] \leqq \int_{2k}^{t_2} \exp\,[C(s-2k)]ds.$$

Hence, since ρ is nondecreasing

$$\eta(t_2) \exp\,[C(t_2-2k)] \leqq \rho(t_1) + \eta(t_2)[\exp\,[C(t_1-2k)] - 1].$$

Therefore, $\eta(t_2) \leqq \rho(t_1)$. It follows that

$$\omega(t_1) = \eta(t_1) \exp\,[C(t_1-2k)] \leqq \rho(t_1) \exp\,[C(t_1-2k)].$$

Since t_1 was chosen arbitrarily from Λ, the lemma is proved.

It is necessary that the upper limit on the sum in (3.2) be retarded. In fact, if $\rho(t) = \text{const.} = c_0$ and

$$\omega(t) = c_0 + ck \sum_{s=2k}^{t} \omega(s)$$

then

$$\omega(t) = c_0(1-ck)^{-t/k}$$

which alternates in sign unless $ck<1$.

The next lemma gives a finite difference analogue of two differential identities used in the proof of the energy inequality (3.1).

Lemma 2. *Let $\phi(x, t)$ and $a(x, t)$ be functions defined on $\overline{D}$. Then at each point $(x, t) \in D$ we have*

$$\phi\phi_{\bar{t}} = \frac{1}{2}[\phi^2]_{\bar{t}} + \frac{k}{2}\phi_{\bar{t}}^2, \tag{3.3}$$

$$\phi_{\bar{t}}a\phi_{x\bar{x}} = [\phi_{\bar{t}}a\phi_x]_{\bar{x}} - a_{\bar{x}}\phi_{\bar{x}}\phi_{\bar{t}} - \frac{\bar{a}}{2}[\phi_{\bar{x}}^2]_{\bar{t}} - \frac{\bar{a}k}{2}\phi_{\bar{x}\bar{t}}^2, \tag{3.4}$$

where $\bar{a}(x, t) = a(x-h, t)$.

Proof.

$$\begin{aligned} k[\phi^2(x, t)]_{\bar{t}} &= \phi^2(x, t) - \phi^2(x, t-k) \\ &= \phi^2(x, t) - \phi(x, t)\phi(x, t-k) + \phi(x, t)\phi(x, t-k) - \phi^2(x, t-k). \end{aligned}$$

Hence,

$$[\phi^2(x, t)]_{\bar{t}} = 2\phi(x, t)\phi_{\bar{t}}(x, t) - k\phi_{\bar{x}\bar{t}}^2$$

which proves (3.3).

Using the difference product rule

$$[\eta(x)\xi(x)]_{\bar{x}} = \eta(x)\xi_{\bar{x}}(x) + \xi(x-h)\eta_{\bar{x}}(x)$$

we obtain

$$\begin{aligned}\phi_{\bar t} a \phi_{x\bar x} &= [\phi_{\bar t} a \phi_x]_{\bar x} - [\phi_{\bar t} a]_{\bar x} \phi_{\bar x} \\ &= [\phi_{\bar t} a \phi_x]_{\bar x} - a_{\bar x} \phi_{\bar t} \phi_{\bar x} - \tilde a \phi_{\bar t \bar x} \phi_{\bar x}.\end{aligned}$$

The identity (3.4) now follows by replacing

$$\tilde a \phi_{\bar t \bar x} \phi_{\bar x}$$

by

$$\frac{\tilde a}{2} [\phi_{\bar x}^2]_{\bar t} + \frac{k \tilde a}{2} \phi_{\bar x \bar t}^2$$

using the identity (3.3).

For any function ϕ defined on $\overline{D}$ we introduce norms $\|\phi\|_{0,t}$, $^*\|\phi\|_{0,t}$ and $\|\phi\|_{1,t}$ as follows:

$$\begin{aligned} \|\phi\|^2_{0,t} &= h \sum_{n=1}^{N} |\phi(nh, t)|^2 = h \sum_{R(t)} \phi^2, \ (R(mk) = R_0(mk) \cup (1, mk)), \\ ^*\|\phi\|^2_{0,t} &= h \sum_{R'(t)} \phi^2, \\ \|\phi\|^{0}_{1,t} &= \|\phi\|^2_{0,t} + \|\phi_{\bar x}\|^2_{0,t} + \|\phi_{\bar t}\|^2_{0,t}. \end{aligned}$$

THEOREM 1 (ENERGY INEQUALITY). *Suppose that $\phi(x, t)$ is a function defined on the lattice $\overline{D}$ which vanishes on $\partial^1 D$ and $\partial^2 D$. Then there exists a constant c_0 depending only on μ_i $(i=0, 1, 2)$ and t_0 such that for all sufficiently small k*

$$\|\phi\|^2_{1,t} \leqq c_0 \left[E(\phi) + k \sum_{s=2k}^{t} \|L\phi\|^2_{0,sk} \right], \tag{3.5}$$

where

$$E(\phi) = h \sum_{R(k)} [\phi_{\bar x}^2 + \phi_{\bar t}^2].$$

Proof. Let $\hat D = \overline{D} \cup [\partial^2 D - (1, k)]$. At each point of D we have

$$\phi_{\bar t} L \phi = \phi_{\bar t} \phi_{t \bar t} - a \phi_{\bar t} \phi_{x \bar x}$$

since ϕ vanishes on $\partial^2 D$. Using the identities of Lemma 2, we obtain the identity

$$\begin{aligned}\phi_{\bar t} L \phi = \frac{1}{2} [\phi_{\bar t}^2]_{\bar t} &- [a \phi_{\bar t} \phi_{\bar x}]_{\bar x} + a_{\bar x} \phi_{\bar t} \phi_{\bar x} \\ &+ \frac{1}{2} \tilde a [\phi_{\bar x}^2]_{\bar t} + \frac{k}{2} (\phi_{t \bar t}^2 + \tilde a \phi_{\bar x \bar t}^2).\end{aligned}$$

The last term on the right is non-negative since $a \geqq \mu_1$ in $\overline{\Omega}$. Hence,

$$\phi_{\bar t} L\phi \geqq \frac{1}{2}\,[\phi_{\bar t}^2]_{\bar t} - [a\phi_{\bar t}\phi_{\bar x}]_{\bar t} + a_{\bar x}\phi_{\bar x}\phi_{\bar t} + \frac{1}{2}\,\tilde a[\phi_{\bar x}^2]_{\bar t} \tag{3.6}$$

Multiplying (3.6) by hk and summing over all lattice points of $\hat D$, we obtain

$$2hk \sum_D \phi_{\bar t} L\phi \geqq hk \sum_{\hat D} [\phi_{\bar t}^2]_{\bar t} + 2hk \sum_D a_{\bar x}\phi_{\bar t}\phi_{\bar x} + hk \sum_{\hat D} \tilde a[\phi_{\bar x}^2]_{\bar t} \tag{3.7}$$

since

$$hk \sum_{\hat D} [a\phi_{\bar t}\phi_x]_{\bar t} = k \sum_{s=2k}^{t_0} a\phi_{\bar t}\phi_x \Big|_{x=0}^{x=1} = 0.$$

Now,

$$\begin{aligned} hk \sum_{\hat D} [\phi_{\bar t}^2]_{\bar t} &= h \sum_{n=1}^{N} [\phi_{\bar t}^2(nh, t_0) - \phi_{\bar t}^2(nh, k)] \\ &= \|\phi_{\bar t}\|_{0,t_0}^2 - \|\phi_{\bar t}\|_{0,0}^2. \end{aligned} \tag{3.8}$$

Summing by parts, we obtain

$$\begin{aligned} hk \sum_{\hat D} \tilde a[\phi_{\bar x}^2]_{\bar t} &= h \sum_{x=h}^{1} k \sum_{s=2k}^{t_0} \tilde a(x, s)[\phi_{\bar x}^2]_{\bar t} \\ &= h \sum_{x=h}^{1} [\phi_{\bar x}^2(x, t_0)\tilde a(x, t_0) - \phi_{\bar x}^2(x, k)\tilde a(x, 2k)] \\ &\quad - hk \sum_{x=h}^{1} \sum_{s=2k}^{t_0} \tilde a_t(x, s)\phi_{\bar x}^2(x, s) \\ &\geqq \mu_1\|\phi_{\bar x}\|_{0,t_0}^2 - \mu_2\|\phi_{\bar x}\|_{0,k}^2 - \mu hk \sum_{\hat D - R(t_0)} \phi_{\bar x}^2. \end{aligned} \tag{3.9}$$

It follows now from (3.7), (3.8) and (3.9) that

$$\begin{aligned} \|\phi_{\bar t}\|_{0,t_0}^2 + \mu_1\|\phi_{\bar x}\|_{0,t_0}^2 \leqq{}& \|\phi_{\bar t}\|_{0,k}^2 + \mu_2\|\phi_{\bar x}\|_{0,k}^2 \\ &+ 2hk \sum_D \phi_{\bar t} L\phi + \mu hk \sum_{\hat D - R(t_0)} \phi_{\bar x}^2 \\ &+ 2\mu hk \sum_D |\phi_{\bar x}|\,|\phi_{\bar t}|. \end{aligned} \tag{3.10}$$

Since

$$2hk \sum_D |\phi_{\bar x}|\,|\phi_{\bar t}| \leqq k \sum_{s=2k}^{t_0} \{\|\phi_{\bar x}\|_{0,sk}^2 + \|\phi_{\bar t}\|_{0,sk}^2\}$$

and

$$2hk \sum_D \phi_{\bar{t}} L\phi \leqq k \sum_{s=2k}^{t_0} \left\{ \left\| \phi_{\bar{t}} \right\|^2_{0,sk} + \left\| L\phi \right\|^2_{0,sk} \right\}$$

we obtain, for all k satisfying

$$(1+\mu)k \leqq 1/2, \qquad uk \leqq \mu_1 - 1/2,$$

$$\begin{aligned} \left\|\phi_{\bar{t}}\right\|^2_{0,t} + \left\|\phi_{\bar{x}}\right\|^2_{0,t} &\leqq 2 \max(1, \mu_2) E(\phi) + 2k \sum_{s=2k}^{t} {}^*\left\| L\phi \right\|^2_{0,sk} \\ &\quad + 2(1+\mu_1)k \sum_{s=2k}^{t-k} \left\{ \left\|\phi_{\bar{x}}\right\|^2_{0,sk} + \left\|\phi_{\bar{t}}\right\|^2_{0,sk} \right\}. \end{aligned} \tag{3.11}$$

Let

$$\omega(t) = \left\|\phi_{\bar{t}}\right\|^2_{0,t} + \left\|\phi_{\bar{x}}\right\|^2_{0,t},$$

$$\rho(t) = 2 \max(1, \mu_2) E(\phi) + 2k \sum_{s=2k}^{t} {}^*\left\| L\phi \right\|^2_{0,sk},$$

$$c = 2(1+\mu_1).$$

Then (3.11) becomes

$$\omega(t) \leqq \rho(t) + ck \sum_{s=2k}^{t-k} \omega(s).$$

Applying Lemma 1, we obtain

$$\left\|\phi_{\bar{t}}\right\|^2_{0,t} + \left\|\phi_{\bar{x}}\right\|^2_{0,t} \leqq (c_0/2)\left[E(\phi) + k \sum_{s=2k}^{t} {}^*\left\| L\phi \right\|^2_{0,sk}\right]$$

where

$$c_0 = 4 \max(1, \mu_2) \exp\left[2(1+\mu_1)(t_0 - 2k)\right].$$

Since

$$\left\|\phi\right\|^2_{0,t} \leqq \left\|\phi_{\bar{x}}\right\|^2_{0,t}$$

we obtain

$$\left\|\phi\right\|^2_{1,t} \leqq c_0 \left[E(\phi) + k \sum_{s=2k}^{t} {}^*\left\| L\phi \right\|^2_{0,sk} \right].$$

This completes the proof of the theorem.

It is readily verified that an analogous energy inequality holds for a corresponding difference approximation to the hyperbolic equation

$$\frac{\partial^2 u}{\partial t^2} - \sum_{i=1}^{\nu} a_i(x_1, \cdots, x_\nu, t) \frac{\partial^2 u}{\partial x_i^2} = 0$$

in ν dimensions.

As an immediate consequence of Theorem 1, we conclude that the implicit difference operator L is stable (in the mean square norm) in the sense of [2; 6; 7]. Note that in the proof of Theorem 1 no relation between the mesh widths h and k was assumed. This means that the difference operator L is stable for all values of the mesh ratio $\lambda = kh^{-1}$.

Another immediate consequence of Theorem 1 is the following approximation theorem [2; 6; 7].

THEOREM 2. *Let $u(x, t)$ be a solution of the hyperbolic equation $Mu = F(x, t)$ in Ω which satisfies the initial and boundary conditions*

$$u(x, 0) = f(x), \qquad \partial u/\partial t(x, 0) = g(x), \qquad u(0, t) = h_1(t), \qquad u(1, t) = h_2(t).$$

Suppose that $\partial^2 u/\partial t^2$ and $\partial^2 u/\partial x^2$ are uniformly continuous in $\bar{\Omega}$. Given $\epsilon > 0$, there exists a pair of integers N and M and a lattice $\bar{D}$ such that

$$\max_{\bar{D}} |u - \phi| < \epsilon,$$

where ϕ is that unique solution of $L\phi = F(x, t)$ in D such that $\phi = u$ on $\partial D \cap B$ and $\phi(x, k) = f(x) + kg(x)$.

If we add to the assumptions of Theorem 2 the requirement that $\partial^4 u/\partial x^4$ and $\partial^4 u/\partial t^4$ are uniformly bounded in Ω then it follows that

$$\max_{\bar{D}} |u - \phi| = O(k + h^2).$$

Theorems 1 and 2 can be extended to quasi-linear hyperbolic equations of the form

$$\tilde{M}u \equiv \frac{\partial^2 u}{\partial t^2} - a(x, t)\frac{\partial^2 u}{\partial x^2} - F\left(x, t, u, \frac{\partial u}{\partial x}, \frac{\partial u}{\partial t}\right) = 0 \tag{3.12}$$

which are approximated by the difference equation

$$\tilde{L}\phi \equiv \phi_{\bar{t}\bar{t}} - a(x, t)\phi_{x\bar{x}} - F(x, t, \phi(x, t), \phi_{\hat{x}}(x, t), \phi_{\bar{t}}(x, t)) = 0 \tag{3.13}$$

where

$$\phi_{\hat{x}}(x, t) = \frac{1}{2}\phi_x(x, t) + \frac{1}{2}\phi_{\bar{x}}(x, t).$$

We assume that $F(x, t, \xi_1, \xi_2, \xi_3)$ is a continuous function with uniformly bounded partial derivatives with respect to its last three arguments.

If $G(x, t) = F(x, t, 0, 0, 0)$, then the energy inequality for the difference equation (3.13) becomes

$$\|\phi\|_{1,t}^2 \leqq c_0\left[E(\phi) + k\sum_{s=2k}^{t}{}^{*}\|G\|_{0,sk}^2\right], \tag{3.14}$$

where the constant c_0 now depends on upper bounds for the derivatives

$$|\partial F/\partial \xi_i|, \qquad (i = 1, 2, 3).$$

To prove (3.14) we first consider the case when F is of the form

$$F(x, t, \xi_1, \xi_2, \xi_3) = G(x, t) + \sum_{i=1}^{3} a_i(x, t)\xi_i \tag{3.15}$$

and the functions $a_i(x, t)$ are uniformly bounded on $\bar{\Omega}$. Replace $L\phi$ in the energy inequality by F. Estimating ${}^*\|F\|_{0,t}^2$ from (3.15) and applying Lemma 1, we obtain (3.14) in this special case. In the general case, we use the integral form of the mean value theorem to write (3.13) as a linear equation, i.e., since

$$\begin{aligned} &F(x, t, \xi_1, \xi_2, \xi_3) \\ &\quad = G(x, t) + \sum_{i=1}^{3} \xi_i \int_0^1 \frac{\partial F}{\partial \xi_i}(x, t, \lambda\xi_1, \lambda\xi_2, \lambda\xi_3)d\lambda. \end{aligned} \tag{3.16}$$

L can be thought of as a linear operator and (3.14) follows from the special case since the partial derivatives $|\partial F/\partial \xi_i|$, $(i=1, 2, 3)$ are uniformly bounded.

We now show that the energy inequality implies that the difference equation (3.13) has a unique solution for all sufficiently small k when ϕ is prescribed on ∂D(3). Consider first the case when F has the form (3.15). It is readily verified that the difference equation (3.13) is equivalent to a system of $(M-2)(N-2)$ linear algebraic equations in $(M-2)(N-2)$ unknowns. If $G\equiv 0$ and $\phi=0$ on ∂D, then these equations are homogeneous. It follows from (3.14) that this homogeneous system has only the trivial solution $\phi=0$ in D. Hence, its determinant is nonzero. This proves that (3.13) has a unique solution when F is of the form (3.15).

Now, consider the case when F is nonlinear. If ϕ is any function defined on $\bar{D}$, its values in D can be thought of as a vector in a Euclidean $(M-2)(N-2)$-dimensional vector space V. The norm $\|\phi\|$ of ϕ considered as a $(M-2)(N-2)$-vector will be

$$\|\phi\| = hk \sum_D \phi^2.$$

Write F in the form (3.16), and consider the resulting linear equation as defining a continuous mapping of the Euclidean space V into itself. This mapping is well defined by the result of the previous paragraph. The energy inequality shows that the disk in V with center at the origin and radius

$$\left\{t_0 c_0 \left[E(\phi) + k \sum_{s=2k}^{t_0} {}^*\|G\|_{0,sk}^2\right\}\right.^{1/2}$$

gets mapped by this transformation into itself. It follows from Brouwer's theorem that this mapping has a fixed point. This fixed point is the desired

(3) We may assume that ϕ vanishes on $\partial^1 D$ and $\partial^2 D$ since this situation can be brought about by the transformation $\psi=\phi-(1-x)\phi(1, t)-x\phi(0, t)$.

solution. The uniqueness of this solution follows immediately from (3.14).

When F is nonlinear, the solution of (3.13) usually is obtained by an iteration procedure. However, it is possible to modify the operator L in such a way that it can be solved simply by inverting the $M-2$ tridiagonal matrices associated with the difference operator L. Let us approximate the expression $\tilde{M}u$ by the expression

$$\hat{L}\phi(x, t) \equiv L\phi(x, t) - F(x, t, \phi(x, t-k), \phi_{\hat{x}}(x, t-k), \phi_{\bar{t}}(x, t-k)), \tag{3.17}$$

where $L\phi$ is given by (2.4). It is readily verified that any solution ϕ of $\hat{L}\phi=0$ satisfies the energy inequality (3.14). Since the nonlinear part of F is evaluated at $t-k$ rather than at t, the solution of $\hat{L}\phi=0$ is obtained at each time step by inverting the $M-2$ tri-diagonal matrices associated with L.

Note that the difference operators $\tilde{L}$ and $\hat{L}$ are consistent with the hyperbolic operator $\tilde{M}$ in the sense of (2.5). This is necessary in order that $\tilde{L}$ and $\hat{L}$ lead to convergent difference approximations to $\tilde{M}u=0$.

4. **Stability of a nonlinear equation.** Although the hyperbolic operator $\tilde{M}$ introduced in the previous section is nonlinear, the restrictions on the partial derivatives of F are quite severe. Consider the nonlinear equation

$$Mu(x, t) = F(x, t, u(x, t)). \tag{4.1}$$

Let us assume that F satisfies a limitation of the form

$$|F(x, t, u)| \leqq G(x, t) + \theta |u|^{\delta}, \tag{4.2}$$

where θ and δ are non-negative constants.

We approximate equation (4.1) by the difference equation

$$L\phi(x, t) = F(x, t, \phi(x, t-k)). \tag{4.3}$$

Note that the nonlinear part of (4.3) has a retarded argument in t. Our proof that (4.3) is stable requires that this argument be retarded.

Before proving the stability of (4.3), we give a generalization of Lemma 1.

Lemma 3. *Let $\omega(t)$ be a nonnegative function defined on the discrete set $\Lambda=\{2k, 3k, \cdots, Mk\}$, $(k>0)$. If $c_1, c_2 \geqq 0$ and*

$$\omega(t) \leqq c_1 + c_2 k \sum_{s=2k}^{t-k} [\omega(s)]^{\delta} \tag{4.4}$$

where $\delta>0$ and $c_1^{\delta-1}(\delta-1)c_2(M-2)k<1$, then

$$\omega(t) \leqq c_1[1 - c_1^{\delta-1}(\delta-1)c_2(t-2k)]^{1/(1-\delta)}. \tag{4.5}$$

Proof. Consider the ordinary differential equation

$$\psi'(t) = c_2[\psi(t)]^{\delta} \tag{4.6}$$

with the initial condition $\psi(2k)=c_1$. By Taylor's theorem, there exists a τ, $0<\tau<1$ such that

$$\begin{aligned}\psi(t+k)-\psi(t) &= k\psi'(t)+(k^2/2)\psi''(t-\tau k)\\ &= k\psi'(t)+(k^2/2)c_2^2\delta[\psi(t-\tau k)]^{2\delta-1}.\end{aligned}$$

Hence, over any interval in which ψ is non-negative, we have

$$\psi(t+k)-\psi(t)\geqq k\psi'(t).$$

It is readily verified that ψ is given by the right side of (4.5).

Consequently,

$$\begin{aligned}\psi(t)-\psi(2k) &= \sum_{s=2k}^{t-k}[\psi(s+k)-\psi(s)]\geqq k\sum_{s=2k}^{t-k}\psi'(s)\\ &= kc_2\sum_{s=2k}^{t-k}[\psi(s)]^{\delta}.\end{aligned}$$

Therefore,

$$\psi(t)\geqq c_1+c_2k\sum_{s=2k}^{t-k}[\psi(s)]^{\delta}. \tag{4.7}$$

It suffices to prove that $\psi(t)\geqq\omega(t)$. In the contrary case, there exists a value of $t>2k$, say $t=t_1$, such that $\omega(t_1)>\psi(t_1)$ and $\omega(t)\leqq\psi(t)$ for $2k\leqq t<t_1$. From (4.4) and (4.7), we obtain

$$0>\psi(t_1)-\omega(t_1)\geqq c_2k\sum_{s=2k}^{t_1-k}\{[\psi(s)]^{\delta}-[\omega(s)]^{\delta}\}.$$

This is impossible since $[\psi(s)]^{\delta}\geqq[\omega(s)]^{\delta}$ in the range $2k\leqq s\leqq t_1-k$. This completes the proof of the lemma.

THEOREM 3. *Let $\phi(x, t)$ satisfy the equation* (4.3) *in D and let ϕ vanish on ∂^1D and ∂^2D. Suppose that*

$$\left\{c_0\left[E(\phi)+k\sum_{s=2k}^{t_0}{}^*\|G\|_{0,sk}^2\right]\right\}^{\delta-1}(\delta-1)2c_0\theta^2(t_0-2k)<1, \tag{4.8}$$

where c_0 is the constant of Theorem 1. *Then for all sufficiently small k,*

$$\|\phi\|_{1,t}^2\leqq K[1-K^{\delta-1}(\delta-1)2\theta^2c_0(t-2k)]^{1/(1-\delta)}$$

where

$$K=c_0\left[E(\phi)+k\sum_{s=2k}^{t_0}{}^*\|G\|_{0,sk}^2\right].$$

Proof. It follows from Theorem 1 that

$$\|\phi\|_{1,t}^2 \leqq c_0\left[E(\phi) + k \sum_{s=2k}^{t} {}^*\|F\|_{0,sk}^2\right].$$

By (4.2)

$$^*\|F\|_{0,sk}^2 \leqq 2\, {}^*\|G\|_{0,sk}^2 + 2\theta^2\, {}^*\|\phi\|_{0,sk-k}^2.$$

Hence,

$$\|\phi\|_{1,t}^2 \leqq K + 2c_0\theta^2 \sum_{s=2k}^{t-k} \|\phi\|_{1,sk}^{2\delta}.$$

The theorem now follows from Lemma 3.

Note that if $0<\delta<1$, then the condition (4.8) is automatically satisfied.

5. **Approximation by ordinary differential equations.** Denote by $\overline{D}_h$ the *strip region* consisting of the coordinate lines $x=nh$, $n=0, 1, \cdots, N$; $Nh=1$. Let $D_h=\Omega\cap\overline{D}_h$. Put $\partial^i D_h=B^i\cap\overline{D}_h$, $(i=0, 1, 2)$, and $\partial D_h=\overline{D}_h-D_h$.

As an approximation to the hyperbolic operator M, we take the *semi-discrete* operator M_h defined as follows:

$$M_h\phi(x, t) \equiv \frac{\partial^2\phi}{\partial t^2}(x, t) - a(x, t)\phi_{x\bar{x}}(x, t). \tag{5.1}$$

The equation $M_h\phi=0$ in D_h is equivalent to a system of $N-1$ second order ordinary differential equations. Clearly M_h is consistent with the hyperbolic operator M.

By modifying slightly the arguments of §2, we obtain the following energy inequality for M_h.

THEOREM 4. *Let $\phi(x, t)$ be defined on $\overline{D}_h$ and let $M_h\phi$ be defined. If ϕ vanishes on $\partial^1 D_h$ and $\partial^2 D_h$, then there exists a constant c_0 depending only on μ_i $(i=0, 1, 2)$ and t_0 such that*

$$\|\phi\|_{1,t}^2 \leqq c_0\left[A(\phi) + \int_0^t {}^*\|M_h\phi\|_{0,\tau}^2 d\tau\right]$$

where

$$A(\phi) = h\sum_{n=1}^{N}\left\{\left|\frac{\partial\phi}{\partial t}(nh, 0)\right|^2 + |\phi_{\bar{x}}(nh, 0)|^2\right\}$$

and

$$\|\phi\|_{1,t}^2 = \|\phi\|_{0,t}^2 + \|\phi_{\bar{x}}\|_{0,t}^2 + \left\|\frac{\partial\phi}{\partial t}\right\|_{0,t}^2.$$

Our main interest in semi-discrete approximations is due to the fact that in many interesting cases an *explicit* error estimate can be obtained. The essential features of this error estimate can be given most simply by restricting ourselves to the system

$$L_h\phi = 0 \text{ in } D_h, \tag{5.2}$$

$$\phi(x, 0) = f(x),$$

$$\frac{\partial \phi}{\partial t}(x, 0) = g(x), \tag{5.3}$$

$$\phi(0, t) = \phi(1, t) = 0.$$

LEMMA 4. *Let $f(x) \in C^4$ and $g(x) \in C^3$ on $0 \leqq x \leqq 1$. Let $a(x, t) \in C^3$ on $\bar{\Omega}$. If ϕ satisfies* (5.2) *and* (5.3), *then there exists a constant C_3 depending only on μ_i ($i=0, 1, 2$), t_0 and the above mentioned derivatives of $f(x)$, $g(x)$ and $a(x, t)$ such that*

$$^*\left\| \phi_{xx\bar{x}} \right\|_{0,t}^2 \leqq C_3. \tag{5.4}$$

Proof. We have from Theorem 3 that

$$\left\| \phi \right\|_{1,t}^2 \leqq C_0 A(\phi). \tag{5.5}$$

Let $\psi = \partial\phi/\partial t$. Then ψ is a solution of the system

$$\frac{\partial^2 \psi}{\partial t^2} - a\psi_{x\bar{x}} = -\frac{1}{a}\frac{\partial \psi}{\partial t}\frac{\partial a}{\partial t}, \tag{5.6}$$

$$\psi(x, 0) = g(x),$$

$$\frac{\partial \psi}{\partial t}(x, 0) = a(x, 0) f_{x\bar{x}}(x), \tag{5.7}$$

$$\psi(0, t) = \psi(1, t) = 0.$$

Applying the energy inequality to the system (5.6), (5.7), we obtain

$$\left\| \psi \right\|_{1,t}^2 \leqq c_0 \left[A(\phi) + \int_0^t \sigma\, ^*\left\| \frac{\partial \psi}{\partial t} \right\|_{0,\tau}^2 d\tau \right] \tag{5.8}$$

where

$$\sigma = \underset{\bar{\Omega}}{\text{l.u.b.}} \left| \frac{1}{a}\frac{\partial a}{\partial t} \right|.$$

It follows from (5.8) that

$$\left\| \psi \right\|_{1,t}^2 \leqq c_0 A(\phi) e^{c_0 \sigma t_0}.$$

Repeating this procedure, we conclude that $\left\| \partial^3\phi/\partial t^3 \right\|_{1,t}$ is uniformly bounded by a constant which is computable from f, g, a and their derivatives. Using this result and equation (5.2) we obtain the existence of a constant C_3 satisfying (5.4).

THEOREM 5. *Let $u(x, t)$ be a solution of the system: $Mu=0$ in Ω and $u(x, 0)=f(x)$, $\partial u(x, 0)/\partial t=g(x)$, $u(0, t)=u(1, t)=0$. Let the assumptions of Lemma 4 hold. Suppose that $\phi^{(p)}(x, t)$ is the corresponding solution of (5.2), (5.3) over the region $\overline{D}_{h_p}$, where $h_p=2^{-p}$. Then*

$$\max_{\overline{D}_{h_p}} \left| u - \phi^{(p)} \right| \leqq \frac{(C_0t_0)^{1/2}\mu_2 C_3}{2} h_{p+1}^2.$$

Proof. Set $\overline{D}_p=\overline{D}_{h_p}$. Set $w=\phi^{(p)}-\phi^{(p+1)}$. Since difference quotients of $\phi^{(p)}$ involve translations which are integer multiples of h_p, we indicate this by writing $\phi^{(p)}_{x,p}$, $\phi^{(p)}_{x\bar{x},p}$ etc. With this notational convention the expression $\phi^{(p+s)}_{x\bar{x},p}$ makes sense if $s\geqq 0$. Now,

$$\frac{\partial^2 w}{\partial t^2} - a w_{x\bar{x},p} = a\left[\phi^{(p+1)}_{x\bar{x},p} - \phi^{(p+1)}_{x\bar{x},p+1}\right]. \tag{5.9}$$

We have the identity

$$\phi^{(p+1)}_{x\bar{x},p} - \phi^{(p+1)}_{x\bar{x},p+1} = \frac{h_{p+1}^2}{4}\phi^{(p+1)}_{xx\bar{x}\bar{x},p+1}. \tag{5.10}$$

Substituting (5.10) into (5.9) yields

$$\frac{\partial^2 w}{\partial t^2} - a w_{x\bar{x},p} = a\frac{h_{p+1}^2}{4}\phi^{(p+1)}_{xx\bar{x}\bar{x},p+1}. \tag{5.11}$$

Applying the energy inequality to (5.11), we obtain

$$\left\|w\right\|_{1,t}^2 \leqq c_0 \int_0^t \frac{h_{p+1}^4}{16}\mu_2^2 {}^*\left\|\phi^{(p+1)}_{xx\bar{x}\bar{x},p+1}\right\|_{0,\tau}^2 d\tau.$$

It follows from Lemma 4 that

$$\left| w \right|_p^2 \equiv \max_{\overline{D}_p} \left| \phi^{(p)} - \phi^{(p+1)} \right|^2 \leqq \frac{c_0t_0\mu_2^2h_{p+1}^4}{16}.$$

By the triangle inequality, we get

$$\begin{aligned} \left| \phi^{(p)} - \phi^{(p+1)} \right|_p &\leqq \sum_{j=p}^{s+p-1} \left| \phi^{(j)} - \phi^{(j+1)} \right|_j \\ &\leqq \frac{(c_0t_0)^{1/2}}{4}\mu_2 c_3 \sum_{j=p}^{s+p-1} h_{j+1}^2 \\ &\leqq \frac{(c_0t_0)^{1/2}\mu_2c_3}{2} h_{p+1}^2. \end{aligned} \tag{5.12}$$

It can be shown that

$$\lim_{p\to\infty} \phi^{(p)} = u.$$

Hence, letting $s \to \infty$ in (5.12), we obtain the explicit error estimate

$$\left| \phi^{(p)} - u \right|_p \leqq \frac{(c_0 t_0)^{1/2} \mu_2 c_3}{2} h_{p+1}^2.$$

References

1. R. Courant, K. O. Friedrichs and H. Lewy, *Über die partiellen Differentialgleichungen der mathematischen Physik*, Math. Ann. vol. 100 (1928) pp. 32–74.

2. J. Douglas, Jr., *On the relation between stability and convergence in the numerical solution of linear parabolic and hyperbolic differential equations*, J. Soc. Indust. Appl. Math. vol. 3 (1955) pp. 42–65.

3. ———, *On the errors in analogue solutions of heat conduction problems*, Quart. Appl. Math. vol. 14 (1956) pp. 333–335.

4. K. O. Friedrichs, *Symmetric hyperbolic linear differential equations*, Comm. Pure Appl. Math. vol. 7 (1954) pp. 345–392.

5. K. O. Friedrichs and H. Lewy, *Über die Eindeutigkeit und das Abhängigkeitsgebiet der Lsosungen beim Anfangöwertproblem linearer hyperbolischer Differentialgleichungen*, Math. Ann. vol. 98 (1927) pp. 192–204.

6. F. John, *On integration of parabolic equations by difference methods*, Comm. Pure Appl. Math. vol. 5 (1952) pp. 155–270.

7. P. D. Lax and R. D. Richtmyer, *Survey of the stability of linear finite difference equations*, Comm. Pure Appl. Math. vol. 9 (1956) pp. 267–293.

8. O. Taussky, *A recurring theorem on determinants*, Amer. Math. Monthly vol. 55 (1949) pp. 672–676.

THE INSTITUTE FOR ADVANCED STUDY,
PRINCETON, NEW JERSEY