UC Berkeley
SEMM Reports Series

Title
On Some Numerical Methods in Continuum Mechanics

Permalink
https://escholarship.org/uc/item/32z707f9

Author
Smith, Paul

Publication Date
1971-11-01

eScholarship.org Powered by the California Digital Library
University of California

REPORT NO.
UC SESM 71-21

STRUCTURES AND MATERIALS RESEARCH
DEPARTMENT OF CIVIL ENGINEERING

ON SOME NUMERICAL METHODS IN CONTINUUM MECHANICS

by

PAUL D. SMITH

NOVEMBER 1971

STRUCTURAL ENGINEERING LABORATORY
UNIVERSITY OF CALIFORNIA
BERKELEY CALIFORNIA

Structures and Materials Research
Department of Civil Engineering
Division of Structural Engineering
and Structural Mechanics

ON SOME NUMERICAL METHODS IN CONTINUUM MECHANICS

by

Paul D. Smith

Dissertation Committee

R. L. Taylor
E. L. Wilson
Associate Professors of Civil Engineering
University of California, Berkeley

B. N. Parlett
Professor of Computer Science
University of California, Berkeley

Structural Engineering Laboratory
University of California
Berkeley, California

November 1971

CONTENTS

APPENDIX

I. INITIAL-BOUNDARY-VALUE PROBLEMS

1. Introduction. In the solution of an initial-value (Cauchy) problem, the time variable plays a special role. At any stage of the process one has at hand one or more functions of some other variables, called space variables, and these functions describe an instantaneous state of the physical system. When a computing machine is used, tabular values of these functions are stored in the storage or memory units of the machine. As time goes on, the function values are replaced by ones describing later states of the system. It is convenient to think of these functions, for fixed t, as represented by an element or point in a Banach (function) space B (68)*, and to denote them by a single symbol u. Each point in the function space represents a state of the system, and the succession of states of the system as time goes on is represented by the motion of the representative point in the function space.

Let A denote a given linear operator that transforms the element u into the element Au by spatial differentiations, matrix-vector multiplications and the like. The initial-value problem is to find a one-parameter family u(t) of elements of B, where t is a real parameter, such that

* Numbers in parenthesis (1) refer to the references, an unparenthesized number refers to an equation in the section being read while I.1.2 refers to equation 2 of Section 1 of Chapter I.

(1) $du(t)/dt = Au(t)$, $0 \leq t \leq T$ and $u(0) = u_o$

where u_o is a given element of B describing the initial state of the physical system. The derivative $d(.)/dt$ is defined in an appropriate manner by e.g. Richtmyer and Morton (52).

Systems involving higher order derivatives with respect to t can be put into the above form in the usual way by introducing lower order derivatives as further unknown functions, see Section I.5.

If there are any boundary conditions in the problem, they are assumed to be linear and homogeneous and taken care of by assuming that the domain of A is restricted to those functions satisfying the boundary conditions.

2. Semigroup Theory. The theory of semigroups of linear operators had its origin in Stone's theorem on groups of unitary operators acting in a Hilbert space (1932). Stone's theorem was motivated by the time dependent solution of Schroedinger's wave equation for quantum mechanics. The study of semigroups of operators, rather than groups of operators, was undertaken by Hille in 1936. However it was not until 1948 that the applicability of the theory was fully appreciated. At that time K. Yosida applied semigroup methods to the diffusion equation. Feller, Hille and Yosida made the theory an integral part of the theory of probability and Friedrichs, Kato, Lax and Phillips applied it with profit to Cauchy problems.

The initial-value problems of mathematical physics

will be used to motivate the theory of semigroups. Recall the abstract Cauchy problem from Section I.1.

Hadamard called an initial-value problem well-set (or well-posed) if

1. There is a unique solution to the problem for some given class of initial data.
2. The solution varies continuously with the initial data.

These two requirements are eminently reasonable on physical grounds. The existence and uniqueness of the solution is an affirmation of the principle of scientific determinism, and the continuous dependence is an expression of the stability of the solution. Stability is essential in view of the impossibility of knowing the precise initial state and the consequent desirability of approximating the true solution with approximate initial data. Put in another way this means that if two initial states u_o^1 and u_o^2 are "close" the two solutions $u^1(t)$ and $u^2(t)$ should be "close" for t finite in order that the process will be called stable.*

Suppose I.1.1 is well-set in the above sense. Let $T(t)$ map the solution $u(s)$ at time s to the solution $u(t+s)$ at time $t+s$. The assumption that A does not depend on time implies that $T(t)$ is independent of s; the physical meaning of this is that the description of the underlying physical system does not vary with time. This

* In recent years a growing theory of physical problems which are not well-posed in the above sense has been developed, Lavrentiev (39). This theory is unnecessary for the physical problems considered here since only well-posed problems will be admitted.

would not be the case, for example, if viscoelasticity rather than elasticity were the physical phenomenon being described.

The solution $u(t+r)$ can be computed as $T(t+r)u_o$ or, alternatively, solve for $u(r) = T(r)u_o$, take this as initial data, and t units of time later the solution becomes $u(t+r) = T(t)(T(r)u_o)$. The uniqueness of the solution implies the semigroup property

(1) $$T(t + r) = T(t)T(r) , \quad t,r \geq 0 .$$

Also,

(2) $$T(0) = I = \text{the identity operator,}$$

this means that the initial condition is assumed; the mapping $t \to T(t)u_o$ is differentiable on $[0,\infty)$; and $d(T(t)u_o)/dt = AT(t)u_o$ so that $u(t) = T(t)u_o$ solves I.1.1; and each $T(t)$ is a bounded linear operator (this reflects the continuous dependence of $u(t)$ on u_o). The initial data u_o should belong to the domain of A, which is assumed to be dense in B.

This mass of definition leads to the notion of a strongly continuous one parameter (or C_o) semigroup of linear operators on a Banach space. A family of operators T having the above properties is a C_o-semigroup of operators and A is called its infinitesimal generator. It is fortunate (or perhaps indicative of deeper results) that in the physically interesting problems of type I.1.1, A will turn out to be the infinitesimal generator of some semigroup.

The concept of a semigroup should be regarded as the suitable abstraction and extraction of the essential mathematical concepts occuring in a large class of initial-value problems, namely those in which A in I.1.1 is a linear operator independent of time. Two examples should help to clarify these concepts.

a. Consider the scalar equation

$$du(t)/dt = au(t) \quad , \quad u(0) = u_o \quad , \tag{3}$$

where u is a real-valued function of the real variable t and a is a real number. As is well-known the solution is

$$u(t) = \exp(at)u_o \tag{4}$$

where one definition of the function exp(.) is given by

$$\exp(x) = 1 + x + x^2/2 + x^3/3! + x^4/4! + \ldots \; . \tag{5}$$

b. Consider equation I.1.1 with A an n by n matrix of real numbers and u(t) a vector of n real-valued functions $u_i(t)$, i=1,...,n of the real variable t. The solution here is the same as in Example a, namely

$$u(t) = \exp(At)u_o \quad , \tag{6}$$

where exp(.) has the same definition as in 5 with the 1 taken as the identity matrix. In this case exp(At) is a matrix. The quantity exp(At) is well defined up to the case where A is a bounded linear operator (of which an n by n matrix of real numbers is a particular example), and t finite.

Each of the quantities exp(at) and exp(At) defines elements of semigroups and the sets $\{\exp(at)\}$ and $\{\exp(At)\}$ where the elements of the sets are distinguished by the

parameter $t \geq 0$, are examples of semigroups. Thus set

(7) $\quad T(t) = \exp(at) \quad$ or $\quad T(t) = \exp(At)$

and concrete examples of semigroups are obtained. It is easy to check that the definitions of a semigroup, i.e., equations 1, 2, etc., are satisfied.

The concept of a semigroup should be thought of then as the extension of the exponential function to a more general setting. Most of the difficulties arise when A in I.1.1 is a differential operator or a matrix of differential operators. Retaining the Banach space setting, A is then unbounded and the above definition of exp(At) no longer makes sense.

In Chapter II examples will be given illustrating how the notion of a semigroup enables approximate solution schemes of unexpected form to be constructed for partial differential equations.

3. Symmetric Hyperbolic Form (SHF). This dissertation deals with the numerical solution of wave propagation problems. The work is applicable to a variety of wave propagation phenomena of classical physics including; accoustics, electromagnetism, electric waves on transmission lines, magnetohydrodynamics, hydrodynamics, elastodynamics, etc. A unified discussion of these phenomena is possible because they are all governed by first order systems of partial differential equations of the symmetric hyperbolic type first systematically studied by Friedrichs (22). The systems can be written in the matrix form

(1) $$ED_t u + \sum A^j D_j u + Bu = f(x,t) ,$$

where $t \in R^1$ (time), $x = (x_1,\ldots,x_n) \in R^n$ (space), $D_t = \partial/\partial t$ and $D_j = \partial/\partial x_j$. $u = u(x,t) = (u_1(x,t),\ldots,u_m(x,t))^*$ is a real m by 1 (column) matrix (if M is a matrix M^* is its transpose) which describes the state of the medium at position x and time t. E, A^j (j=1,...,n) and B are real m by m matrices, characteristic of the medium, which have the properties

A^j, j=1,...,n are constant; if E and B are constant the medium is homogeneous,

(2) E and A^j, j=1,...,n are symmetric, and

E is positive definite.

The function $f(x,t) = (f_1(x,t),\ldots,f_m(x,t))^*$ is a prescribed function specifying the sources acting in the medium.

The wave equations are distinguished among the general first crder systems by possessing a quadratic energy density and corresponding energy conservation law in the sense of the following definitions.

The matrix E defines the quadratic form

(3) $$e = \tfrac{1}{2} u^* E u$$

which is interpreted as an energy density (energy per unit volume) in the applications. The matrices A^j define the quadratic forms

(4) $$s^j = \tfrac{1}{2}\, u^* A^j u$$

which are interpreted as the components of a Poynting vector describing the flow of energy (energy per unit area per unit time). Solutions of 1 satisfy a conservation of energy law which in differential form is

(5) $$D_t e + \sum D_j s^j + \tfrac{1}{2} u^*(B + B^*)u = u^* f \ .$$

The medium is said to be conservative if $B + B^* = 0$. In this case the energy

(6) $$\int_{R^n} e(x,t)dx = \tfrac{1}{2}\int_{R^n} u^*(x,t)\ E\ u(x,t)dx$$

is constant if $f = 0$. The medium is said to be dissipative if $B + B^*$ is positive definite, which implies the energy 6 is a nonincreasing function of t if $f = 0$.

Equations of type 1 can be written in a number of different ways, corresponding to different choices of dependent and independent variables. In most cases they will assume the special canonical form 1 only after a judicious choice of variables. In Section I.5 some equations of interest here will be exhibited in the form 1. For examples of other phenomena in this form, and more details, see Wilcox (65,66).

4. Boundary Conditions for Equations in SHF. In contrast to equations of type I.1.1, in the case where the operator A either is, or is obtained from a second order elliptic operator, the treatment of boundary conditions for equations in SHF is not so simple.

In the ordinary textbook treatment of the partial differential equations of mathematical physics, three sorts of problems are usually handled. There is Dirichlet's problem, the "pure boundary-value problem", which can be solved for the potential equation; Cauchy's problem, the "pure initial-value problem", which can be solved for the

wave equation; and finally the "mixed initial-boundary-value problem", or simply mixed problem, which again is suitable for the wave equation.

A major part of the modern theory of partial differential equations is the generalization of the first two problems. There is now a rather complete theory of boundary value problems for elliptic equations, of which Laplace's equation is the prototype. There is also a well-developed theory of the initial-value problems for equations of hyperbolic type, of which the wave equation is an example. The third classical problem, the mixed problem for the wave equation, remains relatively neglected. Garding, for instance, in the introduction to his lectures on the Cauchy problem (24) remarks that the theory he presents "is not very useful in physical problems. In fact, most of these lead to mixed boundary value problems which so far have not received an adequate mathematical treatment."

The reader should not infer from the above that the situation is hopeless, what is being said is that the theory is not in polished form and what theory there is will not be found in most textbooks. There has been a great deal written on the subject and the following list should provide an introduction; Duff (18), Friedrichs (22), Hersh (30), Kreiss (37), Lax and Phillips (41), Strang (57), Sarason (54), and Thomee (61). For the problem at hand, symmetric hyperbolic systems, a study of the above

references reveals that general conditions for the problem to be well-posed are well known.

The crux of the problem is establishing an a priori "energy" estimate. The devices for this purpose come down essentially to integration by parts; in the mixed problem, boundary terms appear at this point which are not present in the Cauchy problem. If one requires these boundary terms to be of the right sign one can still obtain the requisite inequality. Thus, for a symmetric first-order hyperbolic system, the following sufficient condition has been given by Friedrichs (22) and by Lax and Phillips (41): if the differential operator corresponds to equation I.3.1 and if at each point on the boundary of the domain there is an outward normal vector with components p_j, define $A_p = \sum p_j A^j$. Then the problem is well-posed if on the boundary the solution u is required to lie in a boundary space N satisfying two conditions:

1. For every vector u in N the quantity $\int u^* A_n u \, dx$ is non-negative.
2. N is not properly contained in any other subspace having property 1.

Condition 1 can be interpreted as a statement that for homogeneous boundary conditions the region of interest does not gain energy from its surrondings.

This sufficient condition on N is more stringent than that known to be necessary and sufficient in two variables. Hersh (30), has given necessary and sufficient conditions for the general mixed problem to be

well-posed, but these do not seem to be of interest here because of their complexity.

The more practical aspects of boundary conditions will be discussed in Chapters II and III.

5. Examples. In this section certain equations of classical physics will be exhibited in the form I.1.1 and the SHF I.3.1.

a. Wave Equation. The wave equation in three dimensions

(1) $$D_t^2 u = L\,u \quad , \text{ where } \quad L = D_1^2 + D_2^2 + D_3^2 \ ,$$

can be put in the form I.1.1 by defining $u_t = D_t u$,

(2) $$\frac{d}{dt}\begin{Bmatrix} u \\ u_t \end{Bmatrix} = \begin{bmatrix} 0 & 1 \\ L & 0 \end{bmatrix}\begin{Bmatrix} u \\ u_t \end{Bmatrix} .$$

Equation 1 can be put in the SHF I.3.1 by defining three auxiliary functions v_i,

(3) $$v_i = D_i u \ , \ i = 1,2,3 \ .$$

After differentiating these relations with respect to time the form

$$D_t w + \begin{bmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} D_1 w + \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} D_2 w + \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} D_3 w = 0$$

is obtained where $w = (u_t, v_1, v_2, v_3)^*$. This equation is of the form I.3.1 with $E = I$, $B = 0$ and $f = 0$.

There are three important changes that take place when an equation like 1 is put in the form I.3.1. First, the original independent variables are no longer solved for directly. For example, if equation 1 (in two dimen-

sions) is taken as the equation of motion of a taut membrane, then u is the displacement of the membrane from the rest position. In the SHF the velocity of u is obtained, not the original u. The displacement u could be obtained by appropriate spatial integration of the variables v_i, or easier, by time integration of the velocity. Second, the ability to specify nonzero boundary values of u is lost. Third, and perhaps most important, the seemingly innocent introduction of the auxiliary functions v_i in equation 3 actually represents a major change in viewpoint with respect to the physical process. Equations like 3 represent constitutive equations. In equations of type 1 these constitutive equations are identically satisfied. In the SHF these constitutive equations are solved, and when an approximate solution scheme is used, they are solved approximately, rather than identically satisfied. In Chapter II this will be shown to correspond to the introduction of Reissner's principle in elasticity. These three themes recur in all transformations of this type.

b. Timoshenko Beam. Timoshenko beam theory can be considered the prototype of the theory of elasticity since most important ideas have representation in both theories. The following quantities are defined

u = deflection of the centerline of the beam

du/dx = slope of the centerline of the beam .

g = slope due to bending

$g - du/dx$ = loss of slope, equal to the shear angle

x = independent variable, indicates position along beam.

There are two constitutive equations

(4) $$g - du/dx = V/kAG \quad \text{and} \quad dg/dx = M/EI$$

where A is the cross-sectional area, G the shear modulus, k a factor depending on the shape of the cross-section, EI the bending stiffness, V the shear and M the moment. There are two momentum equations

(5) $$Jg_{tt} = M_x - V \quad \text{and} \quad mu_{tt} = -V_x$$

where J is the rotary inertia, m the mass of the beam per unit length and the subscripts indicate partial differentiation. Differentiating the constitutive equations 4 with respect to time the SHF

$$\begin{bmatrix} 1/EI & 0 & 0 & 0 \\ 0 & 1/kAG & 0 & 0 \\ 0 & 0 & J & 0 \\ 0 & 0 & 0 & m \end{bmatrix} w_t + \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} w_x + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} w = 0$$

is obtained, where $w = (M, V, g_t, u_t)^*$. This equation is of the form

(6) $$Ew_t + Aw_x + Bw = 0 \ .$$

Observe that $B + B^* = 0$, which it will be recalled from Section 3, implies that the system is conservative, i.e., energy is not dissipated. This is what is expected since Timoshenko beam theory is an example of perfect elasticity. The interesting thing is that the theory of symmetric hyperbolic systems is sufficiently general to predict this.

c. Elastodynamics. The primary equations of interest here are those of infinitesimal elasticity, Sokolnikoff (56), hereafter called elasticity. Considering plane strain and Cartesian coordinates x and y there are two equations of motion

$$(7) \qquad \rho u_{tt} = T^{xx}_x + T^{xy}_y \quad \text{and} \quad \rho v_{tt} = T^{yx}_x + T^{yy}_y$$

and three constitutive equations

$$(8) \qquad \begin{aligned} T^{xx} &= (\lambda + 2\mu)u_x + \lambda v_y \\ T^{yy} &= (\lambda + 2\mu)v_y + \lambda u_x \\ T^{xy} &= \mu(u_y + v_x) \end{aligned}$$

where T^{xx}, T^{yy} and $T^{xy} = T^{yx}$ are the stresses, u and v displacements from the reference configuration in the x and y directions respectively, ρ the density, λ and μ the Lame elasticity constants, and subscripts indicate partial differentiation. To indicate how judiciously the variables must be chosen the SHF will be exhibited although an alternate form will be used in the sequel. The dilatational wave velocity and shear wave velocity are

$$(9) \qquad C_1 = ((\lambda + 2\mu)/\rho)^{\frac{1}{2}} \quad \text{and} \quad C_2 = (\mu/\rho)^{\frac{1}{2}}$$

respectively. Then defining the dimensionless velocities

$$\underline{u} = u_t/C_1 \quad \text{and} \quad \underline{v} = v_t/C_1$$

the dimensionless stresses

$$p = \tfrac{1}{2}(T^{xx}+T^{yy})/(\rho C_1^2) \;,\; q = \tfrac{1}{2}(T^{xx}-T^{yy})/(\rho C_1^2) \;,\; r = T^{xy}/(\rho C_1^2)$$

and the constants $a = (C_1/C_2)^2$ and $b = a/(a - 1)$, the SHF is (after differentiating 8 with respect to time)

$$(10) \qquad E w_t + A^x w_x + A^y w_y = 0 \;,$$

where

$$E = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & b & 0 & 0 \\ 0 & 0 & 0 & a & 0 \\ 0 & 0 & 0 & 0 & a \end{bmatrix}, \quad A^x = \begin{bmatrix} 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \end{bmatrix}, \quad A^y = \begin{bmatrix} 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and $w = (\underline{u}, \underline{v}, p, q, r)^*$. The matrix E is positive definite since a and b are positive (b is positive because a is less than 1). This form of these equations was probably first found by Clifton (14). These equations can also be put in the SHF by choosing as dependent variables the two velocities, the dilatation and rotation.

The form of these equations which will be used in all subsequent developments is

$$w_t = Aw_x + Bw_y \tag{11}$$

where

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ c & 0 & 0 & 0 & 0 \\ a & 0 & c & 0 & 0 \\ 0 & b & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & a & 0 & 0 & 0 \\ 0 & c & 0 & 0 & 0 \\ b & 0 & 0 & 0 & 0 \end{bmatrix}, \quad a = \lambda/\rho \ , \ b = \mu/\rho \ , \quad c = a + 2b$$

and $w = (\rho u_t, \rho v_t, T^{xx}, T^{yy}, T^{xy})^*$.

If the constitutive equations 8 are assumed to be satisfied identically and substituted into equations 7 the following equation results

$$w_{tt} = Aw \tag{12}$$

where w is the vector $(u,v)^*$ and

$$\rho A = \begin{bmatrix} (\lambda + 2\mu)D_x^2 + \mu D_y^2 & \mu D_x D_y \\ \mu D_x D_y & (\lambda + 2\mu)D_y^2 + \mu D_x^2 \end{bmatrix}. \tag{13}$$

Then these equations can be put in the form I.1.1 in the same way as in the first example

$$\frac{d}{dt}\begin{Bmatrix} w \\ w_t \end{Bmatrix} = \begin{bmatrix} 0 & I \\ A & 0 \end{bmatrix}\begin{Bmatrix} w \\ w_t \end{Bmatrix}. \tag{14}$$

d. Hydrodynamics. The Eulerian conservation-law form of the equations of hydrodynamics, for the case of two Cartesian variables X and Y, is

$$\partial u/\partial t + \partial f/\partial X + \partial g/\partial Y = 0 , \tag{15}$$

where

$$u = \begin{bmatrix} \rho \\ m \\ n \\ e \end{bmatrix}, \quad f = \begin{bmatrix} m \\ (m^2/\rho) + p \\ mn/\rho \\ (e + p)m/\rho \end{bmatrix}, \quad g = \begin{bmatrix} n \\ mn/\rho \\ (n^2/\rho) + p \\ (e + p)n/\rho \end{bmatrix},$$

m and n are the X and Y components, respectively, of momentum per unit volume, e is the total energy per unit volume, and p and ρ are the pressure and the density. In terms of the fluid velocity components u_t and v_t, $m = \rho u_t$ and $n = \rho v_t$; if E is the internal energy per unit volume, then $e = E + \rho(u_t^2 + v_t^2)/2$.

To linearize the system 15, $f = f(u)$ and $g = g(u)$ are replaced by Au and Bu, where the matrices $A = A(u)$ and $B = B(u)$ are the Jacobians of $f(u)$ and $g(u)$; i.e., $A_{ij} = \partial f_i/\partial u_j$, etc.

The matrices A and B are rather complicated, and it is convenient to discuss the simpler matrices A' and B' which can be obtained from an equivalent form of 15. The resulting system written for two variables is

$$\partial u'/\partial t + A'\partial u'/\partial X + B'\partial u'/\partial Y = 0 \tag{16}$$

where $u' = (\rho, u_t, v_t, p)^*$ and

$$A'=A'(u')=\begin{bmatrix} u_t & \rho & 0 & 0 \\ 0 & u_t & 0 & 1/\rho \\ 0 & 0 & u_t & 0 \\ 0 & \rho c^2 & 0 & u_t \end{bmatrix}, \quad B'=B'(u')=\begin{bmatrix} v_t & 0 & \rho & 0 \\ 0 & v_t & 0 & 0 \\ 0 & 0 & v_t & 1/\rho \\ 0 & 0 & \rho c^2 & v_t \end{bmatrix}$$

after introduction of the sound speed c, see Richtmyer and Morton (52). The linearized version of 15 can be obtained from 16 by a similarity transformation P

$$du = Pdu' \quad , \quad A = PA'P^{-1} \quad , \quad B = PB'P^{-1}.$$

If in equations 11 μ is set equal to 0, and in equations 16 assumptions comparable to that in infinitesimal elasticity are made, i.e., the velocities u_t and v_t are small, and ignoring the first of equations 16 which is balance of mass, equations 11 and 16 can be seen to agree, recalling that $T^{xx} = T^{yy} = -p$. This shows how the conservation-law form of the elasticity equations can be used to solve problems such as a fluid/solid interface. If the fluid is viscous, no modification of equations 11 is required in order to attack the fluid/solid interface problem. Of course, in the numerical implementation of problems with fluid/solid interfaces cognizance will have to be taken of the fact that the fluid theory may not admit negative pressures, i.e., tensile stresses.

II. APPROXIMATE SOLUTION TECHNIQUES

1. Introduction. Following Lax and Wendroff (43), a review of the notions of stability and accuracy will be given. To make the ideas precise a specific rather than the general problem will be considered, the linear hyperbolic system

$$u_t = A\, u_x + B\, u_y \,, \tag{1}$$

where u is a vector function of x,y and t, and A and B are symmetric matrices, assumed constant. Actually, A and B need not be symmetric in order to apply the methods given, it is only necessary that A and B be symmetrizable by the same similarity transformation. This insures that the system 1 is hyperbolic.

On occasion the operator on the right side of 1 will be abbreviated by G and the equation written in the form

$$u_t = G\, u \,, \tag{1}'$$

indicating explicitly only the dependence of u on t. Equation 1 is seen to be a specific example of the abstract initial-value problem I.1.1. At present, the problem of interest is the initial-value one, i.e., find a solution of 1 given u(0).

Explicit difference approximations to 1 of the form

$$v(t + k) = S_k\, v(t) \tag{2}$$

are considered, where v denotes an approximation to u, k is the time increment and the linear difference operator S_k is typically a weighted sum of translations

(3) $$(S_k f)(x,y) = \sum C_{jl} f(x+jh, y+lh) \ .$$

The matrix coefficients C_{jl} depend on A,B and the ratio $r = k/h$. The choice of equal mesh widths in the x and y directions is simply a matter of scaling. The difficulty is to provide the best combination of accuracy, simplicity and stability by a suitable choice of and method of obtaining the weights C_{jl}.

Definition. The difference scheme 2,3 approximates the differential equation 1 with m-th order accuracy if for all smooth solutions u(t) of 1

(4) $$\| u(t+k) - S_k u(t) \| \leq O(k^{m+1})$$

i.e., if after one time step exact and approximate solutions differ only by $O(k^{m+1})$.

Definition. The difference scheme 2 is stable if its solutions are uniformly bounded in a unit time range, i.e., if there exists a constant K such that

(5) $$\| S_k^n \| \leq K$$

for all n,k satisfying $nk = 1$. The norm appearing in these definitions will be taken as the L_2-norm. For a scalar function of a scalar variable x this norm is defined by

(6) $$\| f \|^2 = (f,f) = \int |f(x)|^2 dx$$

where the integration is carried out over the entire x axis. The operation (.,.) is called a scalar product. The scalar product of two functions f,g of a single variable x is defined by

(7) $$(f,g) = \int f(x) g(x) dx \ .$$

For a scalar function u of n variables $x = (x_1,\ldots,x_n)$ the L_2-norm is defined by

$$\|u\|^2 = (u,u) = \int |u(x)|^2 dx \tag{8}$$

where $dx = dx_1 \ldots dx_n$ and the integration is again carried out over the entire space. For a m-vector function u, $u(x) = (u_1(x),\ldots,u_m(x))$ of n variables $x = (x_1,\ldots,x_n)$ the L_2-norm is defined by

$$\|u\|^2 = (u_1,u_1) + \ldots + (u_m,u_m) \tag{9}$$

where (.,.) is defined in 8. The norm $\|G\|$ of an operator G is the one induced by the norm of the functions x on which it operates

$$\|G\| = \sup_{x \neq 0} \|Gx\|/\|x\| \ . \tag{10}$$

The following is well known, Lax and Wendroff (43).

Theorem. Let u and v denote solutions of the exact differential equation 1 and the difference equation 2, respectively, having the same smooth initial values. Then

$$\|u(t) - v(t)\| \leqq O(k^m) \ , \quad t \leqq 1 \ , \tag{11}$$

for all smooth initial values if the difference scheme is stable and accurate of order m.

Accuracy and some sort of stability are necessary as well as sufficient in order that the overall error be of the order 11 according to the Lax Equivalence Theorem. The following version is from Richtmyer and Morton (52).

Theorem. Given a properly posed initial-value problem and a finite difference approximation to it that satisfies the consistency condition 4, stability is the necessary and sufficient condition for convergence.

There is a simple way of expressing the accuracy of a difference scheme; this form shall be derived for the case where the coefficients of the differential and difference equations are constant, i.e., independent of x and y. To start, observe that it suffices to verify the error estimate 11 for a dense set of solutions. These solutions can be chosen as the exponential ones; that is, prescribe u(0) as

(12) $$u(0) = \exp(i(ax + by))f ,$$

where a,b are arbitrary real numbers and f is an arbitrary vector. The corresponding solution of 1 is

(13) $$u(t) = \exp(it(aA + bB))u(0) ,$$

while the corresponding solution of 2 is

(14) $$v(k) = C(ka,kb)u(0) ,$$

where

(15) $$C(ka,kb) = \sum C_{jl}\exp(i(jka + lkb)) .$$

Comparing 13 and 14 accuracy of order m is seen to mean

(16) $$\exp(i(aA + bB)) = C(ka,kb) + O(\|d\|^{m+1})$$

for d near zero, where d is the vector (a,b). The function C defined by 15 is called the amplification matrix of the difference operator 2.

The question of stability of constant coefficient difference schemes will now be discussed. For schemes with variable coefficients see Richtmyer and Morton (52) and Strang (59).

Denote the Fourier transformation in the space variables by T; then

$$TS_k v = C(ka,kb)Tv$$

where (a,b) denotes the dual variables. Repeated applications of the above identity gives

$$TS_k^n v = C^n(ka,kb)Tv \ .$$

Since Fourier transformation is an isometry (i.e. norm preserving) in the L_2-norm, the uniform boundedness of $S_k^n v$ is equivalent to the uniform boundedness of the matrices C^n. Thus the following has been proven:

Theorem. A difference scheme with constant coefficients is stable if and only if all powers of the associated amplification matrix are bounded, uniformly for all real values of (a,b) and all powers of the matrix.

To be able to use the above stability criterion conditions under which a family of matrices has the property that all powers of its elements are uniformly bounded need to be known. As observed by von Neumann, a necessary condition for this is that the eigenvalues of each matrix of the family be not greater than one in absolute value. This condition by itself is not sufficient; there are various additional conditions given in the literature, see (52) and (33) which together with von Neumann's condition guarantee the uniform boundedness of the set $\{C^n\}$. Necessary and sufficient conditions were given by Kreiss (33, 34) and by Buchanan (7). The next theorem is a statement of a different sufficient condition. Kreiss has shown that this theorem is a consequence of a new necessary and sufficient condition formulated by him, see (35,36). A

new proof of the criterion of Kreiss has been given by Morton in (45); the deduction of the following stability criterion from that of Kreiss is given there. A recent paper by Brenner and Thomee (70) succinctly reviews and extends many of these results.

Theorem. Suppose that the field of values of a matrix lies in the unit disk, i.e., that

(17) $$|(Cu,u)| \leq 1$$

for all unit vectors u. Then there exists a constant K depending only on the order of C such that

$$\|C^n\| \leq K , \quad n = 1,2,\ldots .$$

Remark. Since all eigenvalues of C belong to its field of values, 17 implies that C satisfies the von Neumann condition.

2. Lax-Wendroff Schemes. With the aid of condition II.1.16 some difference schemes which are accurate to second order will be constructed. The matrices A and B in II.1.1 are assumed constant. Expanding in a Taylor series near a=b=0,

$$(1)\qquad \exp(i(aA + bB)) \approx I + i(aA + bB) - \tfrac{1}{2}(aA + bB)^2$$

where the symbol $\approx$ denotes equality up to third order terms. Furthermore,

$$(2)\qquad \begin{aligned} &a \approx \sin a \ , \quad b \approx \sin b \ , \quad ab \approx \sin a \sin b \\ &\tfrac{1}{2}a^2 \approx 1 - \cos a \ , \quad \tfrac{1}{2}b^2 \approx 1 - \cos b \ . \end{aligned}$$

Substitute the congruences 2 into the right side of 1 and denote the resulting function by C(a,b):

$$(3)\qquad \begin{aligned} C(a,b) = I + i(A\sin a + B\sin b) - A^2(1 - \cos b) \\ - \tfrac{1}{2}(AB + BA)\sin a \sin b - B^2(1 - \cos b). \end{aligned}$$

By the above construction and II.1.16, C is the amplification matrix of a difference scheme which is accurate to second order.

Given the amplification matrix C of a difference operator S, the difference operator S can be recovered by replacing exp(ia) and exp(ib) in C by translations by the amount h in the x and y directions. For C in 3

$$(3_S)\qquad S=I+\tfrac{1}{2}AD_{1x}+\tfrac{1}{2}BD_{1y}+\tfrac{1}{2}A^2D_{2x}+\tfrac{1}{8}(AB+BA)D_{1x}D_{1y}+\tfrac{1}{2}B^2D_{2y}$$

is obtained, where D_1 and D_2 denote symmetric first and second differences:

$$(4)\qquad \begin{aligned} &D_{1x}f(x) = \tfrac{1}{2}(f(x + h) - f(x - h)) \approx hf_x(x) \ , \\ &D_{2x}f(x) = f(x + h) - 2f(x) + f(x - h) \approx h^2 f_{xx}(x). \end{aligned}$$

A more intuitive way of arriving at the difference operator 3_S is to write, by Taylor's theorem, the

approximation

$$v = u(t + k) \approx u + ku_t + \tfrac{1}{2}k^2 u_{tt} = (I + kD_t + \tfrac{1}{2}k^2 D_{tt})u \ . \tag{5}$$

For solutions of II.1.1, time derivatives can be expressed as space derivatives

$$\begin{aligned} D_t &= AD_x + BD_y \\ D_{tt} &= D_t(D_t) = D_t(AD_x + BD_y) \\ &= A^2 D_{xx} + (AB + BA)D_x D_y + B^2 D_{yy} \ . \end{aligned} \tag{6}$$

Substitute 6 into 5, and express the first and second space derivatives as symmetric first and second differences, respectively. The equation $v = Su$, is obtained where S is given by 3_S. The scheme 3_S is not the only nine-point scheme which is accurate to second order. The scheme

$$S' = S - \tfrac{1}{2}\left(\frac{k}{h}\right)^4 (A^2 + B^2)D_{2x}D_{2y} \tag{7}$$

S given by 3_S, has the same accuracy as S, since the added term in S' is of fourth order.

Lax and Wendroff (43) were able to prove the Theorem. The difference scheme 3_S is stable if

$$\max |e_j(A)|\, k/h = 1/\sqrt{8} \ , \quad \max |e_j(B)|\, k/h = 1/\sqrt{8} \ , \tag{8}$$

and the scheme 7 is stable if

$$\max |e_j(A^2 + B^2)|\, (k/h)^2 = 1/\sqrt{2} \ , \tag{9}$$

where $e_j(G)$ denotes the eigenvalues of the matrix G.

The higher order negative definite term added in scheme 7 is called an artificial viscosity. The effect of such terms has been investigated by several authors, in particular by Kreiss (34) in the linear case, and by

von Neumann and Richtmyer (46) and Lax and Wendroff (42) in the nonlinear case. The results of the above theorem furnish another illustration of the stabilizing effect of artficial viscosity.

It is worth noting that the difference schemes presented above have a somewhat unusual form from a certain standpoint. While the original differential equation II.1.1 has only first derivatives, the difference schemes have second differences and in the case of 7 a product of second differences.

3. Lax-Wendroff Schemes of Strang Type. Following Strang (59), this section will include in the presentation of the Strang schemes a review of the Lax-Wendroff schemes in order to gain perspective.

a. Lax-Wendroff Schemes. In this section several accurate difference schemes are presented. The most natural applications are to nonlinear initial-value problems in two space variables. For these problems, methods which are accurate to only first order are often too crude, and third order methods too complicated. The computations are thus made expensive either by the fine mesh required by a first order scheme in order to provide enough detail, or by the delicate differencing which maintains a high order of accuracy. Second order schemes are an obvious compromise. The starting point is taken as the linear hyperbolic model problem

$$u_t = Au_x + Bu_y \quad , \quad u(0) = u_o \quad , \tag{1}$$

where u is a vector function of x,y and t, and A and B are matrices, assumed constant.

Difference approximations to 1 of the form

$$v(t + k) = S_k v(t) \tag{2}$$

will be considered, where v denotes an approximation to u, k is the time increment and the linear difference operator S_k is typically a weighted sum of translations

$$(S_k f)(x,y) = \sum C_{ij} f(x + ih, y + jh) \quad . \tag{3}$$

The matrix coefficients C_{ij} depend on A,B and the ratio $r = k/h$.

To find the order of accuracy, one compares $S_k u$ with the Taylor expansion of $u(t + k)$. In particular, second order accuracy requires that for smooth solutions u,

$$S_k u = u + k u_t + \tfrac{1}{2}k^2 u_{tt} + O(k^3) . \tag{4}$$

Writing $u(t) = f$ and determining u_t and u_{tt} from the differential equation, it is required that, for all smooth functions $f = f(x,y)$,

$$S_k f \approx f + k(Af_x + Bf_y) + \tfrac{1}{2}k^2(A^2 f_{xx} + (AB+BA) f_{xy} + B^2 f_{yy}) , \tag{5}$$

where the symbol $\approx$ indicates equality up to $O(k^3)$.

First the case of one space variable, $B = 0$, will be reviewed. There the Lax-Wendroff operator

$$L_k^x = I + rAD_{1x} + \tfrac{1}{2}r^2 A^2 D_{2x} \tag{6}$$

is seen to be second order accurate, given

$$D_{1x} = \tfrac{1}{2}(f(x + h) - f(x - h)) \approx hf_x(x) ,$$
$$D_{2x} = f(x + h) - 2f(x) + f(x - h) \approx h^2 f_{xx}(x).$$

The importance of L_k^x derives not only from its accuracy (in this respect it is the optimal combination of $f(x)$, $f(x \pm h)$) but also from its stability. To test for stability one applies the difference operator to exponentials:

$$L_k^x \exp(iax)v = (I + irA\sin ah + \tfrac{1}{2}r^2A^2(1 - \cos ah))\exp(iax)v$$
$$= G_x(ah)\exp(iax)v .$$

Then stability requires that these amplification matrices G have uniformly bounded powers:

$$\|(G_x(ah))^n v\| = \text{const.} \|v\| \tag{7}$$

for all real ah, all $n > 0$, and all vectors v. Behind these definitions lies a very substantial literature as

the discussion in Sections 1 and 2 has indicated; the book of Richtmyer and Morton (52) and the resume by Lax and Wendroff (43) are recommended. Here the stability condition on the particular operator L_k^x is simply recalled. If the eigenvalues $e_j(A)$ of A satisfy

(8) $$\max |e_j(A)| \leq h/k = 1/r ,$$

then L_k^x dissipates energy:

$$\int |L_k^x f(x)|^2 dx \leq \int |f(x)|^2 dx .$$

This is equivalent to 7 holding with the constant equal to one; this property is called strong stability.

The physical basis for this restriction 8 on the ratio r was pointed out by Courant, Friedrichs and Lewy in their famous 1928 paper (16). The left side, $\max |e_j(A)|$ represents the maximum signal speed in the differential equation $u_t = Au_x$, and the right side is the corresponding velocity in the difference equation $v(t + k) = L_k^x v(t)$. If 8 is violated, so that the flow of energy in the approximate solution cannot keep up with the true flow, then convergence must fail. And the only alternative to convergence is instability. A scheme will be called optimally stable if it is stable whenever this C-F-L condition is satisfied (i.e., the signal speed in the difference equation is at least as great as the signal speed in the differential equation). For many schemes (including all those of accuracy ≥ 3) the condition for numerical stability is more severe than this physical argument would suggest; dropping the D_{2x} term in L_k^x, for example, leaves a first

order scheme which is unstable for every $r > 0$. A stable first order scheme will be given in the sequel.

Return now to the two-dimensional system 1. Three of the schemes discussed use the same nine indices, where i and j equal -1,0 or +1, in 3. The C-F-L condition

(9) $$|e_j(A)|\,k/h \leqq 1 \quad , \quad |e_j(B)|\,k/h \leqq 1$$

relating the eigenvalues of A and B to the ratio r, is therefore common to them all.

Lax and Wendroff (43) proposed two schemes for the model problem, the simpler being

$$S_k^1 = I+r(AD_{1x}+BD_{1y})+\tfrac{1}{2}r^2(A^2D_{2x}+(AB+BA)D_{1x}D_{1y}+B^2D_{2y}) \ .$$

The subcsripts indicate the direction in which to take differences. Obviously S_k^1 satisfies 5. It is known to be stable if the ratio permitted by 9 is reduced by $\sqrt{8}$. Numerical experiments by Burstein (9,10) and Clifton (14) suggest that stability holds under weaker restrictions, although 9 itself is definitely too weak (43).

To improve the stability, Lax and Wendroff added a dissipative term, forming

$$S_k^2 = S_k^1 - \tfrac{1}{2}r^4(A^2 + B^2)D_{2x}D_{2y} \ .$$

The accuracy is unchanged, and now stability follows from

(10) $$2r^2e_j(A^2 + B^2) \leqq 1 \ .$$

Again the scheme is not optimally stable; 9 is not sufficient for stability, though 10 may be more than sufficient. Burstein's preliminary verdict was that the improved stability of S_k^2 did not compensate for its increased complexity.

b. Strang Schemes. A third method, suggested by Strang (58), is the symmetric product of the one-dimensional Lax-Wendroff operators:

$$S_k^3 = \tfrac{1}{2}(L_k^x L_k^y + L_k^y L_k^x) \quad . \tag{11}$$

The first question is whether this alternation of one-dimensional operators retains second order accuracy. This can be decided only by a computation:

$$\begin{aligned} S_k^3 f \approx\ & \tfrac{1}{2}((I + kAD_x + \tfrac{1}{2}k^2A^2D_{xx})(I + kBD_y + \tfrac{1}{2}k^2B^2D_{yy}) + \\ & (I + kBD_y + \tfrac{1}{2}k^2B^2D_{yy})(I + kAD_x + \tfrac{1}{2}k^2B^2D_{xx}))f \\ =\ & f + k(Af_x + Bf_y) + \tfrac{1}{2}k^2(A^2f_{xx} + (AB+BA)f_{xy} + B^2f_{yy}) \\ & + \tfrac{1}{4}k^3((AB^2 + B^2A)f_{xyy} + (A^2B + BA^2)f_{yxx} \\ & + k^4/8\ (A^2B^2 + B^2A^2)f_{xxyy} \\ \approx\ & f + k(Af_x + Bf_y) + \tfrac{1}{2}k^2(A^2f_{xx} + (AB + BA)f_{xy} + B^2f_{yy}) \quad , \end{aligned}$$

which is precisely the requirement 5. Optimal stability is easy to prove, since the corresponding amplification matrix is just

$$G^3 = \tfrac{1}{2}(G_x(ah)G_y(bh) + G_y(bh)G_x(ah)) \quad .$$

If 9 holds, then both L_k^x and L_k^y are strongly stable, so S_k^3 is, too:

$$\|G^3v\| \leq \tfrac{1}{2}(\|G_xG_yv\| + \|G_yG_xv\|) \leq \|v\| \quad .$$

By choosing the symmetric product, rather than simply alternating directions by means of $S_k = L_k^xL_k^y$, second order accuracy is retained. Gourlay and Morris (26) have shown how S_k^3 ought to be organized in practice for nonlinear problems as well. They estimate that S_k^1 can be applied about twice as fast as S_k^3, and since the admissible ratios r are roughly in the opposite ratio 1:2, this

leaves a rather delicate balance (in favor of S_k^3). Strang's second scheme (59), however, is intended to reduce this factor nearly to one.

To introduce Strang's second scheme, begin with

$$S_k^4 = L^x_{\frac{1}{2}k} L^y_k L^x_{\frac{1}{2}k} . \tag{12}$$

The first question again is whether this alternation of one-dimensional operators retains second order accuracy. Computing as before:

$$\begin{aligned} S_k^4 f &\approx (I+\tfrac{1}{2}kAD_x+\tfrac{1}{2}(\tfrac{1}{2}kA)^2D_{xx})(I+kBD_y+\tfrac{1}{2}(kB)^2D_{yy})(I+kAD_x+ \\ &\qquad \tfrac{1}{2}(\tfrac{1}{2}kA)^2D_{xx})f \\ &= f + k(Af_x + Bf_y) + \tfrac{1}{2}k^2(A^2f_{xx}+(AB+BA)f_{xy} + B^2f_{yy}) \\ &+ (\tfrac{1}{2}k)^3(A^3f_{xxx}+(A^2B+BA^2+2ABA)f_{xxy}+2(AB^2+B^2A)f_{xyy}) \\ &+ (\tfrac{1}{2}k)^4((A^2B^2+B^2A^2+2AB^2A)f_{xxyy}+(ABA^2+A^2BA)f_{xxxy}+\tfrac{1}{4}A^4f_{xxxx}) \\ &+ (\tfrac{1}{2}k)^5((AB^2A^2+A^2B^2A)f_{xxxyy}+\tfrac{1}{2}A^2BA^2f_{xxxxy})+(\tfrac{1}{2}k)^5\tfrac{1}{2}A^2B^2A^2f_{xxyyxx} \\ &\approx f+k(Af_x+Bf_y)+\tfrac{1}{2}k^2(A^2f_{xx}+(AB+BA)f_{xy}+B^2f_{yy}) \end{aligned}$$

which is again the requirement 5. The asymmetry of S_k^4 may actually be an asset in case u is known to vary more rapidly in one direction than in the other. By working twice as hard, with a second order scheme (i.e., taking $\Delta x = \Delta y/2$) the accuracy in the x-direction is roughly four times better. However, if $\Delta x = \Delta y$, the dispersion will be larger in the x than in the y direction.

The individual operator

$$L^x_{k/2} = I + \tfrac{1}{2}k/h\ AD_{1x} + \tfrac{1}{2}(\tfrac{1}{2}k/h)^2A^2D_{2x}$$

is strongly stable as long as every $|e_j(A)| \leq 2h/k$. Since L^y_k is strongly stable if $|e_j(B)| \leq h/k$, these conditions combine to ensure the strong stability of S_k^4. In fact,

this combination coincides with the C-F-L condition, so the scheme is optimally stable. To see this, imagine S_k^4 expanded as in 3. Because L^x is applied twice, the index i ranges from -2 to +2. Therefore the signal speed in the x-direction is raised to 2h/k in the difference equation, and consequently $|e_j(A)| = 2h/k$ is optimal. In the y-direction nothing is new.

The term optimal stability was introduced by Strang (59). Unfortunately, it is not the "most optimum" stability available. Retaining Strang's terminology as is, the writer would like to introduce the term maxi-optimal. A stable scheme is maxioptimally stable if the signal speeds in the difference equation and the differential equation are the same. Thus scheme 11 is maxi-optimal in the x and y directions, scheme 12 is maxi-optimal in the y direction, while both schemes are optimal. A measure of overall optimallity can be obtained by considering the ratio of the domain of dependence of the difference equation to that of the differential equation, (see III.1). For scheme 11 this ratio is 1.275 and for scheme 12 it is 2.55. For a scheme which is maxioptimal in all directions this ratio is 1.0. Of course, scheme 12 is truly optimal in the sense that no clever modification of it could produce a scheme stable at a larger time step unless the domain of dependence is modified.

As it stands, S_k^4 would consume perhaps half again as much time as S_k^1, at least if one adapts them both to

nonlinear systems in conservation form (see Richtmyer and Morton (52) and Gourlay and Morris (26,27)) and counts the nonlinear evaluations. A simple alteration, however, makes the comparison more nearly even. Let $v(x,y,t) = v(ih,jh,mk) = v_{ij}^m$, where v is the approximation to u. Then the two Strang schemes can be written

$$v_{ij}^{m+1} = \tfrac{1}{2}(L_k^x L_k^y + L_k^y L_k^x) v_{ij}^m \tag{13}$$

and

$$v_{ij}^{m+1} = L_{k/2}^x L_k^y L_{k/2}^x v_{ij}^m \,. \tag{14}$$

Whenever no printout is desired, the combination of $L_{k/2}^x$ at the end of one step and again at the beginning of the next can be replaced by a single operator L_k^x. By applying the operator twice and using the fact that

$$L_{k/2}^x \cdot L_{k/2}^x \longrightarrow L_k^x + O(k^3)$$

then

$$v_{ij}^{m+2} = L_{k/2}^x L_k^y L_k^x L_k^y L_{k/2}^x v_{ij}^m$$

so that six operators have been compounded into five. The order of accuracy is still two, and if this takes place most of the time, the computational complexity is reduced to nearly the minimum, with the factor $\sqrt{8}$ in favor of scheme 14 from the stability condition. The use of L_k^x returns the stability condition to 9 and the scheme is very nearly an alteration of L_k^x and L_k^y. Only at $t = 0$ and around printout do the half-steps enter; the programming for this option presents no difficulty.

Not only is scheme 14 more efficient than 13 in that three rather than four passes have to be made through the

mesh each time step, with additional savings from the compounding of operators, but only half as much storage is required. To see this, note that the computationally efficient way to handle the scheme 13 is to apply L_k^x to each point in the mesh and then L_k^y at each point in the mesh to the results of applying L_k^x . This is because $L_k^x v_{ij}^m$ is needed at points ij+1, ij-1 as well as at point ij to calculate $L_k^y L_k^x v_{ij}^m$ at the point ij. If the computations were not done in this way three times as much effort would be required. The results of $L_k^y L_k^x v_{ij}^m$ and the original data v_{ij}^m need to be saved until the operation $L_k^x L_k^y v_{ij}^m$ is performed and the result v_{ij}^{m+1} obtained. In scheme 14, however, the data v_{ij}^m need not be saved once it is operated on by $L_{k/2}^x$, hence only half the storage is required. If the bias between the x and y directions in scheme 14 is not significant then the scheme 14 has strong points recommending it over scheme 13. Numerical experiments on this point are reported in the Appendix.

Looking back at the expanded versions of $S_k^3 f$ and $S_k^4 f$ observe that again the difference scheme does not resemble the differential equation. In this case there are even sixth order differences in a scheme approximating a differential equation with only first derivatives.

c. Three Dimensional Schemes. In (27), Gourlay and Morris report the natural extension of the Strang schemes to problems in three space dimensions. Thus, for the system

(15) $$u_t = Au_x + Bu_y + Cu_z$$

the extension of 14 is

(16) $$v_{ijk}^{m+1} = L^x_{n/2}L^y_{n/2}L^z_n\, L^y_{n/2}L^x_{n/2}v_{ijk}^m$$

and, as before, the $L^x_{n/2}$ operators combine. In (58), Strang suggests that the extension of 13 would involve all six permutations of L^x_n, L^y_n and L^z_n. According to Gourlay and Morris this is not so. It is sufficient to use

(17) $$v_{ijk}^{m+1} = \tfrac{1}{2}(L^x_nL^y_nL^z_n + L^z_nL^y_nL^x_n)v_{ijk}^m .$$

A cyclical permutation could be used to insure symmetry every three time steps. Observe that the computational advantage of 16 over 17 has dropped to a factor of 6/5 compared to 4/3 in the two dimensional case and that the storage requirements per point has remained the same for the respective schemes. The computational complexity of 16 and 17 over their respective two dimensional counterparts is 5/3 and 3/2 respectively.

d. A First Order Scheme. The optimally stable first order scheme of type 13 has been given by Strang (58). The equations

(18) $$u_t = Au_x + Bu_y \quad , \quad u_t = Eu_w + Fu_z$$

are equivalent under the change of variables

$$w = x + y \ , \ z = x - y \ , \ E = A + B \ , \ F = A - B \ .$$

The scheme

(19) $$v_{ij}^{m+1} = \tfrac{1}{2}(L^w_kL^z_k + L^z_kL^w_k)v_{ij}^m$$

where

$$L_k^w = I + \tfrac{1}{2}k/h\ ED_{1w} \quad \text{and} \quad L_k^z = I + \tfrac{1}{2}k/h\ FD_{1z}$$

is optimally stable. The methods only drawback is its low accuracy. Actually, since in elasticity problems the vector v is composed of velocities and stresses, i.e., stresses will be solved with first order accuracy, this scheme compares favorably with the usual displacement methods in elasticity where displacements are solved with second order accuracy - giving first order accurate stresses.

e. Third Order Schemes. A scheme which is uniformly third order accurate in both the space variables and time has been proposed by Burstein and Mirin (11). For the problem in two space variables 1, the scheme is of the form

$$(20) \qquad v_{ij}^{m+1} = (9/8\ T_{k/3}^x T_{2k/3}^y T_{2k/3}^x T_{k/3}^y - 1/8\ T_k^x T_k^y) v_{ij}^m$$

where the third order one-dimensional Lax-Wendroff operator T_k is defined by

$$(21) \qquad T_k^x = I + k/h\ AD_{1x} + \tfrac{1}{2}(k/h)^2 A^2 D_{2x} + 1/6(k/h)^3 A^3 D_{2x} D_{1x},$$

with a similar definition for T_k^y. Another third order scheme has been given by Dunn (11)

$$(22) \qquad v_{ij}^{m+1} = (2/3(T_{k/2}^x T_k^y T_{k/2}^x + T_{k/2}^y T_k^x T_{k/2}^y) - 1/6(T_k^x T_k^y + T_k^y T_k^x)) v_{ij}^m .$$

The stability conditions for these third order schemes are quite complicated and the reader is referred to Burstein and Mirin (11) for details. The essential complication arises because of the negative weights 1/8 and 1/6.

In a recent paper by Zwas and Abarbanel (71), new third and fourth order schemes are proposed.

f. Nonlinear Problems. The Lax-Wendroff method was actually developed for nonlinear versions of 1, i.e.,

$$u_t = f_x + g_y \tag{23}$$

where f and g are nonlinear vector functions of the vector u. The equations of hydrodynamics in Section I.5 are an example. If the differentiation is carried out in 23 the equation

$$u_t = A(u)u_x + B(u)u_y \tag{24}$$

is obtained, A(u) and B(u) being the Jacobian matrices with respect to f and g respectively. One can consider the class of linear problems

$$u_t = A_o u_x + B_o u_y \tag{25}$$

where $A_o = A(u_o)$ and $B_o = B(u_o)$, u_o being the state about which the motion is linearized. There are methods which for computational efficiency eliminate the need to explicitly calculate the matrices A(u) or B(u), see Gourlay and Morris (26,27) and Richtmyer and Morton (52).

g. Conclusion. It seems the Lax-Wendroff method for equations in three independent variables was actually developed by Butler (12). Butler applied this method to a hydrodynamic problem, Burnat et al (8) to gas dynamic problems, Richardson (51) to problems in hydrodynamics, Clifton (13,14), Estrin (19) and Baltov (2) to elastic and plastic media and Bejda (4,5) to elastic/viscoplastic media. Butler's mehtod was generalized to four independent variables and applied to elastic solids and elastoplastic soils by Recker (48) and Sauerwein (55), respectively.

The method as developed by Butler is somewhat complicated and does not have near the simplicity of Strang's schemes 13 or 14. In fact, none of the authors cited above presented a scheme in a form ready for computation. In addition, the authors cited above did not concern themselves with the more practical problems of material interfaces, irregular boundaries, irregular meshes and computational efficiency. Clifton (14) was aware of the need for work in these directions.

Applying the Lax-Wendroff method to elastic media is a relatively novel approach, as the above (fairly complete) list of authors indicates. Its most striking novelty comes from the fact that the stress equations of motion and constitutive equations are (simultaneously) solved directly. It will be shown in the sequel that this is an application of Reissner's principle in elasticity.

The more common approach is to discretize the displacement equations of motion. These methods are usually of second order accuracy in displacements, hence first order accuracy in stresses. The writer has run problems where the usual displacement methods took eight times as long per time step as did the Strang scheme 13, and scheme 13 has the additional advantage of second order accuracy in stresses. Hence for equivalent accuracies the efficiency ratio would be even greater. Working with the more basic differential form of the conservation laws (which are only one step removed from the most basic integral laws) not only is esthetically pleasing but allows the Lax-Wendroff method to easily solve some problems difficult for displacement formulations, e.g., fluid/solid interfaces. The Lax-Wendroff mehtod as developed by Strang has qualities which recommend it very highly: simplicity, efficiency, accuracy, novelty.

The Lax-Wendroff method is also a bit unusual in that an explicit rather than the more common implicit method is used. Implicit schemes have long been favored in structural applications because they are usually unconditionally stable. Hence larger time steps can be used and the solution advanced to a certain point in time with less computational effort. For problems involving shocks, implicit schemes are seldom justifiable on the grounds of increased stability, because the state changes significantly from one time step to another so that there

is no motivation for using longer time intervals. However, implicit methods are useful in problems where the expected motion is a slow succession of near equilibrium states, in problems where one part of the system has a much larger wave speed than the other parts (hence, is always nearly in equilibrium) and in problems where the important effects are propagated by mechanisms considerably slower than the wave speed.

4. Application to Elasticity. Consideration will now be given to the application of the Strang schemes to the equations of elasticity. It was shown in I.5c that for plane strain the equations of elastodynamics take the form

(1) $$w_t = Aw_x + Bw_y$$

where

$$A = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ c & 0 & 0 & 0 & 0 \\ a & 0 & 0 & 0 & 0 \\ 0 & b & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & a & 0 & 0 & 0 \\ 0 & c & 0 & 0 & 0 \\ b & 0 & 0 & 0 & 0 \end{bmatrix}, \quad a = \lambda/\rho \ , \ b = \mu/\rho$$

$$c = a + 2b$$

and $w = (u_t, v_t, T^{xx}, T^{yy}, T^{xy})^*$. To use the Strang schemes the matrices A^2 and B^2 are needed

$$A^2 = \begin{bmatrix} c & 0 & 0 & 0 & 0 \\ 0 & b & 0 & 0 & 0 \\ 0 & 0 & c & 0 & 0 \\ 0 & 0 & a & 0 & 0 \\ 0 & 0 & 0 & 0 & b \end{bmatrix}, \quad B^2 = \begin{bmatrix} b & 0 & 0 & 0 & 0 \\ 0 & c & 0 & 0 & 0 \\ 0 & 0 & 0 & a & 0 \\ 0 & 0 & 0 & c & 0 \\ 0 & 0 & 0 & 0 & b \end{bmatrix}.$$

The first thing to observe about these four matrices is that they are quite sparse, in fact, each matrix has only a single entry per row. In the canonical SHF I.5.10 the matrices $(A^x)^2$ and $(A^y)^2$ have two entries in the third and fourth rows. This ultimately leads to additional computation compared to the form chosen. The second observation is that the fixed known structure of these four matrices eliminates the need for matrix multiplication. Not only is it unnecessary to multiply by zero but it is unnecessary to test for zero to skip multiplication by zero because their locations in the matrices are known. This also leads to computational savings. It is at this point that explicit difference methods achieve their

speedup over, for example, explicit finite element methods where general matrix methods are used.

The Strang scheme II.3.13 and the scheme II.3.14 when compounding of operators is used have the stability condition

$$\max\left(|e_j(A)|, |e_j(B)|\right) \leq h/k \tag{2}$$

where $e_j(A)$ and $e_j(B)$ are the eigenvalues of A and B in 1 respectively. For isotropic elasticity

$$e_j(A) = e_j(B) = C_1 \text{ and } C_2 \,. \tag{3}$$

The quantities C_1 and C_2 are the dilatational and shear wave velocities defined in I.5c. Since $C_1 \geq C_2 \geq 0$ the stability condition 2 means that the stable time step is the time it takes a dilatational wave to travel the mesh spacing h. Because of the C-F-L necessary condition this is the maximum time step at which an explicit scheme could hope to be stable.

Appropriate boundary conditions for equation 1 have the form

$$\begin{aligned} au_n + bT_n &= g_1 \\ cu_t + dT_t &= g_2 \end{aligned} \tag{4}$$

where a,b,c,d,g_1 and g_2 are given functions defined on the boundary which satisfy $ab \geq 0$ and $cd \geq 0$. Also, both a and b or c and d cannot vanish simultaneously since then the number of boundary conditions would be reduced. u_n and u_t are respectively the normal and tangential components of velocity at the boundary; T_n and T_t are the analogous components of the boundary stresses. The conditions $ab \geq 0$

and $cd \geq 0$ insure that Condition 1 in Section I.4 is met for the homogeneous problem ($g_1 = g_2 = 0$).

Consideration is now given to the computation of the solution to the difference equations at a plane boundary parallel to the x or y axis. It is assumed that the boundary coincides with a line of mesh points of a rectangular mesh which covers the region of interest. Further, the not necessary assumption of equal spacing in the x and y directions is made. In order to calculate the difference solution at mesh points on the boundary the difference method used for interior points must be modified in two respects.

First, centered differences cannot be used at boundary points for approximating spatial derivatives with respect to the coordinate normal to the boundary. In these cases backward or forward differences must be used. For example, if the region $x \geq 0$ is the region of interest, the following (forward) differences can be used for approximating the derivatives at the boundary $x = 0$,

$$(5) \quad \begin{aligned} f_x(0,y) &= (2f(h,y)-\tfrac{1}{2}f(2h,y)-3/2f(0,y))/h + O(h^3) \\ f_{xx}(0,y) &= (f(0,y)-2f(h,y) + f(2h,y))/h^2 + O(h^3) . \end{aligned}$$

The Strang schemes introduce the simplification that the derivative f_{xy} need not be computed, in contrast to the original Lax-Wendroff method.

Second, a choice must be made as to which conditions are to be satisfied at boundary points. When the Strang schemes are applied to equation 1 five equations are

obtained at each point of the mesh; one for each of the quantities u_{tij}, v_{tij}, T^{xx}_{ij}, T^{yy}_{ij} and T^{xy}_{ij} for each ij. To avoid confusion a different symbol will not be used for the velocity u_t in the differential equation 1 and the approximation to the velocity u_{tij} in the difference equations, etc., the presence or absence of the subscripts ij will indicate which quantity is under discussion. Consider the boundary conditions 4 applied to a boundary at x = 0. The boundary conditions 4 then have the form

$$a_{ij}u_{tij} + b_{ij}T^{xx}_{ij} = g_{1ij} \tag{6}$$
$$c_{ij}v_{tij} + d_{ij}T^{xy}_{ij} = g_{2ij}$$

where the velocity normal to the boundary is the x-velocity u_t, the velocity tangent to the boundary is the y-velocity v_t, the normal stress is T^{xx} and the tangential stress is T^{xy}. The Strang schemes could be used to calculate u_{tij}, v_{tij} and T^{yy}_{ij}. The boundary conditions 6 are then used to calculate the remaining unknowns T^{xx}_{ij} and T^{xy}_{ij}. Alternately, the Strang schemes could be used to calculate T^{xx}_{ij}, T^{yy}_{ij} and T^{xy}_{ij}. The boundary conditions 6 are then used to calculate the remaining unknowns u_{tij} and v_{tij}.

A more common type of boundary condition is where, for example, in 6a either a or b equals zero. That is, either normal velocity or normal stress is prescribed, not linear combinations of them. In this case there is no choice as to which variables will be solved using the difference scheme and which using the boundary conditions. The variables whose values are not specified on the

boundary are solved using the difference scheme and the remaining variables are solved using the boundary conditions.

Numerical experiments have been performed on problems with boundaries using the above method of applying boundary conditions and the calculation was found to be stable (at the optimum time step). The Lax Equivalence Theorem (see II.1) then guarantees that the method will converge. The details on the numerical experiments are reported in the Appendix.

It should be remarked that the Strang schemes achieve in a very real but perhaps unexpected way the splitting of the elasticity operator, see Yanenko (67).

5. Conservation Form and Reissner's Principle. In 1950, Reissner (49) announced a stationary variational principle for the equations of elastostatics which has both the stress equations of equilibrium and the stress-displacement gradient relations as Euler equations. In 1961, Reissner (50) noted that, "It may be added that within the framework of the calculus of variations our general variational theorem is nothing but the canonical form of the problem (17,29)". The canonical form is that from which both the principle of minimum potential energy (minimum principle for displacements) and the principle of minimum complementary energy (maximum principle for stresses) can be obtained. Canonical variational principles were first systematically studied by Friedrichs (21).

The equations of elastostatics for plane strain can put in the form

$$(1) \qquad T^{xx}_x + T^{xy}_y - f^x = 0 \quad , \quad T^{yx}_x + T^{yy}_y - f^y = 0$$

which are the equilibrium equations, and

$$(2) \qquad T^{xx} = (\lambda + 2\mu)u_x + \lambda v_y \quad , \quad T^{yy} = (\lambda + 2\mu)v_y + \lambda u_x$$
$$\text{and} \qquad T^{xy} = \mu(u_y + v_x)$$

which are the constitutive (stress-displacement gradient) equations, where T^{xx}, T^{yy} and $T^{xy} = T^{yx}$ are the stresses, u and v displacements and f^x and f^y body forces in the x and y directions respectively, λ and μ the Lame elasticity constants and subscripts indicate partial differentiation.

For the purposes here, an equivalent formulation of Reissner's principle will be used, with the aid of

Galerkin's method in a Hilbert space H, see Mikhlin and Smolitskiy (44). Recall the scalar product (.,.) introduced in SectionII.1. Galerkin's method asserts that to solve the equation

(4) $$Au - f = 0$$

it suffices to solve

(5) $$(Au - f , w) = 0$$

for every element w in H. The solution of 5 corresponds to the notion of a weak solution of 4. One soon finds that the ordinary notion of a solution of 4 is not broad enough to admit realistic solutions in many cases. Thus the concept of a solution is enlarged to include weak solutions, distributional solutions, etc. In many cases what is being done is that the notion of a solution is suitably extended so the physically realistic solutions of the integral conservation laws are admitted as candidates for the differential form of the conservation laws. The idea that all conservation laws are integral in nature has been attributed to Hilbert. As Biot (6) has observed, the concept of a weak solution corresponds to the principle of virtual work. Actually, it suffices to solve 5 for every element w of a set which is dense in H or for every element w of a set whose finite linear combinations are dense.

The equations 1 and 2 can be put in the form 4 with u in 4 equal to $(u,v,T^{xx},T^{yy},T^{xy})^*$, hence the Galerkin form of Reissner's principle is illustrated.

A slight modification shows how to obtain the same "Euler equations" 1 and 2 as in Reissner's principle as a result of minimization of a quadratic functional. The least square method asserts that to solve the equation 4 it suffices to solve

$$\min\,(Au - f, Au - f) = \min \|Au - f\|^2 \tag{6}$$

where u is expanded into its components relative to a basis w_i of H, i.e.,

$$u = \sum_{i=1}^{\infty} (u,w_i)w_i = \sum_{i=1}^{\infty} a_i w_i \tag{7}$$

where the a_i are unknown coefficients.

In the applications, the solution is sought in some finite dimensional subspace W_n of H. The Galerkin method then becomes

$$(A\sum_{i=1}^{n} a_i w_i - f,\ w_j) = 0 \quad , \quad j = 1,\ldots,n$$

where $w_1,\ldots,w_n$ span W_n. This is an equation of the form

$$A_n a = f_n \tag{9}$$

where A_n is an n by n matrix whose elements are defined by $A_n^{ij} = (Aw_i,w_j)$, a is the vector $(a_1,\ldots,a_n)^*$ and f_n the vector $((f,w_1),\ldots,(f,w_n))^*$. Proceeding similarly the least square method gives the equation

$$A_n^* A_n a = A_n^* f_n \tag{10}$$

where A_n, a and f_n have the same definition as before and A_n^* is the transpose of A_n. Comparing equations 9 and 10 the Galerkin and least square method are seen to be quite similar, the least square method requiring more computation. The least square method does have the advantage that the

matrix $A_n^* A_n$ is symmetric while in the present case the matrix A_n is not.

In a similar manner a variational formulation of the Lax-Wendroff method can be obtained. Consider the equations of elastodynamics II.4.1

(11) $$w_t = Aw_x + Bw_y = Gw \ .$$

The Galerkin method can be applied to this equation by setting

(12) $$(dw(t)/dt - Gw(t) \ , \ v) = 0$$

for all v in H. Let $v_1,\ldots,v_n,\ldots$ be a basis of H. An approximate solution of 11 will be sought in the form

(13) $$w_n(t) = \sum_{i=1}^{n} g_i(t) v_i \ ,$$

the $g_i(t)$ will be determined by the system of n linear ordinary differential equations

(14) $$(dw_n(t)/dt - Gw_n(t) \ , \ v_j) = 0 \ , \ j=1,\ldots,n$$

which are of the form

(15) $$Mdg/dt - Kg = 0 \ .$$

M is an n by n matrix defined by $M^{ij} = (v_i, v_j)$, K is an n by n matrix defined by $K^{ij} = (Gv_i, v_j)$ and g is the vector $(g_1,\ldots,g_n)^*$. The matrix M is symmetric since $(v_i, v_j) = (v_j, v_i)$ and positive definite since the v_i form a basis. In second order displacement formulations M corresponds to the mass matrix which is often diagonalized by assuming (for M only) that $(v_i, v_j) = 0$ for $i \neq j$ and making a physical assumption for the diagonal terms of M. This process can be justified using finite difference

concepts (53). In either event the system

(16) $$dg/dt - M^{-1}Kg = 0$$

can be obtained from 15. Let g^m be the approximation to $g(t)$ at the time $m\Delta t$. Then equation 16 can be approximated by

(17) $$(g^{m+1} - g^m)/\Delta t - M^{-1}Kg^{m+1} = 0$$

or

(18) $$(g^{m+1} - g^m)/\Delta t - M^{-1}Kg^m = 0 .$$

Scheme 17 is implicit and 18 explicit.

The only point of this section is to show how methods such as the finite element method (the basis functions are of special form) can be used on the Lax-Wendroff form. Difference methods formed in this way will have all the advantages (which are considerable) and disadvantages of the FEM and in all probability the explicit scheme will not be optimally stable.

6. Semigroup Schemes. A scheme more common than the Lax-Wendroff for the solutions of problems of the motion of an elastic media starts with the second order wave equation

(1) $$w_{tt} - Aw = f$$

corresponding to the displacement equations of motion for the two dimensional (plane strain) case. The operator A is defined in I.5c, $w = w(x,y,t) = (u(x,y,t),v(x,y,t))^*$ is the vector of displacements u and v in the x and y directions respectively and $f = f(x,y,t) = (f^x(x,y,t), f^y(x,y,t))^*$ is the vector forcing function per unit mass.

Using either the finite element method (Zlamal (69)) or by replacing the spatial derivatives in A by differences one can arrive at a semi-discretization of the form

(2) $$d^2w_h(t)/dt^2 - A_hw_h(t) = f_h(t)$$

a system of ordinary differential equations, where the h is used to indicate the discretization with respect to the space variables. If the FEM is used, $A_h = -M^{-1}K$, where M is the mass and K the stiffness matrices. In either case the matrix A_h will contain in its definition a restriction to the correct displacement boundary conditions (assumed fixed in time). That is, the diagonal term of A_h corresponding to the particular degree of freedom to be constrained is set equal to zero, as well as the remaining terms in the row of A_h containing this element, and the prescribed displacement is stored as an initial condition. There can be no forcing term associated with a degree of freedom having prescribed

displacements. The stress boundary conditions can be imposed by spatially integrating them appropriately and including the results in the forcing term f_h .

In practice, solutions are usually obtained by discretization with respect to all the variables. Letting w_h^m approximate $w_h(t)$ at time $m\Delta t$, the discretizations

$$(3) \qquad (w_h^{m+1} - 2w_h^m + w_h^{m-1})/\Delta t^2 - A_h w_h^{m+1} = f_h^{m+1}$$

and

$$(4) \qquad (w_h^{m+1} - 2w_h^m + w_h^{m-1})/\Delta t^2 - A_h w_h^m = f_h^m$$

are obtained, the scheme 3 being implicit and 4 explicit. The implicit scheme 3 is unconditionally stable. The stability of the explicit scheme 4 has been studied by Alterman and Karal (1). They found the stability condition

$$(5) \qquad C_1 \Delta t/\Delta x \leq (1 + (C_2/C_1)^2)^{-\frac{1}{2}}$$

where C_1 and C_2 are the dilatational and shear velocities defined in I.5c and Δx is assumed equal to Δy.

The theory of semigroups (Section I.2) can be used to obtain an explicit scheme for equation 1 of a somewhat different type. Take as the starting point the semi-discretization 2. In the usual way the second order system 2 can be put in the first order form

$$(6) \qquad dz/dt = Fz + f$$

where

$$(7) \qquad F = \begin{bmatrix} 0 & I \\ A_h & 0 \end{bmatrix}, \quad z = \begin{Bmatrix} w_h \\ \dot{w}_h \end{Bmatrix}, \quad f = \begin{Bmatrix} 0 \\ f_h \end{Bmatrix}$$

which is of the form I.1.1. Since F is a bounded linear operator (in fact a finite matrix of real numbers) the solution is well known. The homogeneous solution (f = 0)

is

$$z(t) = \exp(Ft)z(0) \tag{8}$$

where the exponential function is defined in I.2.5. One can define

$$T_h(t) = \exp(Ft) \tag{9}$$

and thus have another example of a semigroup. The h is a reminder of the dependence on the spatial discretization through A_h. In fact, $\{T_h\}$ represents a sequence of semigroups, and the proof of many of the above assertions can be obtained from the theory of convergence of semigroups, see Trotter (63), Kato (32) and Sunouchi (60).

As one writer put it, one will soon run out of patience if the attempt is made to sum the infinite series I.2.5. In practice then, one must be content with an approximation to T_h. Define

$$T_h^n(t) = I + \sum_{k=1}^{n} (Ft)^k/k! \tag{10}$$

and by taking t small hope to compensate for a lack of patience. For example, for n = 2 in (10)

$$T_h^2(t) = \begin{bmatrix} I + A_h t^2/2 & It \\ A_h & I + A_h t^2/2 \end{bmatrix}. \tag{11}$$

Recall that in terms of FEM concepts $A_h = -M^{-1}K$, M the mass and K the stiffness matrices. If n in 10 is taken larger than 2 then powers of A_h will be computed which means that the sparseness of T_h^n will be reduced (zeros will start filling in) and the bandwidth of the submatrices will increase. Another way of looking at this is to observe that this means the domain of dependence is increas-

ing (see III.1) which fits in with the fact that while 9 is valid for any finite t; $T_h^n(t)$ certainly has a t_o beyond which it is unstable. The approximate solution scheme is

(12) $$z^{m+1} = T_h^n(\Delta t)\, z^m .$$

This scheme shows how approximation schemes of arbitrarily high truncation error with respect to time are obtained in a unified manner. It should be emphasized that the solution at time m + 1 is obtained solely from the given state at time m, in contrast to the schemes 3 and 4.

Numerical experiments were performed on the scheme 12 with n = 2, for the equations of elasticity. For the case with $C_2/C_1 = \frac{1}{2}$, the scheme was found to be stable when $C_1 \Delta t/\Delta x = \frac{1}{4}$ and unstable when $C_1 \Delta t/\Delta x = \frac{1}{2}$. These numerical results do not violate the condition 5 for the scheme 4 , but they are not as sharp. The situation is similar to that in the Lax-Wendroff method where the optimum time step II.3.9 has to be reduced by the factor $\sqrt{8}$. It appears that in both the original Lax-Wendroff method and in explicit difference schemes for the displacement equations of motion the factor that reduces the stable time step below optimum is the presence of mixed derivatives coupling the x and y directions. If a splitting scheme for the displacement equations of motion could be found, this situation could possibly be improved.

Because the solution was found to be stable the Lax Equivalence Theorem (see II.1) guarantees that the method

12 will converge. More details on the numerical experiments are found in the Appendix.

The writer is not aware of this sort of solution scheme being proposed for problems of this type before.

The more intuitive method of Lax and Wendroff can be used to obtain the approximation schemes 12. Consider the equation

(13) $$dz(t)/dt = Fz(t) .$$

Expand $z(t)$ in a Taylor's series about $t = 0$

(14) $$z(t) = z(0) + tz_t(0) + \tfrac{1}{2}t^2 z_{tt}(0) + \tfrac{1}{6}t^3 z_{ttt}(0) + \ldots$$

For solutions $z(t)$ of equation 13, the time derivative of z can be obtained by computing Fz. Hence 14 becomes

$$z(t) = z(0) + tFz(0) + \tfrac{1}{2}t^2 F(Fz(0)) + \tfrac{1}{6}t^3 F(F(Fz(0))) + \ldots$$

which is the same as 8.

7. Forcing Terms. A common practical case is the inhomogeneous problem

(1) $\quad du(t)/dt = Au(t) + f(t)\ ,\ 0 \leq t \leq T\ ,\ u(0) = u_o$

corresponding to equations I.1.1 and II.1.1, where $f = f(x_1,\ldots,x_n,t)$ is called the forcing term. If $E(t)$ is the solution operator of equation I.1.1, i.e., $E(t)$ is defined by $u(t) = E(t)u_o$, then the solution of 1 is given by Duhamel's principle, see Thompson (62),

$$(2) \qquad u(t) = E(t)u_o + \int_0^t E(t - s)f(s)ds\ .$$

For example, for the Strang schemes the solution operator is defined by equations II.3.11 and II.3.12, and the complete solution of 1 is given by

$$(3) \qquad v_{ij}^{t+k} = S_k v_{ij}^t + \int_t^{t+k} S_{k-s}\, f(s)\, ds\ .$$

III. PRACTICAL CONSIDERATIONS

1. Domain of Dependence. An important concept central to the study of explicit approximate solution schemes for hyperbolic systems is that of domain of dependence (DOD). The concept of a finite DOD is peculiar to hyperbolic systems and reflects the fact that in hyperbolic systems the speed of propagation of signals is finite. This means that the solution at a given point (x,y) is dependent only on the data within the DOD of the point. All data outside the DOD of (x,t) has no effect on the solution at (x,t). An example should clarify this concept.

Consider the wave equation in one spatial dimension,

$$w_{tt} = c^2 w_{xx} \tag{1}$$

with the initial conditions

$$w(x,0) = w^0(x) \ , \quad w_t(x,0) = w^1(x) \tag{2}$$

which can be taken, for example, as the equation of an infinite string in which case w will be the displacement of the string from the rest position, and $c^2 = T/\rho$ where T is the tension in the string and ρ is the mass of the string per unit length. The functions $w^0(x)$ and $w^1(x)$ define respectively arbitrary, prescribed initial displacement and velocity of the string for $-\infty \leq x \leq \infty$.

The solution of (1), subject to (2) is well known, see Friedrichs (23),

$$w(x,t) = \tfrac{1}{2}(w^0(x+ct)+w^0(x-ct))+\frac{1}{2c}\int_{x-ct}^{x+ct} w^1(s)ds \ . \tag{3}$$

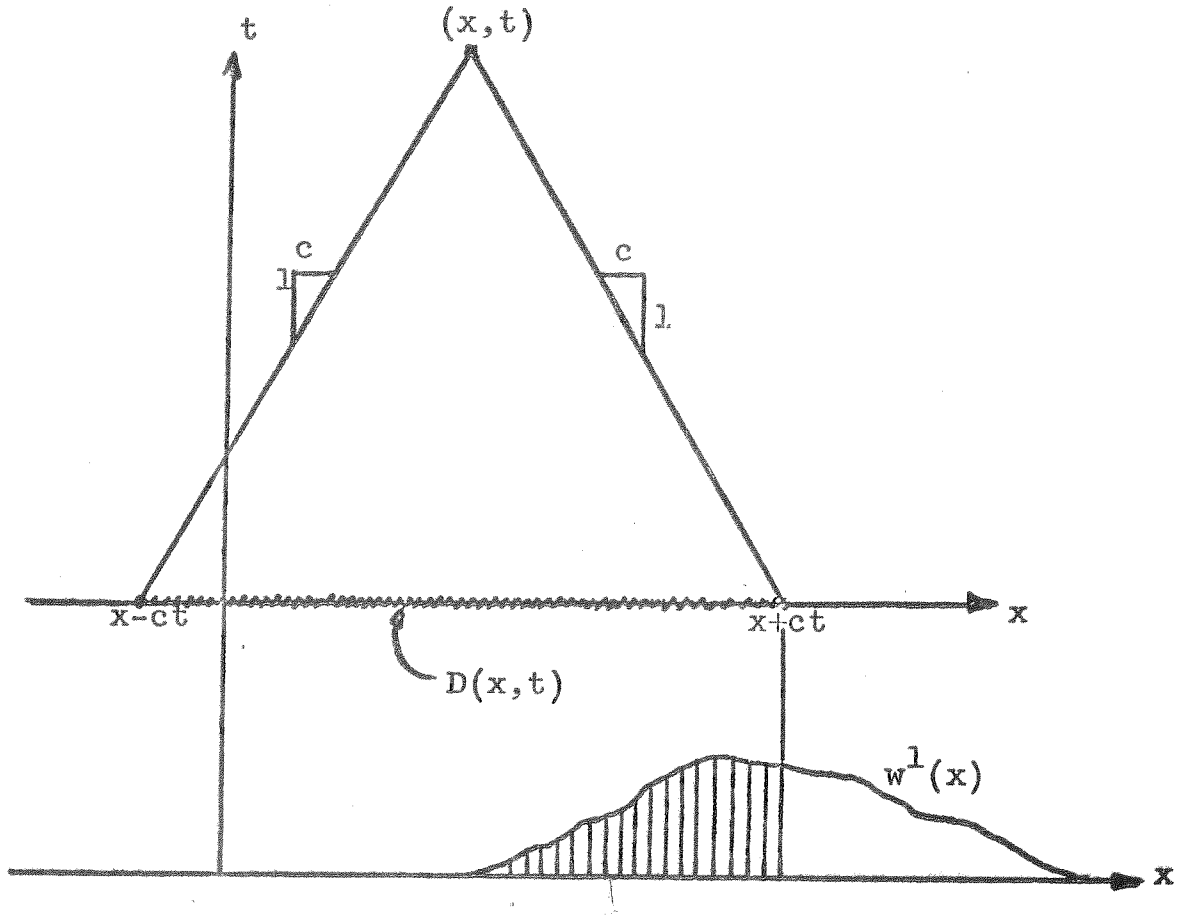


Figure 1

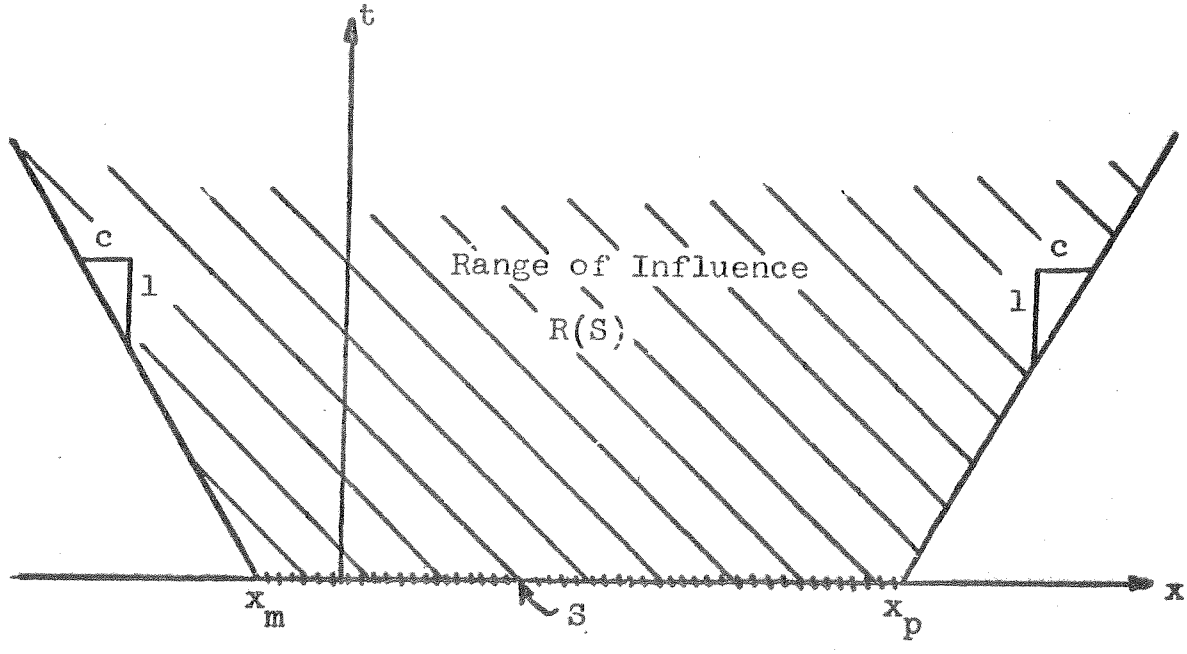


Figure 2

This formula represents the unique solution of the initial value problem 1 and 2, and will be a classical solution if w^1 and w^0 are respectively once and twice differentiable.

It is convenient to describe the motion by discussing the graph of $w(x,t)$ over the x,t-plane, see Figure 1. To interpret the solution geometrically, take the point (x,t) in the x,t-plane and draw downward through (x,t) two lines of slope $+1/c$, $-1/c$. These lines intersect the x-axis at $x-ct$, $x+ct$. The quantity $\frac{1}{2}(w^0(x+ct)+w^0(x-ct))$ in 3 represents the average of the values of the w^0 curve at the points $x+ct$ and $x-ct$. The integral in 3 represents the area under the w^1 curve cut off by $(x-ct)$, $(x+ct)$.

The <u>domain of dependence</u> of the point x_0 at the time t_0 on the initial line $t = 0$ is the segment

$$D(x_0,t_0): x_0 - ct_0 \leqq x \leqq x_0 + ct_0 , \quad t = 0 .$$

The significance of this domain $D(x_0,t_0)$ is that the solution at (x_0,t_0) depends on the initial data only on $D(x_0,t_0)$. This means: the solutions produced by two different sets of initial data agree at (x_0,t_0) if the initial data agree in $D(x_0,t_0)$ even if they differ outside of $D(x_0,t_0)$.

The <u>range of influence</u> of a segment $S: x_m \leqq x \leqq x_p$ on the initial line $t = 0$ consists of all points x at any time $t = 0$ for which

$$R(S): x_m - ct \leqq x \leqq x_p + ct ,$$

see Figure 2. Its significance is that the DOD of any

point outside of it does not contain the segment S. Consequently, the initialdata on the segment S do not influence the solution outside the range R(S). This means: two solutions produced by two different sets of data agree outside of the range R(S) if the data agree outside of the set S. This fact is also frequently stated by saying that the influence of disturbances of the initial data travels with the speed c.

Another notion is the domain of determinacy

$$E(S): x_m + ct = x = x_p - ct$$

of the segment S, see Figure 3. It consists of all points (x,t) whose DOD lies inside of S. The solution at points in E(S) are therefore completely determined by the data in the segment S. To illustrate the significance of the DOD, consider, say at the same time $t=t_o$ two segments, represented by the lines R_1 and R_2 in Figure 4. If the DOD, as in the figure, do not overlap, then the deflections at $t=t_o$ at the two segments of the string will be entirely incoherent; they will be caused by initially prescribed deflections and velocities which are entirely independent of each other.

The existence of the DOD and the range of influence in the x,t-plane is the decisive fact for the wave character of solutions. An initial disturbance, concentrated at t=0 around a point x_o, will at $t=t_1>0$ have "spread" only to points inside the range of influence.

The lines $x-x_o = \pm ct$ in the x,t-plane forming the

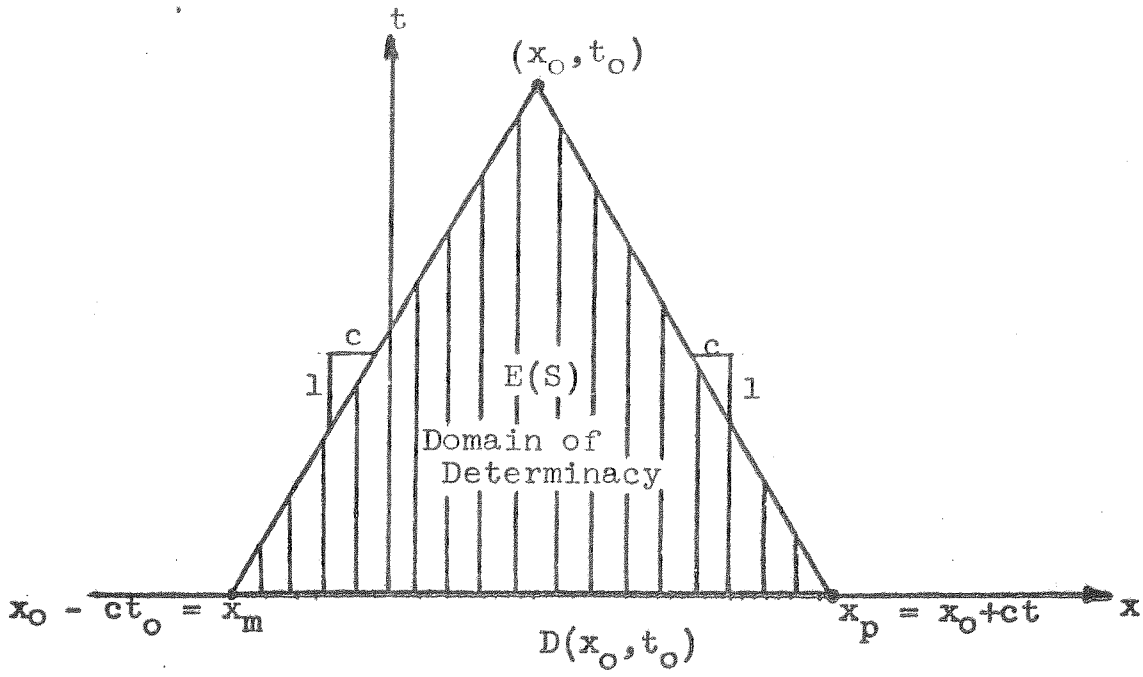


Figure 3

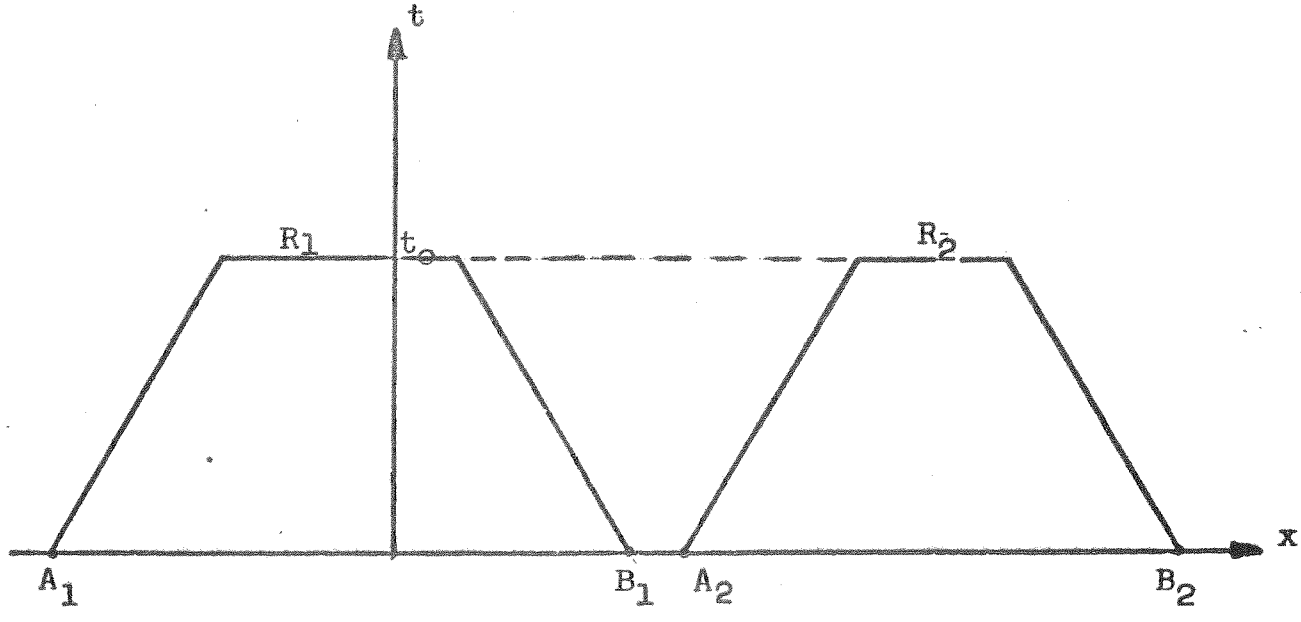


Figure 4

boundary of the domain of influence of the point $(x_o,0)$ are often called the characteristic lines of the differential equation (1). The existence of the wave character of solutions lies in the fact that if the DOD is drawn for every point P of the region R_1 of the x,t-plane it will be found that the value of $w(x,t)$ at these points depends only on the initial conditions in the interval A_1B_1 of the t=0 axis; the same is true for the values of $w(x,t)$ at points of the region R_2 which is determined by a different interval A_2B_2. Hence the values of $w(x,t)$ in the space-time region R_2 are not dependent on the initial values on A_1B_1, while the values of $w(x,t)$ in the space-time region R_1 does not depend on the initial values on A_2B_2. Thus the DOD illustrates a distinguishing trait of the wave equation as compared with, for example, the potential equation. All these considerations show the significance of the constant c; this constant is called the "speed of propagation", "wave speed", "light speed", etc.

If a given external force is acting on the string, the homogeneous equation (1) has to be replaced by

$$w_{tt} - c^2 w_{xx} = f(x,t) \quad , \tag{4}$$

in which $f(x,t)$ represents the external force per unit mass, and the initial conditions are as before (2). The solution is given by Friedrichs (23)

$$w(x,t) = \tfrac{1}{2}(w^{o}(x+ct)+w^{o}(x-ct)) + \frac{1}{2c}\int_{x-ct}^{x+ct} w^{1}(s)ds$$
$$+ \frac{1}{2c}\int_{0}^{t}\int_{x-c(t-s)}^{x+c(t-s)} f(y,s)dyds \ .$$

The domain of integration of the second integral is the domain of determinacy of the point (x,t).

If the wave equation in two spatial dimensions

$$w_{tt} - c^{2}(w_{xx} + w_{yy}) = f(x,y,t) \tag{6}$$

is considered, a similar analysis would reveal the DOD in the x,y-plane to be a circle and the domain of determinacy in x,y,t-space to be a right circular cone. In this case rather than characteristic lines there would be a characteristic surface, the surface of the cone forming the characteristic surface. This cone is called the characteristic conoid or light cone.

If the medium is inhomogeneous, i.e., the coefficient c is a function of spatial location, or if the differential equation is nonlinear; then the characteristic lines are no longer straight.

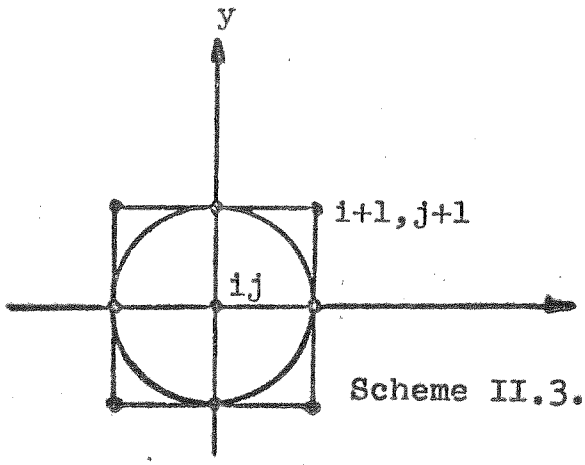
An important aid to the study of some of the practical aspects of difference schemes for hyperbolic systems is to consider the DOD of the differential and difference equations. A result of the famous 1928 paper of Courant-Friedrichs-Lewy (16), is that a necessary condition for stability is that the DOD of the difference equation must contain the DOD of the differ-

ential equation.

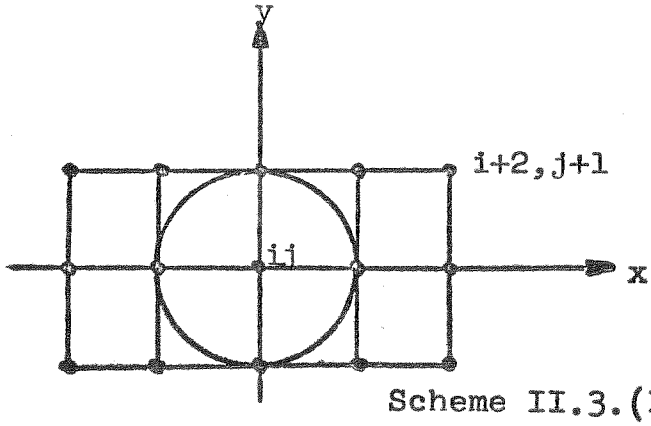
In Figure 5, the DOD in the x,y-plane at time m is shown for v_{ij}^{m+1} for the Strang schemes II.3.13 and II.3.14 . The circular area is the DOD of the differential equation, and the rectangular area outlined by the mesh points is the DOD of the difference equation. The C-F-L condition is seen to be satisfied by each of the schemes.

In the 45 degree directions scheme II.3.13 propagates signals at a speed $\sqrt{2}$ times that in the x and y directions, directions in which signals are propagated at the correct speed for the constant coefficient case with $\Delta x = \Delta y$ = constant. The C-F-L condition implies that all difference schemes will have a precurser wave (dispersion) of some amount. Notice that this $\sqrt{2}$ figure is not reduced as the mesh is refined, it is a property of the scheme. As the mesh is refined, this precurser signal, while nonzero, becomes less and less significant and sensible signals are propagated closer and closer to the correct speed in all directions.

This is precisely the point at which a major difference occurs between one and multi-spatial dimensioned problems. For many problems in one spatial dimension (i.e., linear, constant coefficient) it is possible for the DOD of the differential equation and the difference equation to be identical. The above discussion indicates that this is not possible in problems



Scheme II.3.(13)



Scheme II.3.(14)

Figure 5

Domain of Dependence for Differential Equation and Difference Equation

where the spatial dimension is two or more.

At one time the writer felt that making the DOD of the differential and difference equations more nearly equal than is the case in the scheme II.3.13 was a strategy worth pursuing as a means of increasing accuracy (by definition it would decrease dispersion). It seems that schemes which have this advantage have the disadvantage of computational complexity. Additionally, for nonlinear, and even variable coefficient linear problems, the wave speeds will vary with the spatial location (and time in the nonlinear case) so that it is not clear whether any computational advantage is to be obtained over simpler schemes by focusing in on the DOD concept in the more general case. No results are included here so the matter is very much open.

The DOD concept assists in explaining a phenomenon observed in the solution of implicit difference schemes. Consider an implicit scheme of the form

$$A\, u^{n+1} = u^{n}$$

where A is a matrix and u^n the state vector at time m. The solution will behave as if it were obtained by

$$u^{n+1} = A^{-1} u^{n} .$$

Even if A is banded etc., so that it reflects closely the true DOD, the matrix A^{-1} will in general be full, so that the solution u^{n+1} at a point ij will depend on the solution at time n of all the points in the domain. This leads to the large dispersion observed in implicit

schemes. In effect, the signal speed is infinite in an implicit scheme.

In principle then, one might say that implicit schemes lead to gross errors because of their gross overestimation of the DOD. In practice however, the situation is not quite this bad. If the implicit scheme is a convergent one, sufficient refinement will produce results reflecting the hyperbolic character of the problem. The practical question is: which scheme requires less computational effort?

The situation is analogous to that when the Fourier method is used, i.e., separation of variables and Fourier expansion. If the eigenfunctions of the reduced partial differential equation are available this method seems to work quite well, see Goudreau (25). In general these eigenfunctions are not available, and the implicit solution of the system of ordinary differential equations obtained by a spatial discretization such as the FEM, for example, is similar to using a general non-eigenfunction expansion in the Fourier method. If these functions form a basis, the solutions will converge, but not as rapidly as when using the eigenfunctions. In a functional expansion of this type each function "sees" a disturbance immediately upon application regardless of how far away the disturbance is; this means an infinite speed of propagation.

Put in a slightly different way this means that an

implicit scheme completely ignores the finite speed of propagation character of hyperbolic PDE's and concentrates instead on solving the system of ODE's without regard to their source.

A similar situation can arise in explicit schemes also. For example, consider the semigroup scheme II.6.12 with n = 2 written for the equations of elasticity. A stable time step was found numerically to be $\frac{1}{4}$ the optimum time step. This means that the difference scheme propgates signals a minimum of four times faster than the differential equation. For wave propagation problems in a finite domain, after only a few time steps the DOD can be the entire domain, the situation which occurs in implicit schemes. This fact alone would probably preclude the semigroup scheme II.6.12 from being a useful improvement over the standard scheme II.6.4.

The DOD concept will prove useful in discussing other practical aspects of computing in the sequel.

2. Irregular Mesh. The question arises in practical problems as to how to handle an irregular mesh. A simple case is illustrated in Figure 6, where $\Delta x = \Delta y$ and $k > 1$. Numerical experiments confirm that when applying the Strang scheme II.3.13, for example, at point ij in the figure, interpolation can be used to obtain the needed values at the non-existant mesh points a,b and c and stability maintained. The Lax Equivalance Theorem (see II.1) then guarantees convergence. Actually, the Lax theorem is not needed in this case, since the approximation of derivatives by differences is not restricted to the concept of nodal function values. In most general theorems convergence is only guaranteed in the mean, so a preoccupation with "pointwise" values of an approximate solution can be misleading.

Another strategy which maintains stability is simply to use the values at the point (i+1,j+1) in Figure 6 instead of interpolating between points (i,j+1) and (i+1,j+1) for the needed value at "point" a, etc. The C-F-L DOD requirements are not violated since with either strategy the difference scheme at point ij has the same DOD, namely the rectangular area with corners $(i \pm 1, j \pm 1)$, the weighting of the coefficients C_{ij} in the difference scheme II.3.3 is just slightly different.

The latter strategy requires less storage and computation and the former seems intuitively more accurate; though proving it more accurate is quite another

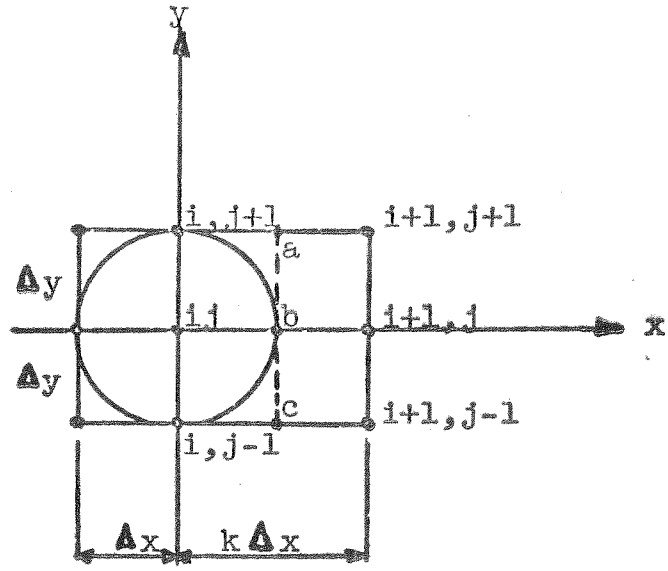


Figure 6

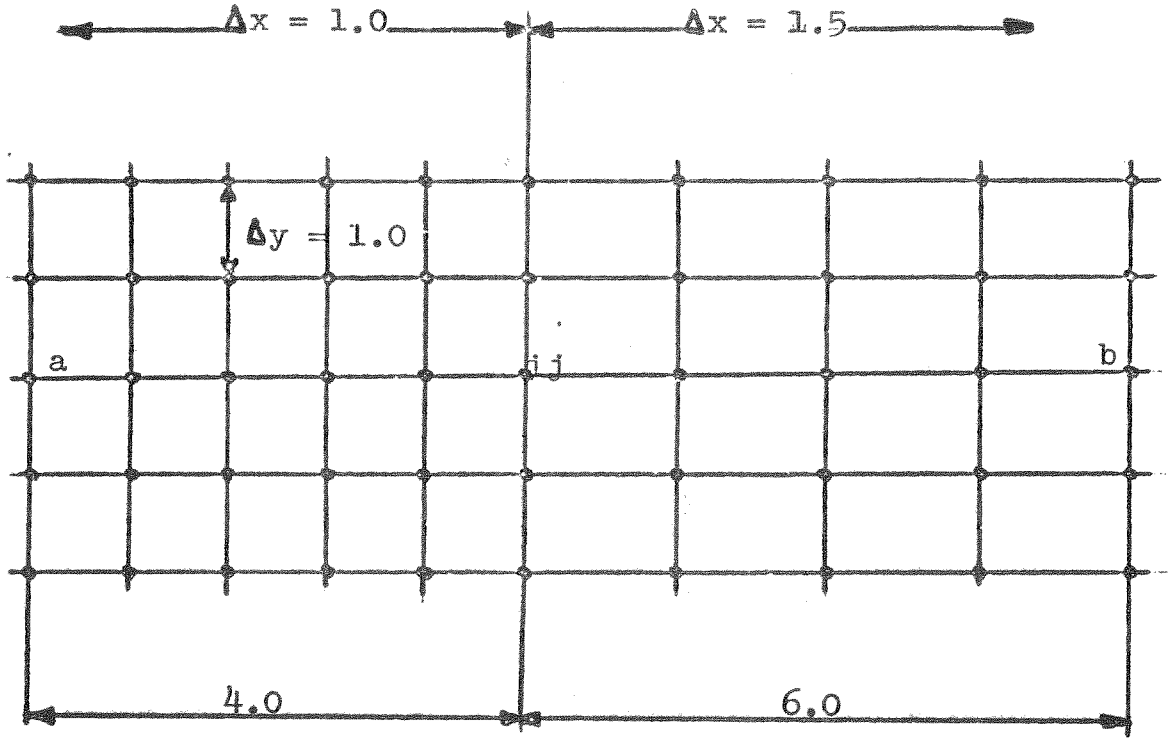


Figure 7

thing and proving it "better" still another. Hence a clear choice between them is difficult if not impossible and remains a matter of taste. In the limit the results would be indistinguishable. An honest numerical analyst cannot call the results of a single approximate analysis a solution until he has refined the mesh (in this case both spatially and temporally) and finds the results essentially unchanged. This latter procedure (refinement to verify convergence) will be made unnecessary when sharp pointwise truncation error estimates become available (a posteriori); and at less computational effort than refinement. Sharp pointwise error estimates are a major unsolved practical problem in numerical analysis.

These "interpolation" methods and variants work on even more severely distorted meshes - the thing to keep in mind are the DOD's.

These considerations can be used to give an interesting example of what can happen if the casual mesh spacing used in equilibrium problems is used in wave propagation problems. Consider Figure 7. Assume the Strang scheme II.3.13 with one of the above interpolation schemes is used, that nonzero initial data is given only at point ij and that there is no forcing term. Assume further that the wave speed is 1 and the time step used is 1. Then at the end of 4 time steps the effect of the initial data will be observed at point a four units to the left of ij and at point b six units to

the right of ij. Thus the use of casual mesh spacing propagates signals at the wrong speeds. This variation in wave speed is ordinarily obtained by a change of material properties; elastic constants, density or both in the case of elasticity. But if the medium is isotropic and homogeneous there is a contradiction and some error. In effect, the mesh spacing has introduced a change of material.

These considerations have possibly an even greater effect in problems with boundaries. In that case the propagation of signals at the wrong speeds might mean that a given quantity, e.g. stress, upon reflection would be diminished rather than increased as it should or vice versa. Hence the solution could even be of opposite sign rather than just a few percent off.

The conclusion is that in wave propagation problems, as opposed to equilibrium or vibration problems, if the medium is isotropic and homogeneous and an explicit scheme is used with Cartesian coordinates the mesh spacing should be uniform and equal in all coordinate directions if at all possible.

An intriguing and elegant approach to irregular meshes and irregular boundaries is described by Lapidus (38). There he performs a transformation of Cartesian space which changes a curved boundary into a straight line, thus reducing the large number of special points and irregularly shaped mesh regions which would otherwise

appear in the difference method calculation. Such transformations are shown to preserve the conservation property of the system of differential equations. This transformation performs on the global level what is routine on the local level in the FEM and bears further investigation as a means of simplifying difference calculations.

In the example illustrated by Figure 7, this method of Lapidus would transform the mesh to one with $\Delta x = 1$ everywhere (Δy remains unchanged at 1) and change the material properties of the medium to the right of the vertical line through point ij in such a way that the medium would no longer be isotropic and homogeneous. This would eliminate the need to perform any interpolation but would necessitate additional storage (fast storage not required) for material constants and a more complicated scheme because of the anisotropy. The method would probably be attractive in those cases with a high degree of irregularity. In any event the inverse transformation would only be required at those relatively few points where output is desired.

3. Material Interfaces. The problem of what to do at a material interface is easily handled by either of the interpolation processes described in Section 2 and some simple observations. Again the DOD concept is central. Two situations will be discussed; the first concerns the case when finite differences are used to perform the spatial discretization and the second where the FEM is used. Only a straight interface is discussed but an interface of arbitrary shape can be handled by the method of Lapidus (Section 2) or by approximating the irregular interface by piecewise linear interfaces.

a. Finite Differences. Consider the Strang scheme II.3.13. Assume that the material on either side of the interface is isotropic and homogeneous and that equal x and y spacing is used in each material region. Further, make the (not necessary) assumption to integrate the equations using the same time step in both material regions. The choice could be made to use the same spatial spacing in both regions, choosing the smaller time step to insure stability. This choice would lead to larger than necessary dispersion or precurser waves in the material with the smaller signal speed. It is more in keeping with the central idea of DOD to use different mesh spacing for different materials. Choose one spacing h_1 in Material 1 and the maximum stable time step k according to equation II.4.2 where the material coefficients used in this equation are those of Material 1. Then

$e(A) = ((\lambda_2 + 2\mu_2)/\rho_2)^{\frac{1}{2}}$ and $ke(A) \leq h_2$ allows the solution of the smallest permissible mesh spacing h_2 in Material 2 commensurate with stability and the time step chosen. Choosing the mesh spacings in this way dilatational signals are propagated at the correct speed in the x and y directions and travel the distance between two adjacent nodes in exactly one time step in both materials (except at the interface).

At the interface there are at least two possibilities:

1. Place the nodes on the interface.
2. Place the interface between two lines of nodes.

These two possibilities as well as the assumptions made previously are depicted in Figures 8 and 9. The values at the "points" marked "i" are obtained by the interpolation process described in Section 2. The presence of an interface is seen to always increase dispersion when an interpolation strategy is used.

In the technical literature the conditions often imposed (1,28) at an interface for a displacement finite difference method (but not a displacement FEM) are the physical ones; i.e., continuity of displacements, normal and shear stresses. If the material coefficients varied continuously with the spatial location this approach would clearly be impractical and only continuity of displacements would be imposed (required by uniqueness) and the stresses would be computed using the material value at the point of interest. Since an abrupt change in material coefficients can be approximated as closely

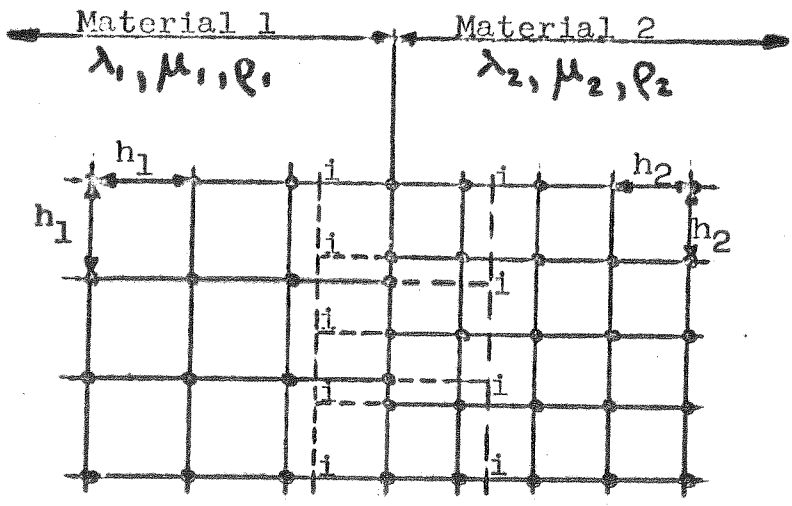


Figure 8

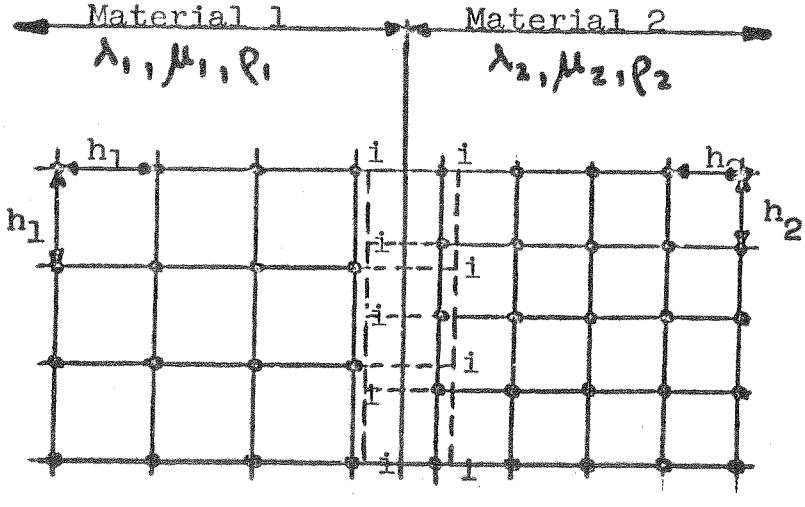


Figure 9

as desired by not only continuous functions but infinitely differentiable ones the problem of a material interface can be considered as that of a differential equation with variable coefficients. Taking this latter attitude allows the difference technique used at interfaces to fit in to existing stability and convergence theory. To be precise - when using scheme II.3.13 to calculate v_{ij}^{m+1} the material coefficients to be used in the matrices A and B are those which exist at the point ij. If the point ij lies on the interface an approach similar in effect to that used in the FEM is to use some average of the material coefficients. Numerical experiments confirm that all of these approaches are stable. In a recent paper by Ciment (72), analytical and numerical results are presented justifying similar ideas.

Strang (58) has reported that in the case of variable coefficients the stability conditions II.3.9 need to be strengthened to a strict inequality. This is the case, for example, if polar or cylindrical coordinates are used even if the material properties are constant. In the numerical experiments at no time did the stability condition need to be strengthened in the case of a material interface.

b. Finite Element. Surprisingly, the strategy which maintains stability in the case where the FEM is used for spatial discretization is relatively more complicated than that for the finite difference case. This is

because the FEM usually handles irregularities so easily, almost automatically, while in the finite difference case one is used to handling them in a more complicated manner.

The same assumptions as in Subsection a will be made, i.e., the material on either side of the interface is isotropic and homogeneous, equal x and y spacing is used in each material region and the same time step is used in both material regions. Again, the choice could be made to use the same spatial spacing in both regions choosing the smaller allowable time step to insure stability. But this choice leads to larger than necessary precurser waves as was discussed in Subsection a. Hence the spatial spacings h_1 and h_2 in regions 1 and 2 will be chosen as in a.

It is at this point that the finite difference methods can operate at a precision that the FEM cannot. If at this point the same method were used as in the finite difference case, then there would in effect be a mixing of finite element and finite difference techniques and quite a complicated one at that. The writer is not aware of explicit schemes for the FEM which can operate at the optimum time step when the spatial dimension is two or more. In some special cases, e.g., propagation of plane waves, the writer has found an explicit FEM to be stable at the optimum time step for quite some time, long enough to solve the problem in fact, but in the general case this cannot be expected. This means that

in general explicit FEM schemes will have somewhat more dispersion than the Strang scheme II.3.13 even under the most ideal circumstances so the strategy described below does not exhibit excessive dispersion. An idea of the amount of dispersion to be expected can be obtained from the stability condition II.6.5.

Returning to the strategy to be used at an interface in a problem when using an explicit FEM scheme, assume the spacings h_1 and h_2 obtained above are ordered such that $h_1 > h_2$ and further that h_1/h_2 is near 2. Modify one of the spacings, say h_2, to h_2' in such a way that $h_1/h_2' = 2$. The time step chosen previously will have to be reduced if $h_2' < h_2$. Then the mesh should be made as shown in Figure 10. The spacing h in the figure has to be strictly greater than h_2' if $h_2' < h_2$ and strictly greater than h_2 if $h_2' > h_2$. This comes about from the DOD requirements and the stiffnesses, i.e., the stiffnesses at the nodes $(i,j) = (4,1),(4,3),(4,5)$, etc. The stiffnesses at these nodes are some sort of average of that at the regular nodes in the interior of the two regions since the interface is on line $i = 4$. The reason h should be kept small is to reduce dispersion, but its computation would probably not be worth the effort and setting $h = h_1$ eliminates all difficulties. There are no "nodes" at $(i,j) = (5,2),(5,4)$, etc. The line of elements between $i = 5$ and 6 can handled by special elements as in Figure 11 or by a special assembly

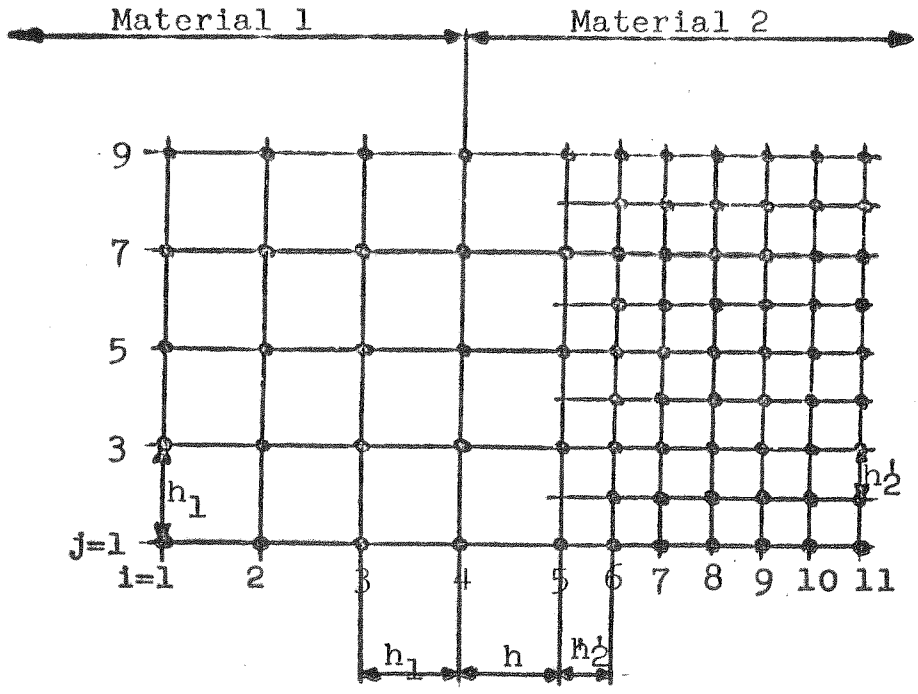


Figure 10

Figure 11

Figure 12

technique using elements as in Figure 12.

If h_1/h_2 is near 3 then three rows of elements will be required in Material 2 per row in Material 1, etc. If h_1/h_2 is near 1, then the same mesh spacing should be used in each material region.

4. Boundaries. In Subsection II.4 a method for treating boundaries when the boundaries are parallel to the coordinate axes was given for the Strang schemes. In this case a FEM needs no modification in order to maintain stability (as long as the DOD requirements are observed) and the boundary conditions can be handled in the usual way. In this section irregular boundaries are treated by piecewise linear boundaries.

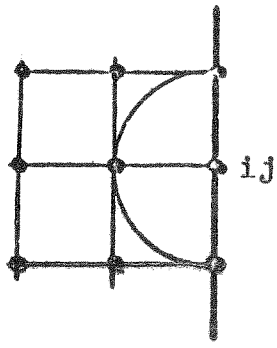
For both difference methods and the FEM the simplist treatment of straight boundaries is illustrated in Figure 13. Here rectangular approximations to the boundary are made.

For the Strang scheme II.3.13 an alternative method is illustrated in Figure 14. Here the solution is advanced at the points on the boundary by transforming to coordinates parallel and perpendicular to the boundary and using the interpolation scheme described in Section 2. Notice that some points have to be "left out" so that the DOD requirements are not violated. The one-sided differences described in Section II.4 for the treatment of boundary points have the unfortunate effect of doubling the DOD in the direction normal to the boundary for the

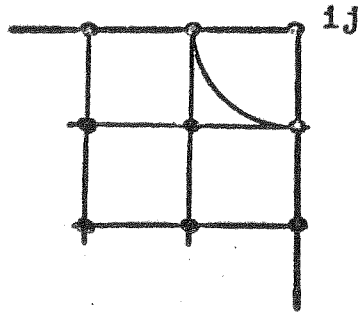
scheme II.3.13. The DOD's for a point ij at a straight boundary and a corner point are shown in Figure 15. Strang (58) has reported that use of first order accurate differences rather than II.4.5 at boundary points does not always destroy overall second order accuracy. If first order one-sided differences were used with the first order scheme II.3.19 at boundary points only, then the ratio of the DOD of the difference scheme to the DOD of the differential equation would have the same value at both boundary and interior points. Even if the use of a first order scheme in this situation resulted in some loss of accuracy there would be the compensating factors of less dispersion and less computational effort because of the use of a simpler scheme on the boundary. Not inconsiderable is that interpolation would have to be performed for fewer points.

An alternative approach to that illustrated in Figure 13 is shown in Figure 16 for the FEM. A row of special elements is placed along the boundary to preclude violation of the DOD requirements.

It is not obvious to the writer that the scheme illustrated in Figure 13 necessarily leads to more error than in the schemes shown in Figures 14 and 16. All interpolatory schemes at irregularities necessarily lead to an increase in dispersion, the DOD cannot be reduced below the C-F-L necessary requirement. As was noted in Section 2 this increase in dispersion can be interpreted



Straight Boundary



Corner

Figure 15

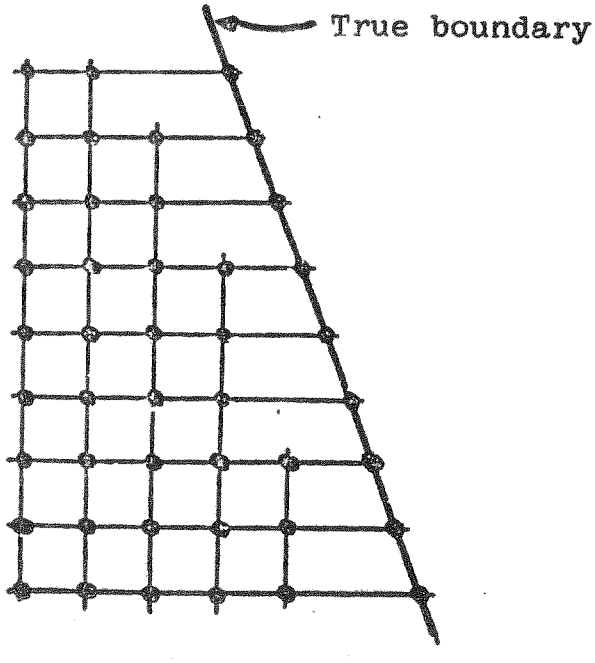


Figure 16

as a change of material properties. Thus the difference between the scheme illustrated by Figure 13 and that by Figures 14 and 16 can be interpreted as a choice between adding or removing material and changing material properties. In reality then, the two approaches are quite similar - the material is modified in some way. In addition, the scheme illustrated in Figure 13 is not interpolatory and it would have the compensating factors that some signals would be reflected from the approximate (rectangular) boundary earlier than they should and others later, rather than always earlier as in the interpolatory schemes.

As was noted in Section 2, the approach of Lapidus can also be used to simplify the boundary treatment.

The rather neat aspect of the DOD concept is that it allows one to see precisely how and when errors are introduced at irregularities although it gives no indication of how much. As the above discussion has indicated, at an irregularity the wave speed in the difference scheme is always increased when interpolation is used.

Other methods could be used to modify the spatial differences at irregularities, see Forsythe and Wasow (20) and Collatz (15).

5. Artificial Viscosity. The stability theory as presented in Chapter II is applicable only to problems whose solutions are smooth. For problems with shocks the calculation is likely to become unstable unless something is done to prevent it. The problem arises because while in any problem where the medium does not tear the displacements are continuous, a function of their spatial derivatives, e.g. stress, need not be. Since the Lax-Wendroff method applied to elastodynamics solves directly for stresses (which may be discontinuous, i.e. shocks) the methods presented in Chapter II can prove ineffective because a (possibly) nonunique solution is being sought.

In order to get around these difficulties, von Neumann and Richtmyer (46) devised an approximate method of calculation in which shocks are taken care of automatically, whenever and wherever they arise. The method is based on certain physical notions concerning real (and therefore nonideal) materials rather than on a direct attempt to incorporate the Rankine-Hugoniot jump condition into the calculations. Specifically, dissipative mechanisms, like viscosity and heat conduction, have a smoothing effect on a shock, so that the surface of discontinuity is replaced by a thin transition layer in which quantities change rapidly, but not discontinuously. The differential equations with the dissiaptive mechanism included, apply in this layer as well as elsewhere, so that no boundary conditions at the shock are

needed. A shock can be thought of as an internal boundary with boundary conditions, but except in the linear case the location of the boundary is unknown. At the same time, the basic conservation laws on which the Rankine-Hugoniot conditions are based are retained and the jump conditions still hold across the transition layer, in the approximation in which this layer is regarded as thin in comparison with the other dimensions occuring in the problem.

The idea of von Neumann and Richtmyer was to introduce a purely artificial dissipative mechanism of such form and strength that the shock transition would be a smooth one extending over a small number (say three or four) of intervals Δx of the space variable, and then to construct difference equations with the dissipation included, but without any necessity for shock fitting. In calculations with such equations, the shocks show up automatically as near discontinuities, across which quantities have very nearly the correct jump, and which travel at very nearly the correct speed (see Richtmyer and Morton (52)).

A mechanism like viscosity is more suitable for this purpose than one like heat conduction, because as Becker (3) showed, with viscosity all quantities vary smoothly (i.e. velocity, stress, mass, energy, etc.). However, Becker also showed that with ordinary viscosity, which is represented by linear terms in the differential

equations, the thickness of the transition layer varies with the shock strength, approaching zero for a very strong shock and infinity for a very weak one. But if the thickness is to be about the same, namely about $(3\text{-}4)\Delta x$ - for all shocks, quadratic terms should be added to the differential equation; this is equivalent to using a small coefficient for weak shocks and a large one for strong shocks. This method has the advantage that the pseudo-viscous terms, being quadratic in the strain rate, are very small in the smooth part between shocks, where one wishes the behavior to approximate that of an ideal material.

The linear viscosity terms have been found necessary to insure stability in the presence of shocks where the differential equation is nonlinear, see Richtmyer and Morton (52). Otherwise, the quadratic viscosity terms have been found sufficient. Some analysts object to artificial viscosity thinking that introducing it will severely distort, if not destroy, the true behavior of the physical system. A calculation by Wilkins (64) is quite illuminating in this respect. The calculation concerned the vibration of an elastic plate, modeled as a two dimensional elastic medium, which is clamped at one end. The plate is set in motion by an initial velocity and oscillates with the correct period compared to known solutions. The calculation was allowed to run for over 10,000 time steps. The wave amplitude and frequency

remained constant. The total energy, kinetic plus internal, remained constant to within 0.1% of the initial energy put into the plate. An artificial viscosity was used to stabilize the grid. The point is to show that the physics of the problem is not influenced by the artificial viscosity.

Again the Strang schemes for the Lax-Wendroff method allow simplifications. The artificial viscosity can be introduced in a one dimensional manner in the one dimensional operators L^x and L^y, see Strang (59) and Richtmyer and Morton (52).

In the numerical experiments it was never found necessary to use artificial viscosity to stabilize the calculation. Perhaps this is because only linear problems were attempted and because the Strang schemes have a built-in viscosity. In many of the problems the transitions were quite sharp, e.g., large negative to large positive values of stress in one spatial spacing.

By way of a partial theoretical justification of the expectation that these procedures give a correct description of problems with shocks, some results of Lax and Wendroff (42) are quoted. The first is expressed in terms of the concept of a weak solution (more precisely a distributional solution) of a system of conservation laws, which will now be defined. Consider the system of conservation laws in one space variable x,

$$u_t + f_x = 0 \quad , \quad u(x,0) = u_o(x) \tag{1}$$

where $f = f(u)$ is a nonlinear vector function of the solution $u = u(x,t)$. The phenomenon of nonexistence of smooth solutions corresponds to the development of shock waves in the physical problem. Mathematically these can be described as curves $x(t)$ across which $u(x,t)$ fails to be continuous, but such that u has a limit as points (x,t) approach the curve from either side. The shock curves must also respect the fact that equation 1 is a conservation law:

$$(2) \qquad d/dt \int_b^a u(x,t)dx = f(b,t) - f(a,t) ,$$

that is, the rate of change of the integral is due to the flux at the end points. Similar relations hold for conservation laws in more than one spatial dimension. A discontinuous solution satisfies 2 if

$$(3) \qquad \begin{aligned} s(u(x+0,t)-u(x-0,t)) &= f(u(x+0,t)) - f(u(x-0,t)) \\ \text{or} \qquad s[u] &= [f(u)] \end{aligned}$$

where s is the shock speed. This is known as the Rankine-Hugoniot jump condition. If $u(x,t)$ is any genuine solution of 1 with given initial condition $u(x,0) = u_0(x)$, and if $w(x,t)$ is any smooth, vector-valued function with the same number of components as u which vanishes outside a finite region of the x,t-plane (such a function is called a test function), then, on taking the scalar product of $w(x,t)$ with equation 1, integrating over all x and all $t \geq 0$, and then integrating by parts, one finds,

$$(4)\quad \int_0^\infty \int_{-\infty}^{\infty} ((w_t,u)+(w_x,f))dxdt + \int_{-\infty}^{\infty} (w(x,0),u_o(x))dx = 0 .$$

Any u(x,t) whether a genuine solution or not, that satisfies this equation for all test functions w(x,t) is called a weak or distributional solution of 1 with initial values $u_o(x)$. It follows from this definition (see Lax (40)) that at a jump discontinuity a weak solution satisfies the jump condition 3. The result of Lax and Wendroff, referred to above, is:

Theorem. If the solution of the Lax-Wendroff difference equations II.2.7 , with or without viscosity terms, converges boundedly almost everywhere to a function v(x,t), then v(x,t) is a weak solution of the system of conservation laws.

Lax and Wendroff (42), put it succinctly, "the strong limit of the approximate solution is a weak solution of the conservation law".

The concept of a weak solution should be looked upon as a mathematical attempt to restore to the differential form of the conservation law the generality of the integral conservation law from which it came.

There is a corresponding nonuniqueness for problems which is rectified by eliminating certain discontinuous solutions as nonphysical, namely those whose entropy decreases across discontinuities. Such an entropy condition, Condition E, was formulated by Oleinik (47) for equation 1 in the case where u is a scalar as fol-

lows: If u^- and u^+ are the limits of u from the left and right respectively, for fixed time, and v is any number between u^- and u^+, then

$$S[v,u^-] \geq S[u^+,u^-]$$

where

$$S[a,b] = \frac{f(a) - f(b)}{a - b} .$$

Recently, Quinn (47) obtained an intriuging theorem which illustrates the fact that physical concepts often have an abstract mathematical counterpart.

To state this theorem two definitions will be needed. Recall from I.2 the notion of a semigroup. A semigroup is said to be contracting in a certain norm, denoted by $\| \ \|$, if

$$\| T(t)u - T(t)v \| \leq K \| u - v \|$$

for all solutions u and v of 1 and all $t \geq 0$, where the constant K is equal to one. In an ordinary semigroup the constant K is simply some finite positive real number. The L_1-norm of u(x,t) at time t is defined by

$$\| u(x,t) \| = \int |u(x,t)| \, dx .$$

Then the statement of Quinn's theorem is:

Theorem. If u(x,t) and v(x,t) are piecewise continuously differentiable solutions of 1 (where u is a scalar) for all x and all $t > 0$, with initial data $u_o(x)$, $v_o(x)$ which are piecewise continuously differentiable and L_1-integrable in x, and if u and v satisfy Condition E at all shocks, then the set of solution operators is a

contracting semigroup in the L_1-norm. Conversely, T(t) contracts only if E is satisfied.

The strong stability of the Strang schemes (II.3) implies that the underlying semigroup is contractive, but in the L_2-norm. If the Strang schemes could be shown to imply the L_1-contractiveness of the underlying semigroup then their use would imply the entropy Condition E is satisfied.

The stress discontinuity at a shock is similar to that at a material interface, hence these general considerations apply to this latter situation also.

REFERENCES

1. Alterman, Z. and F.C. Karal Jr.: Propagation of elastic waves in layered media by finite difference methods. Bulletin of the Seismological Society of America, v 58, n 1, p. 367, February 1968.

2. Baltov, A.: The plane problem for elastic/visco-plastic bodies. Archiwum Mechaniki Stosowanej, v 2, n 18, 1966.

3. Becker, R.: Stosswelle und detonation. Z. Physik, v 8, p. 321, 1922.

4. Bejda, J.: Propagation of two-dimensional stress waves in an elastic/viscoplastic material. Proceedings of 12th International Congress of Applied Mechanics. Stanford: August 26-30, 1968, Editors M. Hetenyi and W.G. Wincenti, Springer-Verlag, 1969.

5. ---------: A difference method for plane problems in dynamic viscoplasticity. In Hyperbolic Equations and Waves, Battelle Seattle 1968 Recontres, Ed. M. Froissart, Springer-Verlag, 1970.

6. Biot, M.A.: Variational principles in heat transfer. Oxford Mathematical Monographs, Oxford University Press, 1970.

7. Buchanan, M.: A necessary and sufficient condition for stability of difference schemes for second order initial value problems. Adelphi College, Dept, Grad. Math., Res. Rep. n 104, 1962.

8. Burnat, M., A. Kielbasinski and A. Wakulicz: The method of characteristics for a multidimensional gas flow. Archiwum Mechaniki Stosowanej, v 3, n 16, p. 179, 1964.

9. Burstein, S.Z.: Numerical methods in multidimensional shocked flows. AIAA J., v 2, p. 2111, 1964.

10. ---------- : Finite-difference calculations for hydrodynamic flows containing discontinuities. J. Comp. Phys., v 2, p. 198, 1967.

11. ------ and A.A. Mirin: Third order difference methods for hyoerbolic equations. J. Comp. Phys., v 5, p. 547, 1970.

12. Butler, D.S.: The numerical solution of hyperbolic systems of partial differential equations in three independent variables. Proc. Roy. Soc., London, A 255, p. 232, 1962.

13. Clifton, R.J.: Analysis of dynamic deformations of elastic/plastic solids under conditions of plane strain. Ph.D. Thesis, Carnegie Institute of Technology, 1963.

14. ----------: A difference method for plane problems in dynamic elasticity. Quart. Appl. Math., XXV, n 1, 1967.

15. Collatz, L.: Numerical treatment of differential equations. Springer-Verlag, Berlin, 1960.

16. Courant, R., K.O. Friedrichs and H. Lewy: Uber die partiellen differenzengleichungen der mathematischen physik. Math. Ann., v 100, p. 32, 1928.

17. ---------- and D. Hilbert: Methods of mathematical physics, v 1. Interscience Publishers, New York, 1953.

18. Duff, G.F.D.: Mixed problems for linear systems of first order equations. Canad. J. Math., v 10, p. 127, 1958.

19. Estrin, M.I.: The equations of the dynamics of a compressible medium. Soviet Phys., Dokladi, v 5 n 6, p. 1349, 1961. (Translation of Dokladi Akad. Nauk SSR (N.S.) v 135 n 1, p. 36, 1960, by Amer. Phys. Inc., New York)

20. Forsythe, G.E. and W.R. Wasow: Finite difference methods for partial differential equations. Wiley, New York, 1960.

21. Friedrichs, K.O.: Ein verfahren der varionsmechnung, das minimum eines interals als das maximum eines anderen ausdruches darzustellen. Nachr. Ges. Wiss Gottingen, p. 13, 1929.

22. ----------: Symmetric hyperbolic linear differential equations. Comm. Pure Appl. Math. v 11, p. 333, 1954.

23. ----------: The theory of wave propagation. Inst. Math. Sci., New York University, 1957. Notes by C. Chester.

24. Garding, L.: Cauchy's problem for hyperbolic equations. Univ. Chicago, 1958.

25. Goudreau, G.L.: Evaluation of discrete methods for the linear dynamic response of elastic and visco-elastic solids. University of California (Berkeley) SESM Report 69-15, June, 1970.

26. Gourlay, A.R. and J.Ll. Morris: A multistep formulation of the optimized Lax-Wendroff method for nonlinear hyperbolic systems in two space variables. Math. Comp. v 22 n 104, p. 715, 1968.

27. ----------: On the comparison of multistep formulations of the optimized Lax-Wendroff method for nonlinear hyperbolic systems in two space variables. J. Comp. Phy. 5, p. 229, 1970.

28. Habberstad, J.L.: A two-dimensional numerical solution for elastic waves in variously configured rods. J. Appl. Mech. p. 62, March, 1971.

29. Hellinger, E.: Die allgemeinen ansatze der mechanik der kontinua. Enzyclopadie der Mathematischen Wissenschaften IV, 4, p. 654, 1914.

30. Hersh, R.: Mixed problems in several variables. J. Math. and Mech., v 12, n 3, p. 317, 1963.

31. Kato, T.: Estimation of iterated matrices, with application to the von Neumann condition. Numer. Math., v 2, p. 22, 1960.

32. --------: Perturbation theory for linear operators. Springer Verlag, New York, 1966.

33. Kreiss, H.O.: Uber die approximative losung von linearen partiellen differentialgleichungen mit hilfe von differenzengleichungen. Trans. Roy. Inst. Tech. Stockholm, n 128, 1958.

34. ----------: Uber die losung des Cauchyproblems. Acta Math., v 101, p. 180, 1959.

35. ----------: Uber die stabilitatdefinition fur differenzengleichungen die partielle differentialgleichungen approximieren. BIT, v 2, p. 153, 1962.

36. ----------: On difference approximations of the dissipative type for hyperbolic differential equations. Comm. Pure Appl., v 17, p. 335, 1964.

37. ----------: Initial boundary value problems for hyperbolic systems. Comm. Pure Appl. Math., v 23, p. 277, 1970.

38. Lapidus, A.: A detached shock calculation by second-order finite differences. J. Comp. Phy. v 2, p. 154, 1967.

39. Lavrentiev, M.M.: Some improperly posed problems of mathematical physics. Springer-Verlag, Berlin, 1967.

40. Lax, P.D.: Hrperbolic systems of conservation laws, II. Comm. Pure Appl. Math., v 10, p. 537, 1957.

41. --------- and R.S. Phillips: Local boundary conditions for dissipative symmetric linear differential operators. Comm. Pure Appl. Math., v 13, p. 427, 1960.

42. --------- and B. Wendroff: Systems of conservation laws. Comm. Pure Appl. Math., v 13, p. 217, 1960.

43. ----------: Difference schemes for hyperbolic equations with high order of accuracy. Comm. Pure Appl. Math., v 17, p. 381, 1964.

44. Mikhlin, S.G. and K.L. Smolitskiy: Approximate methods for solution of differential and integral equations. American Elsevier, New York, 1967.

45. Morton, K.W.: On a matrix theorem due to H.O. Kreiss. Comm. Pure Appl. Math., v 17, p. 375, 1964.

46. von Neumann, J. and R.D. Richtmyer: A method for the numerical calculation of hydrodynamical shocks. J. Appl. Phys., v 21, p. 232, 1950.

47. Quinn, B.K.: Solutions with shocks: an example of an L_1-contractive semigroup. Comm. Pure Appl. Math, v 26, p. 125, 1971.

48. Recker, W.W.: A numerical solution of three-dimensional problems in dynamic elasticity. Ph.D. Thesis, Dept. Civil Eng., Carnegie Mellon University, 1967.

49. Reissner, E.: On a variational theorem in elasticity. J. Math. and Phy., v 29, p. 90, 1950.

50. ----------: (article in) Problems of continuum mechanics. SIAM, 370, 1961.

51. Richardson, D.J.: The solution of two-dimensional hydrodynamic equation by the method of characteristics. Methods in Computational Physics, v 3, Academic Press, 1964.

52. Richtmyer, R.D. and K.W. Morton: Difference Methods for Initial-Value Problems, Interscience, New York, 1967.

54. Sarason, L.: On hyperbolic mixed problems. ARMA, v 18, p. 310 1964.

53. Lions, J.L.: (chapter in) Mathematics Applied to Physics, ed. by E. Roubine, Springer-Verlag, New York, 1970.

55. Sauerwein, H.: Anisotropic waves in elasto-plastic soils. Intern. J. Eng. Sci., v 5, p. 455, 1967.

56. Sokolnikoff, I.S.: Mathematical theory of elasticity. McGraw Hill, 1956.

57. Strang, W.G.: Difference methods for mixed boundary-value problems. Duke Math. J., v 27, p. 221.

58. ----------: Accurate partial difference methods I: linear Cauchy problems. ARMA v 27, p. 221, 1963.

59. ----------: On the construction and comparison of difference schemes. SIAM J. Numer. Anal., v 5 n 3, p. 506, 1968.

60. Sunouchi, H.: Convergence of semi-discrete difference schemes of abstract Cauchy problems. Tohoko Math. J., v 22, p. 394, 1970.

61. Thomee, V.: A difference method for a mixed boundary problem for symmetric hyperbolic systems. ARMA, v 13, p. 122, 1965.

62. Thompson, R.J.: Difference approximations for inhomogeneous and quasi-linear equations. J. Soc. Indust. Appl. Math., v 12, n 1, p. 189, 1964.

63. Trotter, H.F.: Approximation of semigroups of operators. Pacif. J. Math., v 8, p. 887, 1958.

64. Wilkins, M.L.: Finite difference scheme for calculating problems in two space dimensions and time. J. Comp. Phys., v 5, p. 406, 1970.

65. Wilcox, C.H.: Wave operators and asymptotic solutions of wave propagation problems of classical physics. ARMA, v 22, p. 37, 1966.

66. ----------: Steady-state wave propagation in homogeneous anisotropic media. ARMA, v 25, p. 201, 1967.

67. Yanenko, N.N.: The method of fractional steps. Springer-Verlag, New York, 1971.

68. Yosida, K.: Functional analysis. Springer-Verlag, New York, 1968.

69. Zlamal, M.: On the finite element method. Numer. Math., v 12, p. 394, 1968.

70. Brenner, P. and V. Thomee: Contractive difference schemes for symmetric hyperbolic systems. Math. Comp., v 25, n 114, p. 205, 1971.

71. Zwas, G. and S. Abarbanel: Third and fourth order accurate schemes for hyperbolic equations of conservation law form. Math. Comp., v 25, n 114, p. 229, 1971.

72. Ciment, M.: Stable difference schemes with uneven mesh spacings. Math. Comp., v 25, n 114, p. 219, 1971.

APPENDIX

A.1 Refinement of Strang Scheme II.3.13. Probably one of the most difficult plane problems to solve using a difference scheme with a square mesh is that of a point source. If the point source has axial symmetry then cylindrical coordinates should be used enabling a reduction to one spatial dimension - a much simpler, almost trivial problem compared to the two dimensional one. If the point source does not have axial symmetry the signals from it still propagate out in concentric circles (assuming the medium to be isotropic and homogeneous). The difficulty lies in capturing a circular pattern with one made up of squares. If a difference scheme can solve this problem reasonably well then it should be able to solve a problem with an arbitrarily shaped wave front, since each point on the front can be considered as a point source. In the numerical experiments problems with the initial data distributed in some sort of front seemed to converge more rapidly as the mesh was refined than those with a point source as initial data.

The problem chosen to exhibit numerical results is depicted in Figure A.1. The medium is a homogeneous, isotropic elastic body in plane strain which is ten units square. Only one-quarter of the body was modelled since symmetry about the x-axis and asymmetry about the y-axis was utilized. The initial data is taken as a unit velocity in the x-direction at the center, all other initial

data zero.

The material constants for the elastic medium are chosen as $\lambda = 0.5$, $\mu = 0.25$ and $\rho = 1.0$. Using these constants the dilatational wave speed (I.5c) is 1.0 and the shear wave speed 0.5. The optimum time step (numerically equal to the mesh spacing) is used in all cases.

Results are presented in Figure A.2 for the stress T^{xx} at point o in Figure A.1 for the Strang scheme II.3.13. This is the worst possible direction for this scheme; a 45 degree direction from the data. The convergence is faster at the points a and b in Figure A.1. The results are presented in the form of the history at the point.

In Figure A.2 plots are presented for square meshes ranging from 6 by 6 (a total of 36 points) to 81 by 81 (a total of 7561 points) for the problem in Figure A.1. The coarsest mesh thus has a mesh spacing of 1.0 and the finest a spacing of 0.0625. The spacing is the same in both the x and y directions.

For the case of a 41 by 41 mesh (computation at 40 x 40 = 1600 points) the storage required is 56100_8 (14246_8 of this is for the program). The running time for 80 time steps on a CDC 6400 is 150_{10} seconds.

The degree of difficulty of the problem is confirmed by the numerical expenditure required.

Although a finite problem is modeled, it can be thought of as an infinite medium in the following way. No care was taken to model the boundary, in fact the solution is simply set to zero initially at all points

on the boundary and never modified. Since the difference scheme has a finite speed of propagation and solutions are observed at point o in Figure A.1 only up to the time ($t = 3.0$) when "bad signals" are received from the boundary, the solutions will be the same at point o no matter how much further away the boundary is. This simple device breaks down, for the most part, when an implicit scheme is used.

This type of problem is probably one of the best numerical approaches to comparing schemes per se, since many other perturbing influences, e.g. boundaries, are absent.

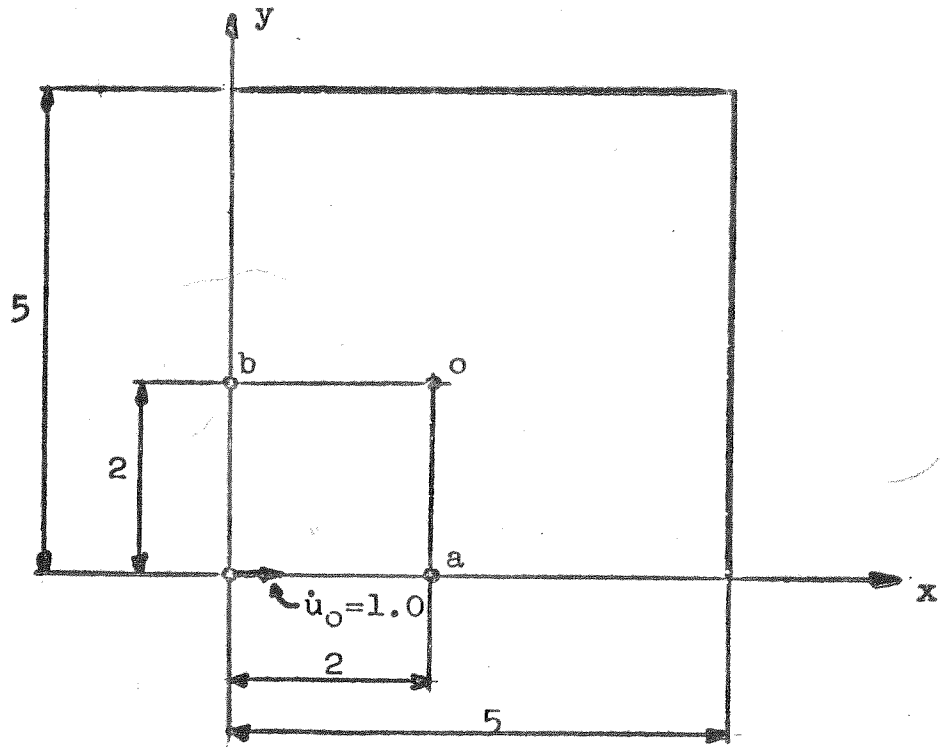


Figure A.1

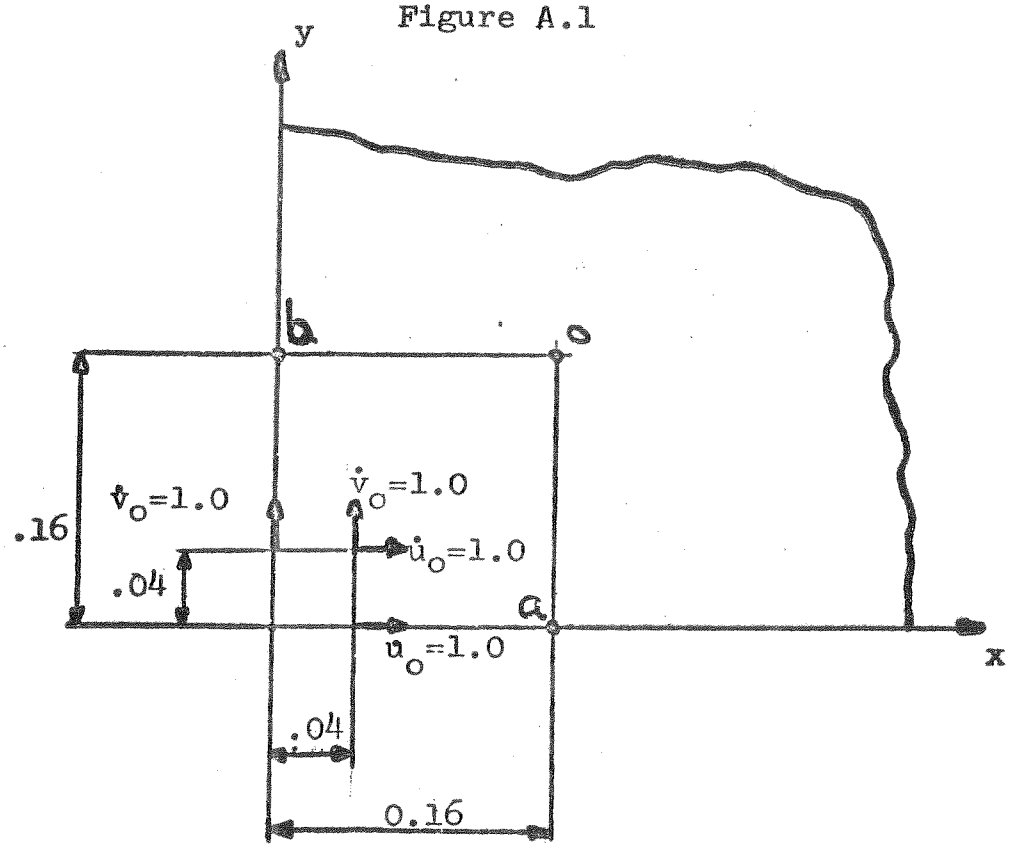


Figure A.3

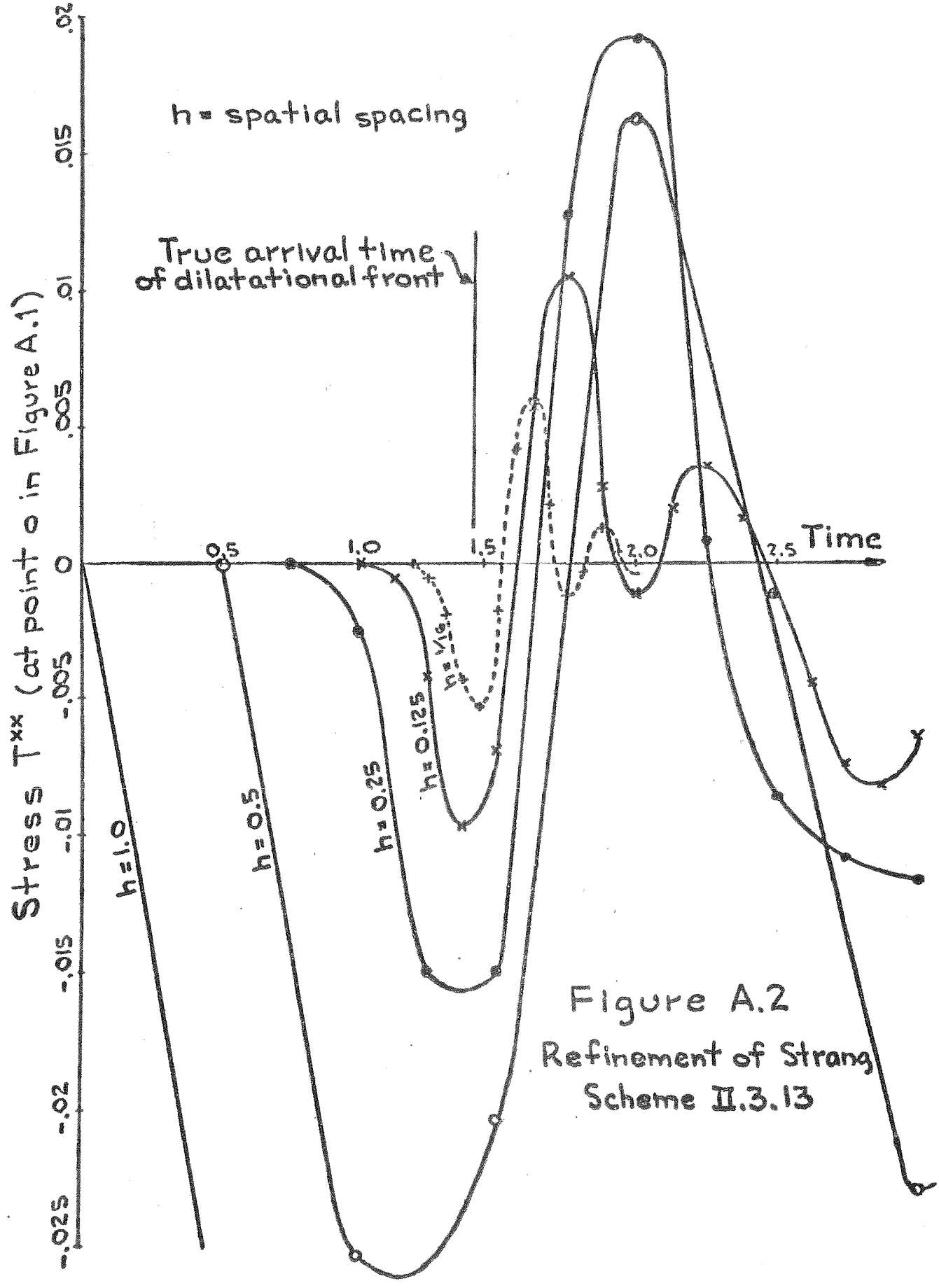


Figure A.2

Refinement of Strang Scheme II.3.13

A.2 Comparison of Strang Schemes II.3.13 and II.3.14.

Numerical experiments comparing the Strang schemes II.3.13 and II.3.14 were performed using the problem shown in Figure A.3. The medium and material constants are the same as in the problem in A.1. The nonzero initial conditions are as shown in Figure A.3 and in this case symmetry was used about both the x and y axes.

Results are presented in Figure A.4 for the stress T^{xx} at the point a in Figure A.3 for the two schemes. Again the results are in the form of the history at the point. Results from only a single mesh are given, with $\Delta x = \Delta y = 0.01$. The results are virtually identical for the two schemes at point b in Figure A.3 and, surprisingly, quite close at point o. The results at point a represent the poorest agreement between the two schemes. Because of its larger DOD (see Figure 5) scheme II.3.14 does not exhibit the sharp wave fronts in the x-direction as well as scheme II.3.13. For some reason the numerical disparity between the two schemes is greater for negative stresses than for positive ones. This is true also for T^{yy} (not shown) at this point, ($T^{xy} = 0$ at this point). The optimum time step (numerically equal to the spatial spacing) is used for both schemes.

An indication of the relative efficiency of the two schemes can be given for the case of a 37 by 37 mesh (computation at 36 x 36 = 1296 points). For scheme II.3.13 the total storage required is 50600_8 (14137_8 of

this is for the program); the comparable figures for scheme II.3.14 are 33400_8 and 14157_8. The running time for 50 time steps on a CDC 6400 was 75_{10} and 58_{10} seconds for schemes II.3.13 and II.3.14 respectively. For computational efficiency the scheme II.3.14 has a definite advantage over II.3.13.

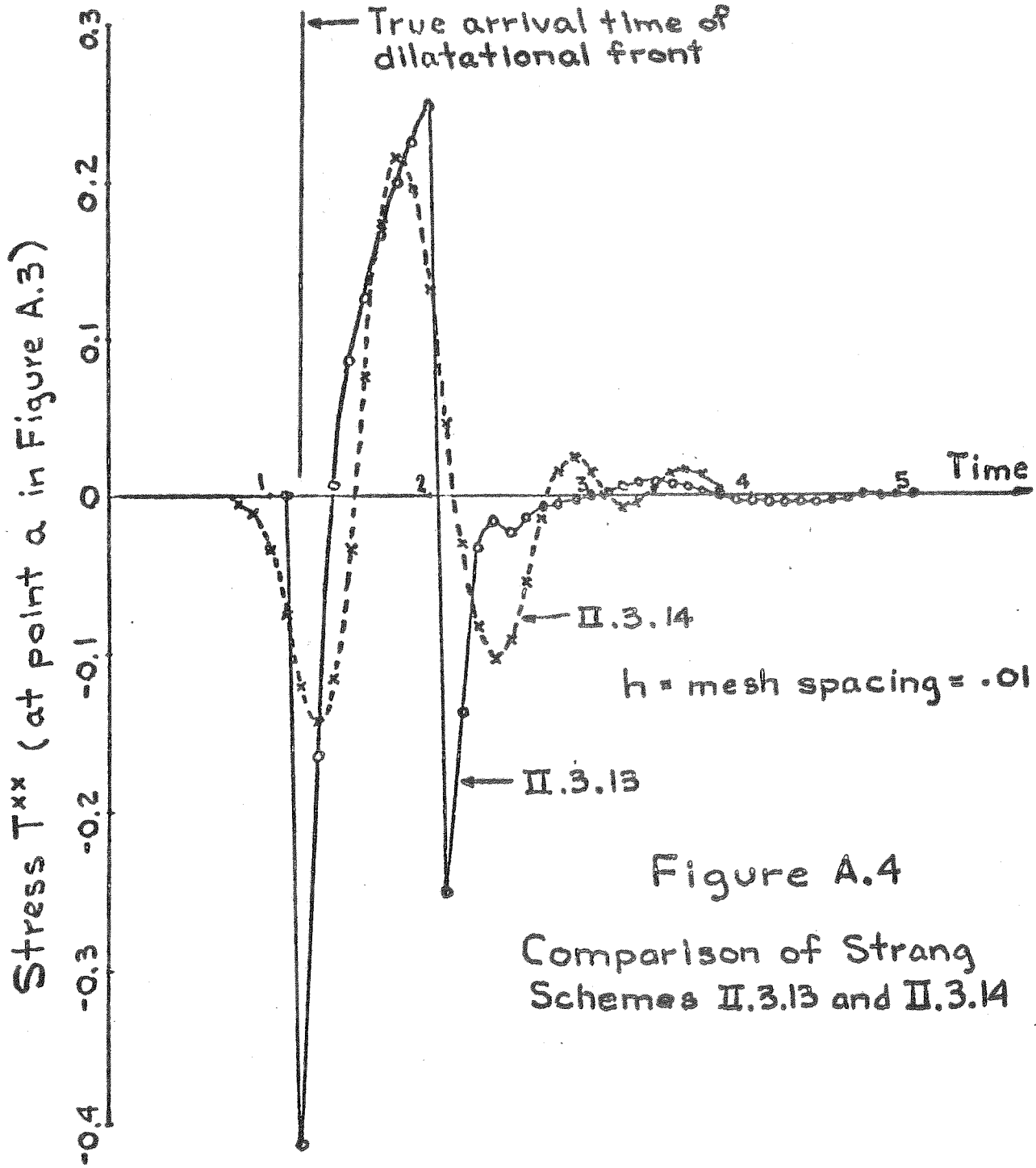


Figure A.4

Comparison of Strang Schemes II.3.13 and II.3.14

A.3 Problem with Boundary. To verify numerically the treatment of boundary conditions suggested in II.4 (one-sided differences at the boundary points) the problem described in A.1 was run with a stress-free boundary condition. The results agree with that in A.1 (as they should) up to the time when signals reflected from the boundary were observed. The optimum time step was used with scheme II.3.13. No evidence of instability was observed and the calculation was allowed to continue for 40 time steps (35 after the initial data reached the boundary). The running time was 1.26 times that of scheme II.3.13 in A.1 and 1.53 times as much total storage was required. The problem had a relatively coarse mesh (6 by 6, 11 boundary points, 25 interior points). For finer meshes (for the same problem) the figures 1.26 and 1.53 would be reduced as the ratio of boundary points to interior points was reduced. In addition, for simplicity the most efficient implementation of the scheme was not used for the points on the boundary.

A.4 Problem with Material Interface. To verify numerically the treatment of material interfaces suggested in III.3 (i.e., interpolation at an interface) a problem similar to that in A.1 was run, see Figure A.5. Only the scheme II.3.13 was used.

In Material 1, $\lambda_1 = 0.5$, $\mu_1 = 0.25$, $\rho_1 = 1.0$ and in Material 2, $\lambda_2 = 0.125$, $\mu_2 = 0.0625$, $\rho_2 = 1.0$. This means the dilatational wave speed in Material 1 is 1.0 and in Material 2 it is 0.5. According to the arguments in III.3 the mesh spacing in Material 1 is chosen twice that in Material 2, i.e., $h_1 = 1.0$ and $h_2 = 0.5$. The optimum time step was used (numerically equal to 1.0) in both regions, and no evidence of instability was observed. The history of the solution was observed at the points a and b in Figure A.5. The results at point a agreed with that in A.1 (as they should) up to the time when signals reflected from the interface were observed at point a in Figure A.5.

A second problem on the mesh in Figure A.5 was run to verify numerically that the problem of a solid/fluid interface could be solved. In Material 1 (solid) the material constants were the same as above but in Material 2 (fluid), $\lambda_2 = 0.25$, $\mu_2 = 0.0$ and $\rho_2 = 1.0$. This means the dilatational wave speed in Material 1 is 1.0 and in Material 2 the wave speed (there is only one) is 0.5.

In addition, the numerical scheme was modified so that tensile stresses (negative pressures) were precluded in Material 2 (fluid). The condition $\mu_2 = 0.0$ automatically imposed the condition $T^{xy} = 0$ as well as $T^{xx} = T^{yy}$ in the fluid.

The optimum time step was used (numerically equal to 1.0) and no evidence of instability was observed.

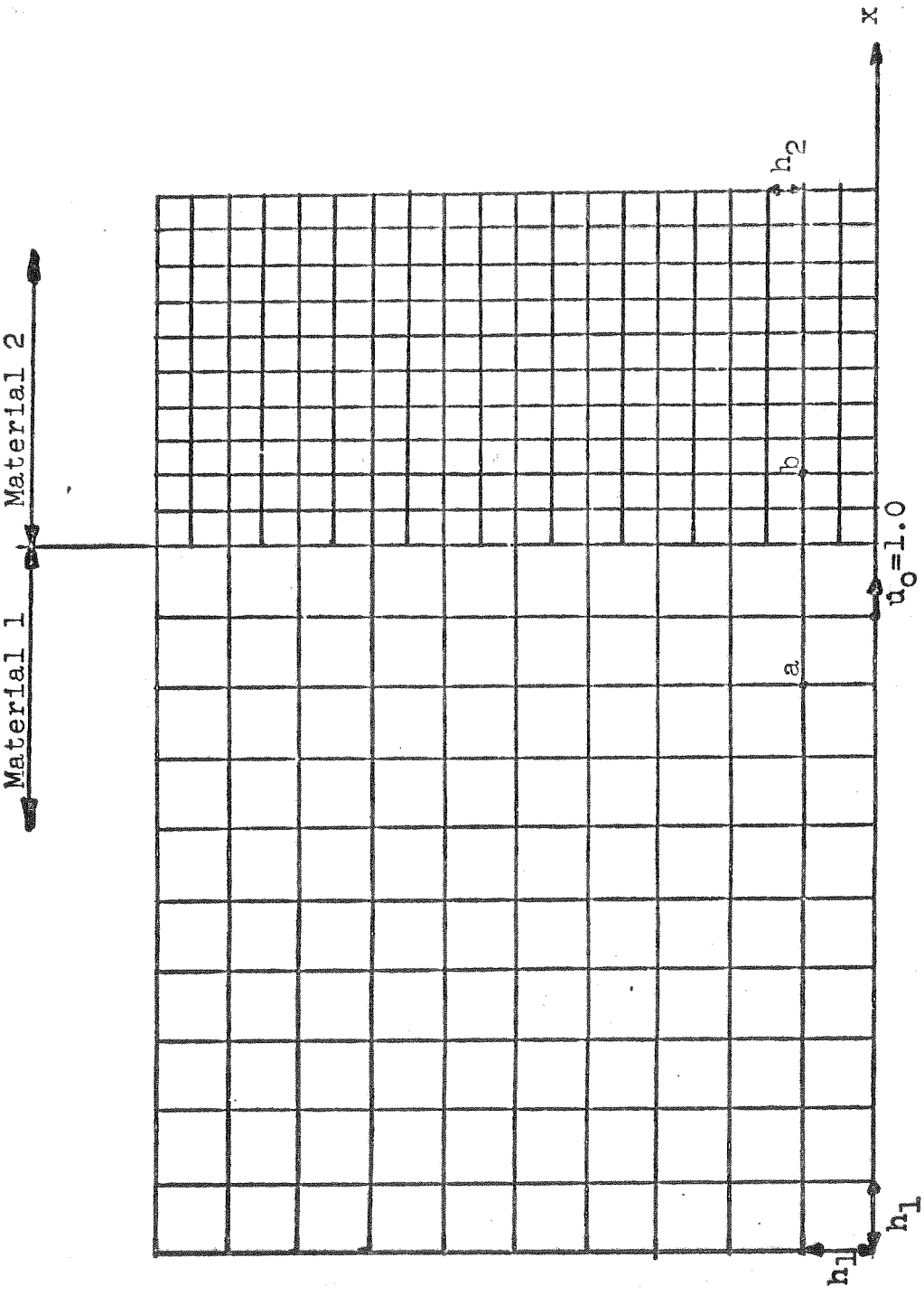


Figure A.5

A.5 Refinement of Semigroup Scheme II.6.12. Numerical experiments similar to that in Section A.1 were performed using the semigroup scheme II.6.12 on the same problem (see Figure A.1). In II.6.12, n was taken as 2 (i.e., T_h^n was defined by II.6.11). Finite differences were used to perform the spatial discretization. That is, second order accurate central differences were used to approximate the differential operators in I.5.13. As was noted previously, the resulting scheme was found to be unstable at $\frac{1}{2}$ and stable at $\frac{1}{4}$ the optimum time step. The stress T^{xx} at point o in Figure A.1 for three mesh spacings using II.6.12 are shown in Figure A.6. The mesh spacings are 0.5, 0.25 and 0.125. In addition a portion of the most refined results from Figure A.2 (scheme II.3.13) are included for comparison.

The fact that the time step used ($\frac{1}{4}$ optimum) is so suboptimal leads to the large precurser wave observed. This scheme has only first order accurate stresses although the same nine point stencil (DOD) is used as in scheme II.3.13. The numerical results indicate that scheme II.6.12 does not give as good results as scheme II.3.13 when the same mesh is used (predictable from the known truncation errors). The indication is that 2 to 4 times the mesh refinement (4 to 16 times as many points) is required for scheme II.6.12 to give the same accuracy as scheme II.3.13. This is due not only to the truncation error but also the large precurser phenomenon in II.6.12.

For the case of a 69 by 69 mesh (computation at 68 x 68 = 4624 points) the storage required was 127200_8 (13542_8 of this was for the program). The running time for 70 time steps on a CDC 6400 was 191_{10} seconds.

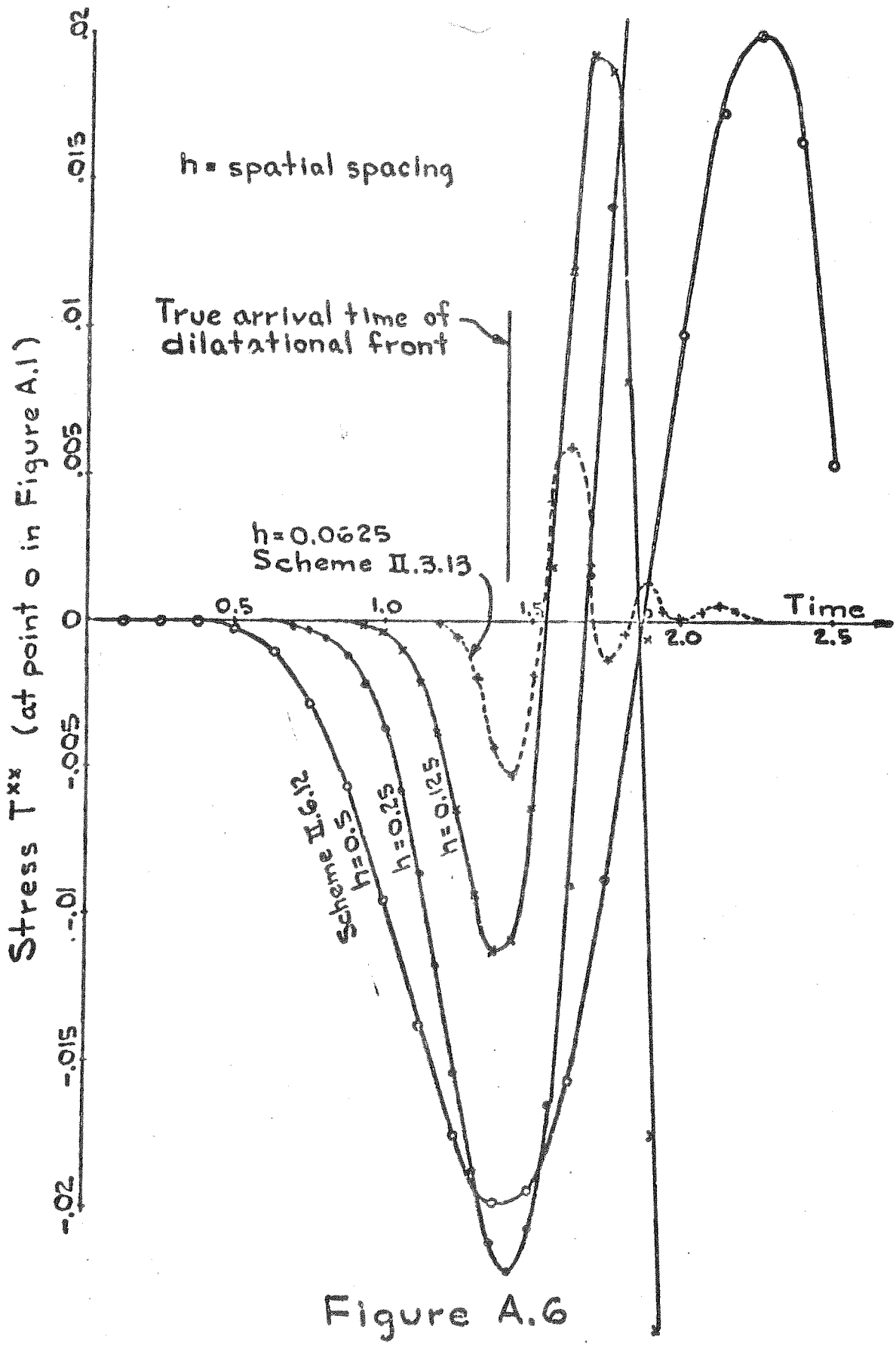


Figure A.6

Refinement of Semigroup Scheme II.6.12