ABSTRACT

An implicit finite difference scheme approximating a third order partial differential equation is examined. The scheme is derived, shown to be consistent, and its stability properties are analyzed. The partial differential equation is a parabolic approximating equation to the reduced wave equation.

Analysis of an Implicit Finite Difference Solution to an Underwater Wave Propagation Problem

D.F. St. Mary[1] and Ding Lee[2]

YALEU/DCS/TR-279

July 1983

[1]Department of Mathematics, University of Massachusetts, Amherst, MA 01003. This research was done while the author was a Visiting Fellow at the Research Center for Scientific Computation, Computer Science Department, Yale University. Research supported in part by NSF Grant PRM-8115410 and ONR contract N00014-82-K-0184.

[2]Naval Underwater Systems Center, New London Laboratory, New London, CT 06320. Research supported by ONR contract N00014-83-WR-30315 and NUSC - Independent Research.

I. INTRODUCTION:

Parabolic Equation (PE) approximations to the reduced wave equation (Helmholtz equation) are used extensively in the prediction of long range sound propagation in ocean environments. They have also been used in laser beam propagation, quantum mechanics, electromagnetic diffraction and propagation, plasma physics, optical waves, and seismic waves. Historically, the "small angle" parabolic wave equation, recognized as the standard PE, was first introduced by Tappert and Hardin[9,10] and solved by a "Split-step" technique utilizing a fast Fourier transform. A "wide angle" parabolic wave equation, which in a certain sense encompasses the standard PE, was introduced by Claerbout[2] using a rational function approximation to a square root operator. Estes and Fain[4] pursued the solution of the wide angle equation using the fast Fourier transform approach after expanding the denominator of the rational function by a series and approximating. Greene[6], and Gilbert, Lee, Botseas[5] solved the wide angle equation by implicit finite difference schemes. The implicit finite difference scheme Gilbert-Lee-Botseas used is the Crank-Nicolson Implicit Finite Difference method, identified as the "IFD" model[7]. Recently, the wide angle capability has been incorporated into the IFD model and is fully available for real applications[1]. The IFD scheme allows the numerical solution of more general problems than the Split-step Fourier algorithm in that IFD can handle boundary conditions other than Dirichlet conditions. The IFD model is increasingly being used in various applications, and thus it becomes desirable that the theoretical validity of the scheme be completely examined. Lee, Botseas, Papadakis[8] and Gilbert, Lee, Botseas[5] briefly discussed the stability of the IFD scheme, but a comprehensive examination had not yet been carried out.

It is the purpose of this paper to perform a complete analysis of the well-posedness of the IFD scheme for the solution of the wide angle ocean acoustic wave equation. In this paper we consider the wide angle equation (including the standard PE) to be a third order partial differential equation, i.e., we do not approximate the denominator of the rational function; rather, we do a careful Crank-Nicolson derivation of the IFD model making no simplifying assumptions on the coefficients in the equation. We show that the scheme is consistent with the partial differential equation, derive its discretization error, and examine its stability properties.

The paper begins with an example to demonstrate the importance of the wide angle capability. Results of this example are obtained through the application of several different methods: the Split-step Fourier algorithm, the Implicit Finite Difference (IFD) code, the "exact" reference solution of the reduced wave equation by the Fast Field Program (FFP) and the normal mode method.

II. A SHALLOW WATER SOUND PROPAGATION PROBLEM:

This problem is presented to show the importance of the wide angle capability, and also to show the accuracy of the Crank-Nicolson finite difference solution. In Figure 1, the sound speed profile is described where the sound speed is 1500 m/s in the water column and 1590 m/s in the bottom.

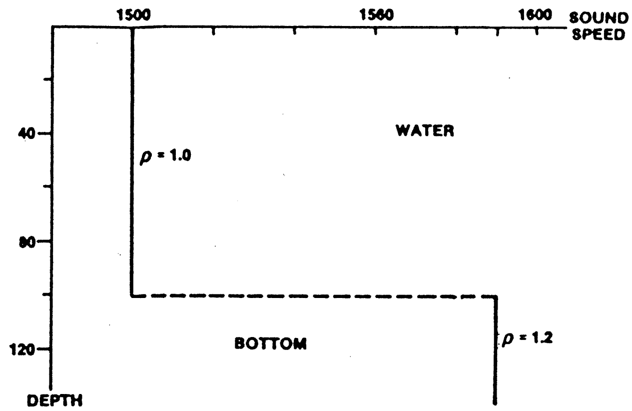


Figure 1: Sound Speed Profile

This problem has a range-independent environment and consists of an isovelocity water column over an isovelocity half-space bottom. Both the source and the receiver are placed at the same depth, 90.5 m below the surface. The source frequency is 250 Hz. There is no attenuation in the water, but an attentuation of 0.5 dB/λ in the bottom. There is a density change from 1.0 gm/cm^3 to 1.2 gm/cm^3 in the bottom. The propagation loss was calculated up to 10 km. The wide angle capability is important because the maximum angle of propagation is approximately 19°.

We use the fast field[3] (FFP) exact solution as a benchmark solution, and a normal mode[3] solution is used as a reference solution. For purposes of this discussion the Split-step small angle and the IFD small angle are identical. From Figure 2 below, it is very clear that without the wide angle capability (denoted by "IFD Small Angle"), a phase error is evident. However, with the wide angle capability (denoted by "IFD Wide Angle"), the Crank-Nicolson finite difference solution gives excellent agreement with the benchmark exact solution.

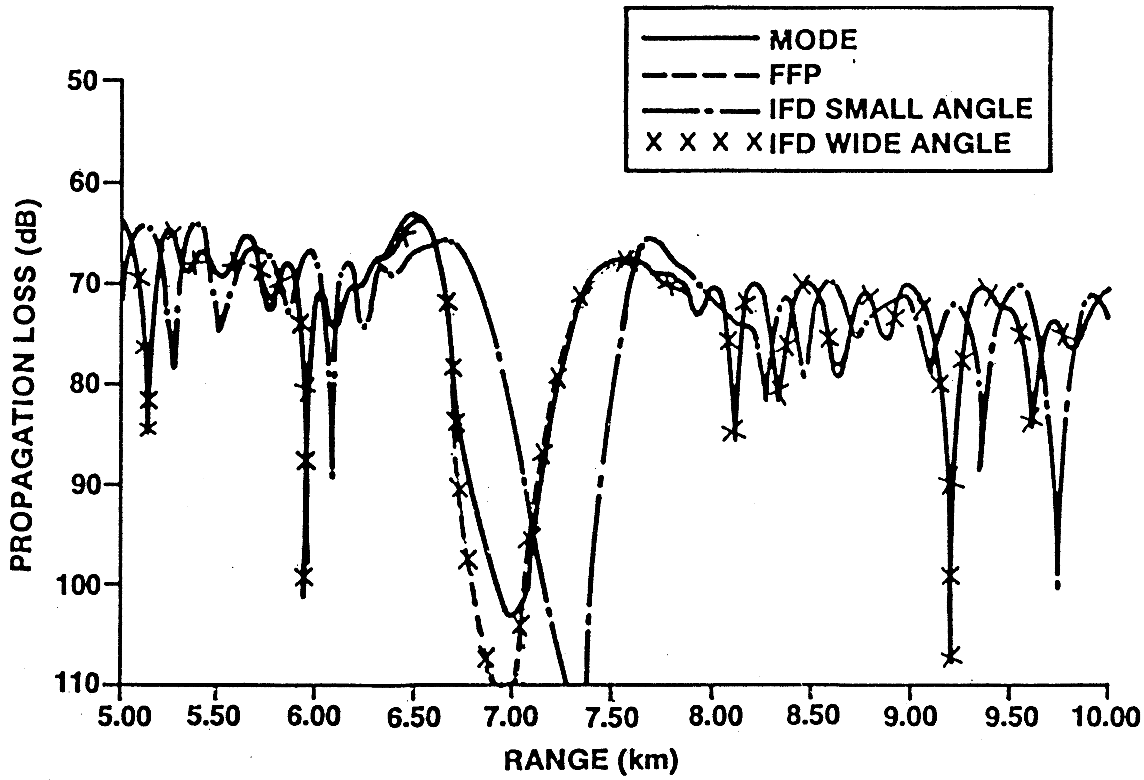


Figure 2: Comparison of Solutions

III. DERIVATION OF IFD:

The parabolic approximation is usually written in the form

$$\begin{aligned} \partial u/\partial r &= ik_0[(1 + pL)/(1 + qL) - 1]u, \\ Lu &= [(n^2(r,z) - 1) + (1/k_0^2)(\partial^2/\partial z^2)]u \end{aligned} \tag{1}$$

where p, q, k_0 are real parameters, $p \neq q$, $k_0 \neq 0$, $i = \sqrt{-1}$ and $n(r,z)$ is a real valued function of the real variables r, z. The function $n(r,z)$ represents the index of refraction of the medium and k_0 an average wave number. The choice of the parameters $p = 1/2$, $q = 0$ yields the "small angle" approximation of Tappert, and the choice $p = 3/4$, $q = 1/4$ yields the "wide angle" equation due to Claerbout.

For our purposes we choose to write (1) as a third order partial differential equation

$$(1 + qL)(\partial/\partial r)u = ik_0(p - q)Lu \ . \tag{2}$$

The boundary conditions associated with a solution $u = u(r,z)$ of (1) and (2) are

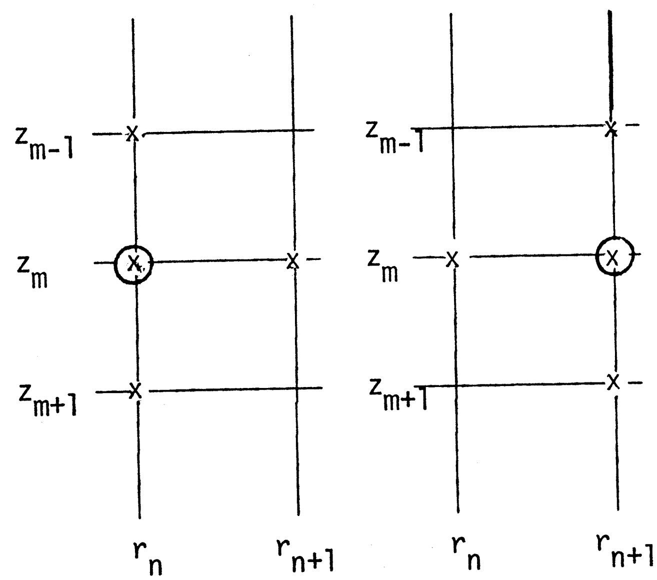
$$\begin{aligned} u(r_0,z) &= f(z), && 0 < z < B \\ u(r,0) &= 0, && r \geq r_0 \\ \alpha u(r,B) + \beta u_z(r,B) &= \gamma, && \beta \neq 0, \quad r \geq r_0, \end{aligned} \tag{3}$$

α,β,γ are real, $r_0 \geq 0$, here $z = 0$ and $z = B$ represent the surface and bottom, respectively, of a flat bottom ocean wave guide, $f(z)$ is the source, and r is the range of propagation.

We choose the standard grid on the wave guide, $h = \Delta z$, $M\Delta z = B$, M an integer, $k = \Delta r$, and for $z_m = mh$, $r_n = r_0 + nk$, m, n integers, $u(r_n,z_m) = u_m^n$. We shall use the letter "n" in two different ways, as a counter on the range variable r, and to designate the index of refraction $n(r,z)$, the context will make it clear which is intended in each case. A standard way in which the Crank-Nicolson approximation is derived for traditional parabolic partial differential equations is to take the average of the classic explicit (forward) difference approximation and the (backwards) implicit approximation. In order to motivate the application of this procedure to (2) we shall briefly describe its application to a parabolic equation in standard form, namely the standard equation PE,

$$u_r = cu + du_{zz}, \qquad c = ik_0(n^2 - 1)/2, \quad d = i/2k_0 \ . \tag{4}$$

Consider the two stencils



The first of these is used to make the forward approximation based at the point (r_n, z_m) and the second to make the backward approximation based at (r_{n+1}, Z_m). The difference equations are

$$[u_m^{n+1} - u_m^n]/k = c_m^n + u_m^n + [u_{m+1}^n - 2u_m^n + u_{m-1}^n]/h^2, \tag{5a}$$

and

$$[u_m^{n+1} - u_m^n]/k = c_m^{n+1} \; u_m^{n+1} + d[u_{m+1}^{n+1} - 2u_m^{n+1} + u_{m-1}^{n+1}]/h^2. \tag{5b}$$

Note that the left hand sides of these equations are the same. The Crank-Nicolson approximation to (4) is obtained on taking $[(5a) + (5b)]/2$.

In order to begin to carry out this development for (2) we need to define the forward and backward discretizations of $\frac{\partial^2}{\partial z^2}\left(\frac{\partial u}{\partial r}\right)$ associated with the two stencils. These are the standard centered difference in z and forward (backward) difference in r combined in a natural manner keeping in mind the base point of the stencil in each case. The two difference approximations are equal, as above, and have the value

$$\{(u^{n+1}_{m+1} - 2u^{n+1}_m + u^{n+1}_{m-1})/h^2 - (u^n_{m+1} - 2u^n_m + u^n_{m-1})/h^2\}/k \ .$$

Henceforth, we shall use the notation $(\delta^2 u)^n_m = u^n_{m+1} - 2u^n_m + u^n_{m-1}$. It is not difficult to prove that for arbitrary sufficiently differentiable functions $\phi(r,z)$, the truncation error of the above approximation is given by:

$$(\phi_{rzz})^n_m - \frac{\{(\delta^2\phi)^{n+1}_m - (\delta^2\phi)^n_m}{kh^2} = -(\phi_{rrzz})^n_m \frac{k}{2} - (\phi_{r4z})^n_m \frac{h^2}{12} - (\phi_{rr4z})^n_m \frac{kh^2}{24} + O(k^2 + h^4) \quad (6)$$

as $k \to o$, $h \to o$ independently of the manner in which h, k approach zero. The analogues of (5a) and (5b) are then determined to be

$$[1 + q((n^2)^n_m - 1)](u^{n+1}_m - u^n_m)/k + q[(\delta^2 u)^{n+1}_m - (\delta^2 u)^n_m]/(kk_o^2h^2)$$
$$= ik_o(p - q)[((n^2)^n_m - 1)u^n_m + (\delta^2 u)^n_m/(k_o^2h^2)], \quad (7a)$$

$$[1 + q((n^2)^{n+1}_m - 1)]\,(u^{n+1}_m - u^n_m)/k + q[(\delta^2 u)^{n+1}_m - (\delta^2 u)^n_m]/(kk_o^2h^2)$$
$$= ik_o(p - q)[((n^2)^{n+1}_n - 1)u^{n+1}_m + (\delta^2 u)^{n+1}_m/(k_o^2h^2)] \ . \quad (7b)$$

Now taking the average of (7a), (7b) and simplifying one obtains the Crank-Nicolson-like difference equation approximation to (2),

$$(\bar{b}/h^2)u^{n+1}_{m-1} + (\overline{a1}^n_m - 2(\bar{b}/h^2)u^{n+1}_m + (\bar{b}/h^2)u^{n+1}_{m+1}$$
$$= (b/h^2)u^n_{m-1} + (a0^n_m - 2(b/h^2))u^n_m + (b/h^2)u^n_{m+1} \quad , \quad (8)$$

where

$$b = q/k_o^2 + ik(p - q)/2k_o,$$
$$a1 = a1(r,z;k) = 1 + q[(n^2(r,z) + n^2(r + k,z))/2 - 1]$$
$$+ ikk_o(p - q)(n^2(r + k,z) - 1)/2 \qquad (8a)$$
$$a0 = a0(r,z;k) = 1 + q[(n^2(r,z) + n^2(r + k,z))/2 - 1]$$
$$+ ikk_o(p - q)(n^2(r,z) - 1)/2 ,$$

and e.g. $\bar{b}$ is the complex conjugate of b, $(a1)_m^n = a1(r_n,z_m;k)$. Note that if n is independent of range then $a1 \equiv a0$.

We shall now consider the discretization of the boundary conditions (3). The conditions $u(r_o,z) = f(z)$ and $u(r,0) = 0$ are trivially obtained on taking $u_m^o = f(z_m)$, $m = 1,2,\ldots, M - 1$ and $u_o^n = 0$, $n = 1,2,\ldots$. The bottom boundary condition is discretized using a central difference in order to obtain a second order approximation,

$$u_z(r,B) \approx [u(r,(M + 1)h) - u(r,(M - 1)h)]/2h \quad ,$$

so the condition becomes

$$\alpha u_M^n + \beta[u_{M+1}^n - u_{M-1}^n]/(2h) = \gamma, \quad n = 1,2,\ldots \quad . \qquad (9)$$

The "fictitious" term u_{M+1}^n can clearly be expressed in terms of the "real" unknowns u_M^n, u_{M-1}^n of the problem. In order to encompass (9) into the matrix formulation of the problem, set $m = M$ in (8) and make the substitution, from (9),

$$u_{M+1}^n = u_{M-1}^n - 2\alpha h u_M^n/\beta + 2h\gamma/\beta \ ,$$

to obtain

$$2(\bar{b}/h^2)u_{M-1}^{n+1} + (\overline{a1}_M^n - 2(\bar{b}/h^2)(1 + \alpha h/\beta))u_M^{n+1}$$

$$= 2(b/h^2)u_{M-1}^n + (a0_M^n - 2(b/h^2)(1 + \alpha h/\beta))u_M^n + g_{h,k} \qquad (10)$$

$$g_{h,k} = (b - \bar{b})2\gamma/\beta h = 2i\gamma k(p - q)/\beta h k_o .$$

Taking (10) into consideration and the surface boundary condition mentioned earlier, the system (8) can be expressed in the form

$$\begin{bmatrix} \overline{A1}_1^n & \bar{B} & & & \\ \bar{B} & \overline{A1}_2^n & \bar{B} & & \\ & \cdot & \cdot & \cdot & \\ & & \cdot & \cdot & \cdot \\ & & \bar{B} & \overline{A1}_{M-1}^n & \bar{B} \\ & & & 2\bar{B} & \overline{A1}_M^n \end{bmatrix} \begin{bmatrix} u_1^{n+1} \\ u_2^{n+1} \\ \cdot \\ \cdot \\ u_{M-1}^{n+1} \\ u_M^{n+1} \end{bmatrix} = \begin{bmatrix} A0_1^n & B & & & \\ B & A0_2^n & B & & \\ & \cdot & \cdot & \cdot & \\ & & \cdot & \cdot & \cdot \\ & & B & A0_{M-1}^n & B \\ & & & 2B & A0_M^n \end{bmatrix} \begin{bmatrix} u_1^n \\ u_2^n \\ \cdot \\ \cdot \\ u_{M-1}^n \\ u_M^n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \\ g_{h,k} \end{bmatrix} \qquad (11)$$

$A1_m^n = a1_m^n - 2(b/h^2)$, $m = 1,\dots,M - 1$; $A1_M^n = a1_M^n - 2(1 + \alpha h/\beta)b/h^2$;

$A0_m^n = a0_m^n - 2(b/h^2)$, $m = 1,\dots,M - 1$; $A0_M^n = a0_M^n - 2(1 + \alpha h/\beta)b/h^2$;

$B = b/h^2$.

Let $\bar{D}_1^n$, D_0^n be the $M \times M$ diagonal matrices having diagonal entries $(\overline{a1}_1^n, \overline{a1}_2^n, \ldots, \overline{a1}_{M-1}^n, \overline{a1}_M^n)$ and $(a0_1^n, a0_2^n, \ldots, a0_{M-1}^n, a0_M^n)$ respectively, and let T, an $M \times M$ tridiagonal matrix, and u^n be given by the following:

$$
T = \begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & & & -1 & 2 & -1 \\ & & & & -2 & 2(1 + \frac{\alpha h}{\beta}) \end{bmatrix} \qquad u^n = \begin{bmatrix} u_1^n \\ u_2^n \\ \cdot \\ \cdot \\ \cdot \\ u_{M-1}^n \\ u_M^n \end{bmatrix} .
$$

Then the system (11) becomes

$$
(\bar{D}_1^n - (\bar{b}/h^2)T)u^{n+1} = (D_0^n - (b/h^2)T)u^n + g, \qquad (12)
$$

g the obvious vector.

IV. CONSISTENCY:

A difference equation approximation to a partial differential equation is said to be (unconditionally) consistent with the differential equation if the difference equation approaches the differential equation as the mesh size approaches zero, independently of the manner in which the mesh size approaches zero. More precisely, (8) is consistent with (2) if

$$\left| \frac{1}{k}\left\{\frac{\bar{b}}{h^2}\phi_{m-1}^{n+1} + (\overline{a1}_m^n - 2\frac{\bar{b}}{h^2})\phi_m^{n+1} + \frac{\bar{b}}{h^2}\phi_m^{n+1} - \frac{b}{h^2}\phi_{m-1}^n - (\overline{a0}_m^n - 2\frac{b}{h^2})\phi_m^n - \frac{b}{h^2}\phi_{m+1}^n\right\} - \left\{(1 + qL)\frac{\partial\phi}{\partial r} - ik_0(p - q)L\phi\right\}_m^n \right| \tag{13}$$

approaches zero as $h, k \to 0$, independently of the manner in which h,k approach zero, for arbitrary "net" functions $\phi(r,z)$ having sufficient differentiability. The factor $1/k$ is present since in the derivation of (8) we previously cleared the k from the denominator. In order to help simplify (13) we shall express, a0, a1, and b in terms of their constituent parts. Let

$$Ra = Ra(r,z;k) = 1 + q[(n^2(r,z) + n^2(r + k,z))/2 - 1],$$

$$Ia = Ia(r,z) = ik_0(p - q)(n^2(r,z) - 1)/2,$$

$$Rb = q/k_0^2,\ Ib = i(p - q)/2k_0,$$

$$c = c(r,z;k) = q[n^2(r + k,z) - n^2(r,z)]/2,$$

then

$$a0(r,z;k) = Ra(r,z;k) + kIa(r,z),$$

$$\overline{a1}(r,z;k) = Ra(r,z;k) - kIa(r + k,z),\ b = Rb + kIb,$$

$$(1 + qL)\phi_r = (Ra - c)\phi_r + Rb\phi_{rzz}, \quad \text{and}$$

$$ik_0(p - q)L\phi = 2(Ia\phi + Ib\phi_{zz}) \ .$$

Observe that the standard Taylor approximation applied to n^2 in the r variable yields $c(r,z;k) = q[n_r^2(r,z)k + O(k^2)]/2$. It follows that (13) may be expressed in the form

$$\left| Ra_m^n \left\{ \frac{\phi_m^{n+1} - \phi_m^n}{k} - (\phi_r)_m^n \right\} + (Rb - kIb) \left\{ \frac{(\delta^2\phi)_m^{n+1} - (\delta^2\phi)_m^n}{kh^2} \right\} + \left\{ c_m^n \right\} (\phi_r)_m^n \right.$$

$$- Ia_m^n \left\{ \phi_m^{n+1} - \phi_m^n \right\} - 2Ib \left\{ \frac{(\delta^2\phi)_m^n}{h^2} \right\} - \left\{ \phi_m^{n+1} \right\} \left\{ Ia_m^{n+1} - Ia_m^n \right\}$$

$$\left. - Rb(\phi_{rzz})_m^n + 2Ib(\phi_{zz})_m^n \right| .$$

Now each term appearing in brackets, { --- }, can be expanded using a standard Taylor approximation, the centered difference approximation, or (6) yielding

$$\left| Ra_m^n\{(\phi_{rr})_m^n k/2 + O(k^2)\} + (Rb - kIb)\{(\phi_{rzz})_m^n + (\phi_{rrzz})_m^n\, k/2 + O(h^2 + k^2)\} \right.$$

$$+ \{q[(n_r^2)_m^n\, k/2 + O(k^2)\}(\phi_r)_m^n - Ia_m^n\{(\phi_r)_m^n k + O(k^2)\} - 2Ib\{(\phi_{zz})_m^n + O(h^2)\}$$

$$\left. -\{\phi_m^n + O(k)\}\{(Ia_r)_m^n k + O(k^2)\} - Rb(\phi_{rzz})_m^n + 2Ib(\phi_{zz})_m^n \right|$$

$$= \left| (k/2)\{Ra\phi_{rr} + Rb\phi_{rrzz} + q(n^2)_r\phi_r - 2(Ia\phi_r + Ib\phi_{rzz}) - 2Ia_r\phi\}_m^n + O(h^2 + k^2) \right|$$

$$= \left| (k/2)\{\partial[(1 + qL)\phi_r - ik_o(p - q)L\phi]/\partial r\}_m^n + O(h^2 + k^2) \right| , \qquad (14)$$

where the last equality uses the fact that $kc\phi_{rr} = O(k^2)$. It follows immediately from the equality of (13) and (14) that the range dependent index of refraction case, $n = n(r,z)$, Crank-Nicolson difference scheme (8) is unconditionally consistent with the partial differential equation (2).

Further, the truncation error or local discretization error can be obtained as the magnitude of the difference, at a point (r_n,z_m), between the differential equation and the difference equation both evaluated with the net function $\phi = u$ the exact solution of the partial differential equation. Again, the equality of (13) and (14) yields immediately that the local discretization error of (8) is $O(h^2 + k^2)$.

V. STABILITY:

We now turn to the question of the stability of the scheme. We shall use the matrix system (12) which encompasses both the system (8) and the boundary conditions. The system is said to be stable if an error (round-off, etc.) made at the n^{th} step does not magnify uncontrolled in its propagation to the $(n + j)^{th}$. In our case, this translates into showing that if (12) is written in the form $u^{n+1} = Bu^n + g$, B an $M \times M$ matrix, g an M-vector, then (assuming for the moment that B is constant) the eigenvalues of B are less than or equal to unity in magnitude.

The system (12) can be transformed so as to replace the nonsymmetric matrix T with a symmetric one. Let P be the $M \times M$ diagonal matrix having diagonal entries $[1,---,1,1/\sqrt{2}]$ then $S = PTP^{-1}$ is the symmetric tridiagonal matrix having exactly the same entries as T except in the lower right 2×2 block, in those positions S is of the form,

$$S(\text{lower block}) = \begin{bmatrix} 2 & -\sqrt{2} \\ -\sqrt{2} & 2(1 + \alpha h/\beta) \end{bmatrix} .$$

Since D_1, D_0, P are diagonal matrices and hence commute, (12) can be written in the form

$$P^{-1}(\bar{D}_1^n - (\bar{b}/k^2)S)Pu^{n+1} = P^{-1}(D_0^n - (b/h^2)S)Pu^n + g,$$

and thus we have a symmetric problem

$$(\bar{D}_1^n - (\bar{b}/h^2)S)v^{n+1} = (D_0^n - (b/h^2)S)v^n + Pg \tag{15}$$

where $u^n = P^{-1}v^n$. It follows that the matrix B (actually B^n) alluded to earlier can be represented by

$$B^n = (\bar{D}_1^n - (\bar{b}/h^2)S)^{-1}(D_0^n - (b/h^2)S),$$
$$v^{n+1} = B^n v^n + g^n. \tag{16}$$

Clearly, system (12) is stable if and only if system (16) is stable.

We need first to show that $(\bar{D}_1^n - (\bar{b}/h^2)S) \equiv A$, n is fixed but arbitrary, is nonsingular. If there exists an M-vector γ, $\gamma = (\gamma_1, \ldots, \gamma_M)$, such that $A\gamma = 0$, i.e. $\gamma^* S\gamma = (h^2/\bar{b})\gamma^* \bar{D}_1^n \gamma$, then the left side is real since S is symmetric and the imaginary part of the right side is a sum, on m, of terms of the form $[k(p - q)/(2k_0)]\{1 + q[n^2(r,z_m) - n^2(r + k,z_m)]/2\}|\gamma|^2$. We would like to conclude that the imaginary part is never zero, and thus that $\gamma = 0$. Now $p \neq q$ is a standing hypothesis, and q is thought of as being of magnitude less than one. We now wish to call upon an assumption which is frequently made long before this point in the discussion of the whole area, namely, "the index of refraction is slowly varying in range". Invoking this assumption to obtain $|n^2(r,z_i) - n^2(r + k,z_i)| < 2$ then yields A nonsingular. We remark that (8) is the first general statement of IFD which has not utilized the prior imposition of the simplifying "slowly varying in range" assumption.

We referred above to the case in which the coefficient matrix B^n in (16) is constant; from (8a), it is clear that this occurs precisely when

$n(r,z) \equiv n(z)$, i.e. the index of refraction is independent of range. This is the so called "layered medium" case. In this case the stability would follow upon showing that the eigenvalues of B all have magnitude less than or equal to unity, but we shall pursue the general case for the moment. Doing the standard backward recursion on (16) yields

$$v^{n+1} = B^n B^{n-1} \ldots B^o v^o + B^n B^{n-1} \ldots B^1 g^o + \ldots + B^n B^{n-1} g^{n-2} + B^n g^{n-1} + g^n,$$

where v^o is the vector of "correct" initial values, and repeating the process using an initial vector v_*^o containing errors we then obtain

$$e^{n+1} = v^{n+1} - v_*^{n+1} = B^n B^{n-1} \ldots B^o (v^o - v_*^o) = B^n B^{n-1} \ldots B^o e^o, \tag{17}$$

the formula for the propagation of errors. The finite difference scheme is stable if e^{n+1} remains bounded as n increases indefinitely. Clearly the boundedness of the e^n is attained if we can show that the matrices B^n are bounded in norm by 1 for all n. It does not appear to be sufficient to simply show that the eigenvalues of each B^n are less than or equal to one (it is not the case that $\|B\| \leq 1$ when all the eigenvalues of B have magnitude less than or equal to 1, that result does hold when B is Hermitian though). If in addition to all eigenvalues having magnitude less than or equal to 1, the associated eigenvectors constitute a complete orthonormal basis then the matrix has norm less than or equal to one.

PROPOSITION 1: Let (μ_j, γ^j) satisfy $B\gamma^j = \mu_j \gamma^j$ with $|\mu_j| \leq 1$ for $j = 1,2,\ldots,M$, and let $\{\gamma^j\}$ be a complete orthonormal set, then $\|B\| \leq 1$.

PROOF: Use the definition of $\|B\|$, $\|B\| = \sup\{\|B\gamma\| : \|\gamma\| \leq 1\}$. If γ is an arbitrary complex M-vector with $\|\gamma\| \leq 1$ then γ can be expressed in the form $\gamma = \sum_{j=1}^{M} \alpha_j \gamma^j$, where $\|\gamma\|^2 = \sum_{j=1}^{M} |\alpha_j|^2$ by the orthonormality. But then

$$B\gamma = \Sigma \alpha_j \mu_j \gamma^j$$

and hence

$$\|B\gamma\|^2 = \Sigma |\alpha_j \mu_j|^2 \leq \Sigma |\alpha_j|^2 \leq 1.$$

Our objective is to show that the hypothesis of Proposition 1 is satisfied for a large class of matrices of the form (16). A property which is related to this question is that of normality of matrices - a matrix A is <u>normal</u> if it commutes with its adjoint, i.e., $A^*A = AA^*$, which holds if and only if A is similar to a diagonal matrix via a unitary matrix.

PROPOSITION 2: If a nonsingular matrix A is normal, say $P^*AP = D$, P unitary, D diagonal, then each of the eigenvalues of $A^{*-1}A$ has magnitude 1 and P diagonalizes $A^{*-1}A$.

PROOF: Since P is a unitary matrix $P^* = P^{-1}$. Thus $A = PDP^*$, $A^* = P\bar{D}P^*$, $A^{*-1} = P\bar{D}^{-1}P^*$ and hence

$$A^{*-1}A = P\bar{D}^{-1}P^*PDP^* = P\bar{D}^{-1}DP^* \quad .$$

We have shown that $A^{*-1}A$ is similar to a diagonal matrix, the diagonal elements (eigenvalues) of which are all of the form $z/\bar{z}$ for some complex number z, and hence have magnitude 1.

We remark that, since the columns of P constitute a complete orthonormal set of eigenvectors for A, the hypothesis of Proposition 1 with $B = A^{*-1}A$ is satisfied when A is normal.

The first case we choose to consider is when $n(r,z) \equiv n$ is constant, then (16) is of the form

$$B^n \equiv B = (\bar{a}I - (\bar{b}/h^2)S)^{-1}(aI - (b/h^2)S),$$

i.e., $D_1^n = D_0^n = aI$, $a = 1 + q(n^2 - 1) + ikk_0(p - q)(n^2 - 1)/2$, I the identity matrix. We take $A = (aI - (b/h^2)S)$, then $A^{*-1} = (\bar{a}I - (\bar{b}/h^2)S)^{-1}$ and it is trivial to show that A satisfies the normality property $A^*A = AA^*$. Thus, <u>if the index of refraction is constant the</u> IFD <u>scheme is unconditionally stable for all values of the parameters</u> $\alpha,\beta,\gamma,p,q,k_0$.

Next we shall consider the layered medium case, namely $n(r,z) \equiv n(z)$, i.e., the index of refraction is independent of range. In this case $D_1^n = D_0^n = D$ and

$$A = D - (b/h^2)S.$$

Since $A^* = \bar{A}$ it is trivial to show that $AA^* = A^*A$ if and only if $\overline{AA^*} = AA^*$, i.e., the imaginary parts of the elements of the product AA^* are all zero. AA^* is a five diagonal matrix and the main diagonal and the two off-off diagonals are obviously real. The off diagonal elements, H_m, $\bar{H}_m$, where

$$H_m = \overline{(a_m - 2(b/h^2))}(b/h^2) + (a_{m+1} - 2(b/h^2))\bar{b}/h^2,$$

$m = 1,2,\ldots,M-2$, and

$$H_{M-1} = \sqrt{2}(\overline{(a_{M-1} - 2(b/h^2))}(b/h^2) + (a_M - 2(1 + (\alpha h/\beta))\bar{b}/h^2) ,$$

$a_m = 1 + q(n^2(z_m) - 1) + ikk_0(p - q)(n^2(z_m) - 1)/2$ are also real when $\alpha = 0$. Thus, <u>the layered medium case of IFD is unconditionally stable when</u> $\alpha = 0$.

A slightly different kind of analysis yields some information in a range dependent case. Let $n(r,z) \equiv n(r)$, i.e., independent of depth, then $D_1^n = a1^n I$ and $D_0^n = a0^n I$, I the $M \times M$ identity matrix and hence

$$B^n = (\overline{a1}^n I - (\bar{b}/h^2)S)^{-1}(a0^n I - (b/h^2)S).$$

Now if λ, γ is an eigenvalue-eigenvector pair for S, then $(a0^n - \lambda(b/h^2))$, γ is an eigenvalue-eigenvector pair associated with $(a0^n I - (b/h^2)S)$ and similarly $1/(\overline{a1}^n - \lambda(\bar{b}/h^2))$, γ is associated with $(\overline{a1}^n I - (\bar{b}/h^2)S)$. It follows that the eigenvalue-eigenvector pairs associated with B^n are $\mu^n = (a0^n - \lambda(b/h^2))/(\overline{a1}^n - \lambda(\bar{b}/h^2))$, γ. Since the matrix S is symmetric, its eigenvalues are real and there exists a complete set of real orthonormal eigenvectors of S. The eigenvalues of S can be approximated either numerically or using Gerschgorin's theorem; thus any condition on the index of refraction for which $|\mu^n| \leq 1$ for $n = 0,1,2,\ldots$ yields a stability condition in this case.

VI. AN UNSTABLE VARIANT.

One might be tempted to consider using only the forward difference approximation scheme (7a) to solve (2). This scheme is easily written in matrix form - we discover that it is implicit, not explicit as we might have hoped. Consider (7a) expressed as follows:

$$bu_{m-1}^{n+1} + a_m^n u_m^{n+1} + bu_{m+1}^{n+1} = (b + ib')u_{m-1}^n + (a_m^n + ia'^n_m)u_m^n + (b + ib')u_{m+1}^n \ , \tag{18}$$

where

$$a_m^n = 1 + q((n^2)_n^m - 1) - 2q/k_o^2h^2, \qquad b = q/k_o^2h^2 \ ,$$

$$a'^n_m = kk_o(p-q)((n^2)_m^n - 1 - 2/k_o^2h^2), \qquad b' = kk_o(p-q)/k_o^2h^2 \ .$$

We shall examine the simplest possible case, $n(r,z) \equiv n$ a constant, $a_m^n = a$, $a'^n_m = a'$, and both the surface and the bottom boundary conditions zero, then we can legitimately apply the Fourier series method (von Neumann's method) to analyze stability. We examine solutions of (18) of the form $u_m^n = e^{\alpha(nk)}e^{i\beta(mh)}$. Substituting this expression into (17) and simplifying, one obtains the amplification factor

$$\xi = e^{\alpha k} = \frac{(a + ia') + (b + ib')e^{i\beta h} + (b + ib')e^{-i\beta h}}{a + be^{i\beta h} + be^{-i\beta h}}$$

$$= 1 + i\left(\frac{a' + b'(2\cos\beta h)}{a + b(2\cos\beta h)}\right) .$$

Clearly $|\xi| > 1$ for all combinations of the step sizes h, k, thus (7a) is "unconditionally unstable", and this is the case for all allowable choices of the parameters p, q, k_o, n.

ACKNOWLEDGEMENT:

The authors wish to express their appreciation to Martin Schultz and Youcef Saad for valuable conversations during the preparation of this paper.

REFERENCES

[1]G. Botseas, D. Lee, K. E. Gilbert, "IFD: Wide Angle Capability", Naval Underwater Systems Center Tech. Report #6905 (1983).

[2]J. F. Claerbout, Fundamentals of Geophysical Data Processing, McGraw-Hill Co., New York, 206-207, 1976.

[3]J. A. Davis, DeWayne White, R. C. Cavanagh, editors, "NORDA Parabolic Equation Workshop", Naval Ocean Research and Development Activity, TN-143, p. 53, 1982.

[4]L. E. Estes, G. Fain, "Numerical Technique for Computing the Wide-Angle Acoustic Field in an Ocean with Range-Dependent Velocity Profile", J. Acoust. Soc. Am., Vol. 62, No. 1, 33-43, 1977.

[5]K. E. Gilbert, D. Lee, G. Botseas, "Theoretical Development and Implementation of the Wide-Angle Parabolic Wave Equation", in preparation for J. Acoust. Soc. Am., 1983.

[6]R. Green, see[3], p. 80.

[7]D. Lee, G. Botseas, "IFD: An Implicit Finite Difference Computer Model for Solving the Parabolic Equation", Naval Underwater Systems Center, TR-6659, 1982.

[8]D. Lee, G. Botseas, J. S. Papadakis, Finite-Difference Solution to the Parabolic Wave Equations, J. Acoust. Soc. Am., Vol. 70, No. 3, 795-800, 1981.

[9]F. D. Tappert, "The Parabolic Approximation Method", in Wave Propagation and Underwater Acoustics, edited by J. B. Keller, J. S. Papadakis, Lecture Notes in Physics, Vol. 70, Springer-Verlag, Heidelberg, 1977.

[10]F. D. Tappert, R. H. Hardin, in "A Synopsis of the AESD Workshop on Acoustic Modeling by Non-Ray Tracing Techniques", edited by C. W. Spofford, AESD TN-73-05, 1973.