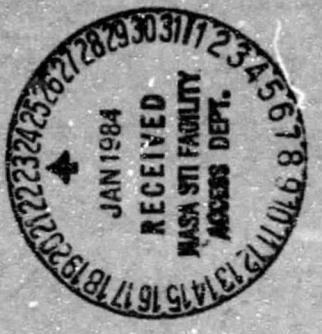
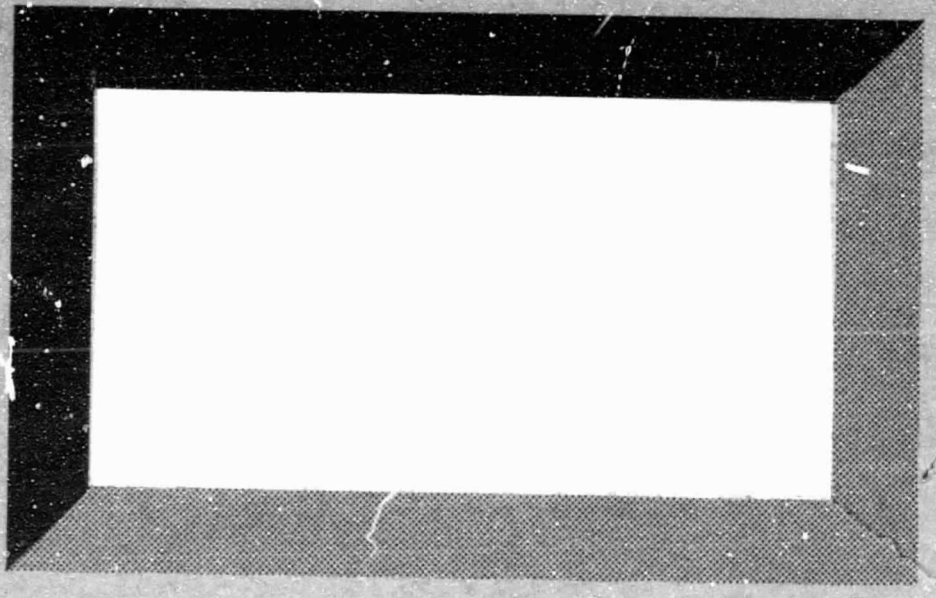
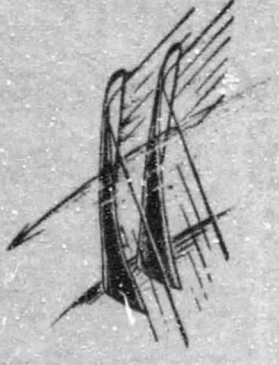






GAS TURBINE & PLASMA DYNAMICS LABORATORY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS

ASYMPTOTIC ANALYSIS OF NUMERICAL WAVE PROPAGATION IN FINITE DIFFERENCE EQUATIONS

by

Michael Giles

W. T. Thompkins, Jr.

GT&PDL Report No. 171 March 1983

This research, carried out in the Gas Turbine and Plasma Dynamics Laboratory, MIT, was supported by the NASA Lewis Research Center under Grant No. NAG3-9.

Abstract

An asymptotic technique is developed for analysing the propagation and dissipation of wave-like solutions to finite difference equations. It is shown that for each fixed complex frequency there are usually several wave solutions with different wavenumbers and the slowly varying amplitude of each satisfies an asymptotic amplitude equation which includes the effects of smoothly varying coefficients in the finite difference equations. The local group velocity appears in this equation as the velocity of convection of the amplitude. Asymptotic boundary conditions coupling the amplitudes of the different wave solutions are also derived.

A wavepacket theory is developed which predicts the motion, and interaction at boundaries, of wavepackets, wave-like disturbances of finite length. Comparison with numerical experiments demonstrates the success and limitations of the theory.

Finally an asymptotic global stability analysis is developed which gives results which agree with other stability analyses and which can be applied to a wider range of problems.

Acknowledgements

This research was supported by NASA Lewis Research Center under grant No. NAG3-9 with Technical Monitor Dr. R. V. Chima, and by a Scholarship for M. Giles from the Kennedy Memorial Trust.

Contents

1. Introduction

Consider the following very simple problem and numerical solution. The partial differential equation is

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0 \qquad (1.1)$$

where c is a positive constant. The domain considered is $0<x<1$. The initial condition is

$$u(x,0) = \exp[-200(x-0.5)^2] \cos(kx) \qquad (1.2)$$

with k=80. This form of distribution is usually called a wavepacket. The cos(kx) term defines the oscillation of a group of waves and the $\exp[-200(x-0.5)^2]$ term is an amplitude 'envelope'.

The upstream condition is

$$u(0,t) = 0 \qquad (1.3)$$

The solution of this problem is

$$u(x,t) = \begin{cases} u(x-ct,0) & ct < x < 1 \\ 0 & 0 < x < ct \end{cases} \qquad (1.4)$$

The numerical solution uses a uniform grid with computational domain $0<j<200$ and a trapezoidal scheme.

$$U_j^{n+1} - U_j^n + \frac{r}{4}\left((U_{j+1}^{n+1} + U_{j+1}^n) - (U_{j-1}^{n+1} + U_{j-1}^n)\right) = 0 \qquad (1.5)$$

where $$r = \frac{c\Delta t}{\Delta x} \qquad (1.6)$$

In this example r=1. The initial condition is

$$U_j^0 = u(x_j,0) \qquad (1.7)$$

and the upstream boundary condition is

$$U_0^n = 0 \qquad (1.8)$$

In addition a numerical boundary condition is required at the downstream boundary. For this condition space extrapolation is used.

$$U_{200}^n = U_{199}^n \qquad (1.9)$$

Figure 1 shows the numerical solution at intervals of 60 time steps with each plot drawn to the same scale. The first two plots show the initial wavepacket travelling downstream in the direction of the physical characteristic. Corresponding wavecrests are labelled a-e and it can be seen that the propagation velocity for the wave crests is greater than for the amplitude envelope. Note for example that the amplitude maximum lies approximately midway between crests b and d at n=60 but at n=120 the maximum is clearly nearer crest b. At n=180 the numerical disturbance is interacting with the downstream boundary. The solution appears to be the sum of two waves, one with the original wavelength , and one with a very much shorter wavelength. At n=240 there is a reflected wavepacket of wavelength slightly greater than 2. and the plots at n=300,360 show that this wavepacket travels back up the domain at approximately the same speed as the original wavepacket. This solution is clearly numerical and not physical since the analytic, physical solution moves from left to right across the domain and then out the downstream boundary. The analytic equation does not have any solutions with waves travelling from right to left. At n=420 the wavepacket is interacting with the upstream boundary, and at n=480 there is a reflected wavepacket with the original wavelength. This completes one cycle. If the solution was continued the wavepacket would travel down to the downstream boundary and then reflect again into a

ORIGINAL PAGE IS
OF POOR QUALITY

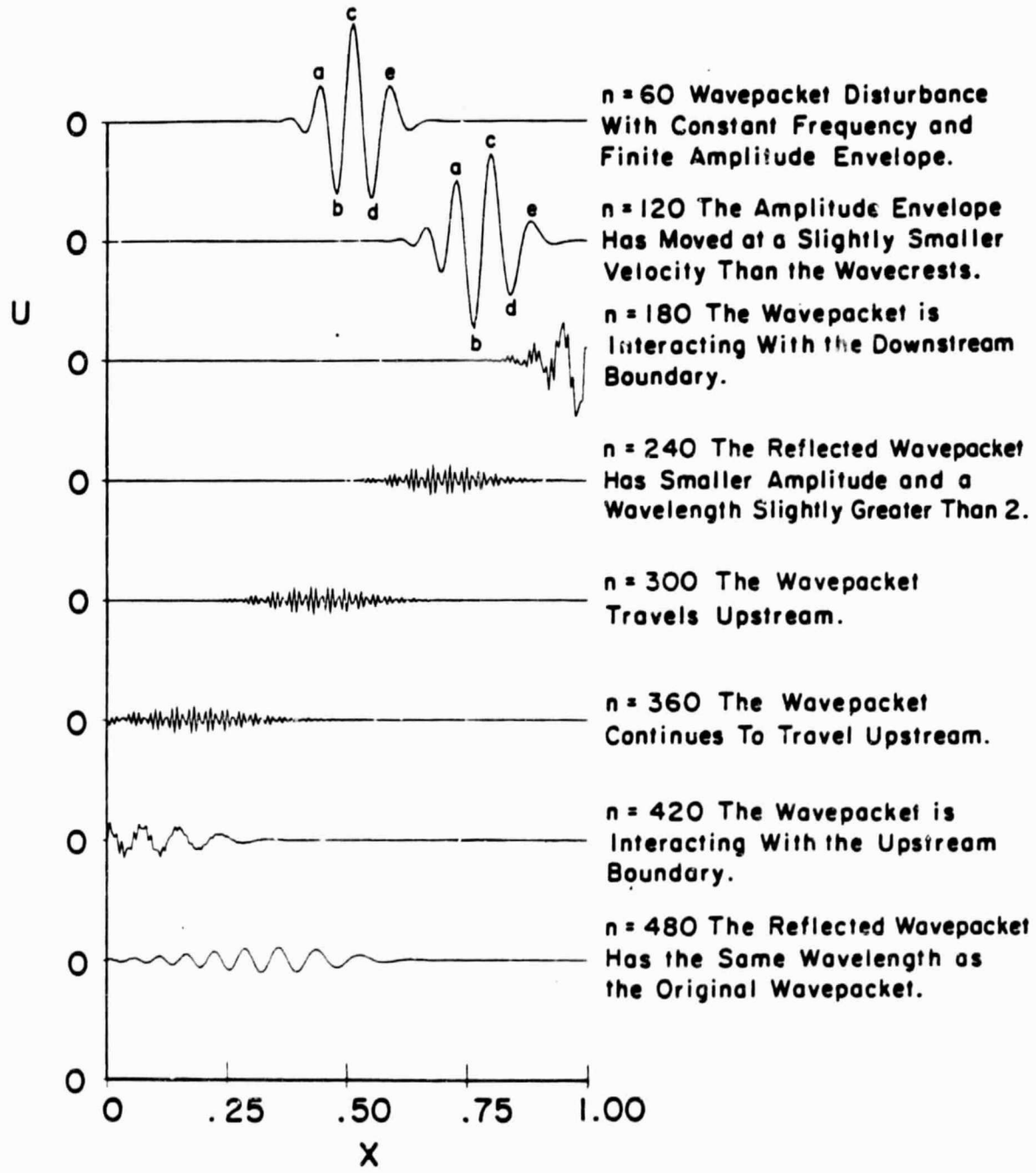


FIGURE 1. NUMERICAL SOLUTION OF CONVECTION EQUATION

wavepacket with short wavelength, and a decreased amplitude, travelling upstream.

The qualitative and quantitative prediction of the behaviour of numerical solutions in problems such as the above is one of the two objectives of this paper. The second objective is a global stability analysis incorporating boundary conditions and smoothly varying coefficients and predicting both stability and accurate asymptotic estimates of convergence rates.

To achieve these aims a technique is developed to analyse the approximate time evolution of an amplitude modulated wave, i.e. a wave with fixed frequency and a slowly varying amplitude. Chapter 2 derives the theory for partial differential equations, while chapter 3 derives the theory for finite difference equations incorporating smoothly varying coefficients and boundary conditions. In the case of dispersive, non-dissipative wave propagation, it is found that the amplitude is convected at the local group velocity, a principle which is well understood in partial differential equations.

Chapter 4 applies the theory to the motion of wavepackets which are wave-like disturbances of finite length and constant frequency such as in the earlier example. Chapters 5 and 6 derive global stability analyses with different levels of asymptotic approximation. Chapters 7-9 develop further topics and examples including comparisons between numerical experiments and theoretical predictions.

Throughout this paper a finite operator notation is used which greatly simplifies analysis and is a neccessity

for general proofs. Since there is no universally accepted standard notation Appendix 1 details the notation used.

Very little previous work appears to have been done along the lines of this paper. The concept of group velocity in partial differential equations is well understood and is explained in many texts [1,2]. The asymptotic approach of chapter 2 is not common due to the advantages of other methods but is discussed by Whitham [1]. Kentzer [3] has discussed the use of group velocity in analysing finite difference equations but does not derive a general equation for the amplitude or calculate the quantitative effects of boundary conditions. Vichnevetsky and Bowles [4] derive the the group velocity in finite difference equations using an approach which is valid only for constant coefficients. They also derive amplitude reflection coeffficients at boundaries and discuss some of the examples given in this paper.

2. Amplitude Analysis of Partial Differential Equations

2.1 Fourier Analysis

Consider a homogeneous partial differential equation

$$L\,u(x,t) = 0 \qquad -\infty < x < \infty\,,\ t > 0 \tag{2.1}$$

where L is a constant linear differential operator defined by,

$$L = \sum_{m,n} C_{mn} \left(\frac{\partial}{\partial x}\right)^m \left(\frac{\partial}{\partial t}\right)^n \tag{2.2}$$

and the coefficients C_{mn} are constants.

An eigenfunction of the operator L is a function u(x,t) satisfying

$$L\,u = \lambda u \tag{2.3}$$

where λ is a constant called the eigenvalue.

An eigenmode is a solution of the homogeneous equation (2.1) i.e. it is an eigenfunction with eigenvalue zero.

$$\frac{\partial}{\partial x} \exp[i(kx-\omega t)] = ik \exp[i(kx-\omega t)] \tag{2.4a}$$

$$\frac{\partial}{\partial t} \exp[i(kx-\omega t)] = -i\omega \exp[i(kx-\omega t)] \tag{2.4b}$$

$$\left(\frac{\partial}{\partial x}\right)^m \left(\frac{\partial}{\partial t}\right)^n \exp[i(kx-\omega t)] = (ik)^m(-i\omega)^n \exp[i(kx-\omega t)] \tag{2.4c}$$

$$\sum_{m,n} C_{mn}\left(\frac{\partial}{\partial x}\right)^m\left(\frac{\partial}{\partial t}\right)^n \exp[i(kx-\omega t)] =$$

$$\sum_{m,n} C_{mn}\,(ik)^m(-i\omega)^n \exp[i(kx-\omega t)] \qquad (2.4d)$$

Thus $\exp[i(kx-\omega t)]$ is an eigenfunction of $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial t}$, $\left(\frac{\partial}{\partial x}\right)^m\left(\frac{\partial}{\partial t}\right)^n$ and L with eigenvalues ik, $-i\omega$, $(ik)^m(-i\omega)^n$, and $\sum_{m,n} C_{mn}(ik)^m(-i\omega)^n$ respectively

Hence,

$$u(x,t) = \exp[i(kx-\omega t)] \qquad (2.5)$$

is an exact solution of (2.1) provided

$$\sum_{m,n} C_{mn}\,(ik)^m\,(-i\omega)^n = 0 \qquad (2.6)$$

This relation between k and ω is called the dispersion relation.

Examples of dispersion relations are ;

Surface waves on deep water $\omega^2 = |gk|$ (2.7a)

Acoustic waves $\omega^2 = c^2k^2$ (2.7b)

Waves propagating along a waveguide $\omega^2 = c^2(k^2 + k_0^2)$ (2.7c)

where g, c and k_0 are constants.

A general solution of (2.1) is a superposition of eigenmodes which in the case of a partial differential equation is expressed as an integral over all the wavenumbers k of the sum of all the eigenmodes with wavenumber k.

$$u(x,t) = \int_{-\infty}^{\infty} \sum_{n=}^{N} A_n(k) \exp[i(kx-\omega_n t)]\, dk \qquad (2.8)$$

If the dispersion relation is of order N in ω , i.e. it contains powers of ω up to ω^N , then

$$\omega_1(k) \ , \ \omega_2(k) \ , \ \ldots\ldots \ , \ \omega_N(k)$$

are the N values of ω which satisfy the dispersion relation for a given value of k and the $A_n(k)$ are the corresponding constant amplitudes of those eigenmodes.

If $A_n(k)$ is non-zero for all k,n then a neccessary and sufficient condition for u(x,t) to remain bounded and not increase exponentially is that each eigenmode must remain bounded. Splitting ω into its real and imaginary components gives ,

$$\omega = \omega_R + i\omega_I \tag{2.9}$$

$$\exp[-i\omega t] = \exp[-i\omega_R t + \omega_I t] \tag{2.10}$$

Thus the condition that every eigenmode remain bounded, and hence a general solution remain bounded, is $\omega_I < 0$ for all k,n.

This analysis is lacking in three respects. The first is that in some situations the initial disturbance is zero except for a finite region and one wants to know the time evolution of this disturbance, in particular the propagation velocity for the energy. The second failing is that when the initial-value problem is replaced by an initial-value / boundary-value problem with boundary conditions at x = 0,1 there is no easy way to include the effect of the boundary conditions in this stability analysis. The third failing is that $\exp[i(kx-\omega t)]$ is an eigenfunction of L only when the coefficients C_{mn} are constant. The analysis breaks down when the coefficients are non-constant.

The resolution of these problems requires the analysis of a wave of constant frequency with an amplitude which varies over a characteristic length scale much greater than the wavelength.

2.2 Asymptotic Amplitude Equation

The problem now being considered is

$$L(x)\ u(x,t) = 0 \qquad -\infty < x < \infty\ ,\ t > 0 \tag{2.11}$$

where L(x) is a non-constant linear differential operator defined by,

$$L(x) = \sum_{m,n} C_{mn}(x) \left(\frac{\partial}{\partial x}\right)^m \left(\frac{\partial}{\partial t}\right)^n \tag{2.12}$$

and the coefficients $C_{mn}(x)$ are slowly varying functions of x.

The theory calculates the approximate evolution of a wavetrain with waves of a constant frequency ω and a slowly varying amplitude, so u(x,t) is written as ,

$$u(x,t) = A(x,t)\ \exp[i\Psi(x,t)] \tag{2.13}$$

where A(x,t) is the slowly varying amplitude and $\Psi(x,t)$ is the phase of the wave which is related to the frequency ω and wavenumber k by

$$\frac{\partial \Psi}{\partial t} = -\omega \tag{2.14}$$

$$\frac{\partial \Psi}{\partial x} = k \tag{2.15}$$

The frequency ω is constant but the wavenumber k will vary slowly with x because of the slowly varying coefficients $C_{mn}(x)$ so the above relations can be integrated to give ,

$$\Psi(x,t) = \int_0^x k(\xi)\ d\xi \ - \omega t \tag{2.16}$$

To explain the asymptotic approximations which are made two characteristic length scales L_k and L_A and one characteristic time scale T_A are defined. L_k is the length scale for variations in k, L_A is the length scale for variations in the amplitude A, and T_A is the time scale for variations in A. Numerical values for L_k, L_A and T_A are are given by ,

$$L_k = \min \left(k \Big/ \frac{\partial k}{\partial x} \right) \tag{2.17a}$$

$$L_A = \min \left(A \Big/ \frac{\partial A}{\partial x} \right) \tag{2.17b}$$

$$T_A = \min \left(A \Big/ \frac{\partial A}{\partial t} \right) \tag{2.17c}$$

The asymptotic approximations used in this theory are

$$L_k >> k^{-1} \qquad L_A >> k^{-1} \qquad T_A >> \omega^{-1}$$

which imply

$$\frac{\partial k}{\partial x} << k^2 \qquad \frac{\partial A}{\partial x} << Ak \qquad \frac{\partial A}{\partial t} << A\omega$$

A Taylor series expansion of A and Ψ about a point (x_0, t_0) gives ,

$$A(x,t) = A_0 + \frac{\partial A}{\partial x_0}(x-x_0) + \frac{\partial A}{\partial t_0}(t-t_0) + H.O.T \tag{2.18}$$

$$\Psi(x,t) = \Psi_0 - \omega(t-t_0) + \int_{x_0}^{x} k(\xi)\, d\xi$$

$$= \Psi_0 - \omega(t-t_0) + \int_{x_0}^{x} \left[k_0 + \frac{\partial k}{\partial x_0}(\xi - x_0) + H.O.T \right] d\xi$$

$$= \Psi_0 - \omega(t-t_0) + k_0(x-x_0) + \frac{1}{2}\frac{\partial k}{\partial x_0}(x-x_0)^2 + \text{H.O.T.} \qquad (2.19)$$

Subscript $_0$ denotes terms evaluated at (x_0,t_0).

The H.O.T. , higher order terms, includes terms like $\frac{\partial^2 A}{\partial x^2}$, $\frac{\partial^2 A}{\partial t^2}$, $\frac{\partial^2 k}{\partial x^2} A$ which are $O\{A(L_A k)^{-2}, A(T_A \omega)^{-2}, A(L_k k)^{-2}\}$ and are neglected in this asymptotic approximation.

$$\exp[i\Psi(x,t)] =$$

$$= \exp\left(i\Psi_0 + ik_0(x-x_0) - i\omega(t-t_0) + \frac{i}{2}\frac{\partial k}{\partial x_0}(x-x_0)^2 + \text{H.O.T} \right)$$

$$= \exp[i\Psi_0 + ik_0(x-x_0) - i\omega(t-t_0)] \left(1 + \frac{i}{2}\frac{\partial k}{\partial x}_0 (x-x_0)^2 \right) + \text{H.O.T.} \qquad (2.20)$$

Hence ,

$$u(x,t) = \exp[i\Psi_0 + ik_0(x-x_0) - i\omega(t-t_0)] \left(1 + \frac{i}{2}\frac{\partial k}{\partial x_0}(x-x_0)^2 \right) \left(A_0 + \frac{\partial A}{\partial x_0}(x-x_0) + \frac{\partial A}{\partial t_0}(t-t_0) \right) + \text{H.O.T.}$$

$$= \exp[i\Psi_0 + ik_0(x-x_0) - i\omega(t-t_0)] \cdot \left(A_0 + \frac{\partial A}{\partial x_0}(x-x_0) + \frac{\partial A}{\partial t_0}(t-t_0) + \frac{i}{2} A_0 \frac{\partial k}{\partial x_0}(x-x_0)^2 \right) + \text{H.O.T.} \qquad (2.21)$$

To evaluate derivatives of $u(x,t)$ at (x_0,t_0) a two-variable version of Leibnitz's rule is used.

$$\left(\frac{\partial}{\partial x}\right)^m \left(\frac{\partial}{\partial t}\right)^n [\, f(x,t)\, g(x,t)\,] =$$

$$\sum_{p=0}^{m} \sum_{q=0}^{n} \frac{m!}{p!(m-p)!} \frac{n!}{q!(n-q)!} \left(\left(\frac{\partial}{\partial x}\right)^p \left(\frac{\partial}{\partial t}\right)^q f \right) \left(\left(\frac{\partial}{\partial x}\right)^{m-p} \left(\frac{\partial}{\partial t}\right)^{n-q} g \right)$$

(2.22)

Let

$$f(x,t) = \exp[i\Psi_0 + ik_0(x-x_0) - i\omega(t-t_0)] \tag{2.23}$$

$$g(x,t) = A_0 + \frac{\partial A}{\partial x_0}(x-x_0) + \frac{\partial A}{\partial t_0}(t-t_0) + \frac{i}{2} A_0 \frac{\partial k}{\partial x_0}(x-x_0)^2 \tag{2.24}$$

Then,

$$\left(\left(\frac{\partial}{\partial x}\right)^p \left(\frac{\partial}{\partial t}\right)^q f\right)_0 = (ik_0)^p (-i\omega)^q \exp[i\Psi_0] \tag{2.25}$$

$$g_0 = A_0 \tag{2.26}$$

$$\frac{\partial g}{\partial x_0} = \frac{\partial A}{\partial x_0} \tag{2.27}$$

$$\frac{\partial g}{\partial t_0} = \frac{\partial A}{\partial t_0} \tag{2.28}$$

$$\frac{\partial^2 g}{\partial x^2_0} = iA_0 \frac{\partial k}{\partial x_0} \tag{2.29}$$

and all other derivatives of $g(x,t)$ evaluated at (x_0,t_0) are zero. Hence,

$$\left(\frac{\partial}{\partial x}\right)^m \left(\frac{\partial}{\partial t}\right)^n u(x,t)_0 = \exp[i\Psi_0] \cdot$$

$$\left(A_0(ik_0)^m(-i\omega)^n + m\frac{\partial A}{\partial x_0}(ik_0)^{m-1}(-i\omega)^n + n\frac{\partial A}{\partial t_0}(ik_0)^m(-i\omega)^{n-1} \right.$$

$$\left. + i\frac{m(m-1)}{2} A_0 \frac{\partial k}{\partial x_0}(ik_0)^{m-2}(-i\omega)^n \right) + \text{H.O.T.} \tag{2.30}$$

and so,

$$L(x)\ u(x,t) = \exp[i\Psi] \sum_{m,n} C_{mn} \Big(A(ik)^m(-i\omega)^n + m \frac{\partial A}{\partial x} (ik)^{m-1}(-i\omega)^n + n \frac{\partial A}{\partial t} (ik)^m(-i\omega)^{n-1} + i\frac{m(m-1)}{2} A \frac{\partial k}{\partial x} (ik)^{m-2}(-i\omega)^n \Big) + H.O.T. \quad (2.31)$$

To satisfy the homogeneous equation (2.12) the amplitude A(x,t) must satisfy ,

$$a_0(k,\omega,x)\ A + a_1(k,\omega,x) \frac{\partial A}{\partial t} + a_2(k,\omega,x) \frac{\partial A}{\partial x} + a_3(k,\omega,x) A \frac{\partial k}{\partial x} = 0 + H.O.T. \quad (2.32)$$

where ,

$$a_0(k,\omega,x) = \sum_{m,n} C_{mn}(x) (ik)^m(-i\omega)^n \quad (2.33a)$$

$$a_1(k,\omega,x) = \sum_{m,n} C_{mn}(x)\ n\ (ik)^m(-i\omega)^{n-1} = i \frac{\partial a_0}{\partial \omega} \quad (2.33b)$$

$$a_2(k,\omega,x) = \sum_{m,n} C_{mn}(x)\ m\ (ik)^{m-1}(-i\omega)^n = -i \frac{\partial a_0}{\partial k} \quad (2.33c)$$

and $$a_3(k,\omega,x) = \sum_{m,n} C_{mn}(x)\ i \frac{m(m-1)}{2} (ik)^{m-2}(-i\omega)^n = -\frac{i}{2} \frac{\partial^2 a_0}{\partial k^2} \quad (2.33d)$$

Because of the asymptotic assumptions ,

$$A >> \omega^{-1} \frac{\partial A}{\partial t} ,\ k^{-1} \frac{\partial A}{\partial x} ,\ k^{-2} A \frac{\partial k}{\partial x}$$

so (2.32) can only be satisfied if

$$a_0(k,\omega,x) = 0 \qquad (2.34)$$

This is the dispersion relation between k and ω. k is now a slowly varying function of x due to the slow variation in the coefficients. Thus the characteristic length scale L_k is related to some characteristic length scale L_c for variations in the coefficients.

Neglecting the H.O.T. and dividing by a_1 gives the asymptotic amplitude equation.

$$\frac{\partial A}{\partial t} + c_g \frac{\partial A}{\partial x} = \varepsilon A \qquad (2.35)$$

where $c_g = a_2 / a_1$ (2.36)

and $\varepsilon = - a_3 \frac{\partial k}{\partial x} / a_1$ (2.37)

If c_g is real the left hand side of (2.35) is a Lagrangian-type total time derivative with respect to an observer moving with velocity c_g. If the coefficients C_{mn} are all constant, k is constant, $\varepsilon = 0$ and so the amplitude A is constant along rays moving with velocity c_g. In a wavepacket individual wavecrests move with phase velocity ω/k , usually denoted c_p , but the wavepacket, or amplitude 'envelope' , moves with velocity c_g. For this reason c_g is called the group velocity.

Because the group velocity is the propagation velocity for the amplitude and energy of the wavepacket the group velocity is often more important than the phase velocity. One example is the Sommerfeld radiation condition which states that the waves generated by a fixed source have a group velocity directed away from the source. In some

unusual cases the phase velocity of the waves is actually directed towards the source. A second example is that the group velocity never exceeds the speed of light which is the limiting speed of propagation of information while the phase velocity can exceed the speed of light , as happens in wave propagation along an electromagnetic waveguide.

To link this derivation of group velocity to other derivations the dispersion relation (2.34) is differentiated with x held constant.

$$da_0 = \frac{\partial a_0}{\partial k}\, dk + \frac{\partial a_0}{\partial \omega}\, d\omega$$

$$= 0 \qquad (2.38)$$

$$\text{Hence } \left(\frac{\partial \omega}{\partial k}\right)_{x \text{ const}} = - \frac{\partial a_0}{\partial k} \Big/ \frac{\partial a_0}{\partial \omega}$$

$$= a_2 / a_1$$

$$= c_g \qquad (2.39)$$

The most common method of showing that $\left(\frac{\partial \omega}{\partial k}\right)_{x \text{ const}}$ is the group velocity uses the method of stationary phase which is well explained in the available literature [1,2]. The usual approach is to combine the dispersion relation , the definition of the group velocity and some physical principle such as energy conservation to calculate the propagation of energy. The approach given above is not usually used partly because sometimes the exact partial differential equation is not known and the dispersion relation has been determined by asymptotic methods (e.g. water waves) or from empirical data (e.g. seismic waves). This approach is however suited to analysing finte difference schemes in which the exact finite difference equations are known and there is no general equivalent to

the principle of energy conservation.

3. Amplitude Analysis of Finite Difference Equations

3.1 Fourier Analysis

Consider a homogeneous finite difference equation

$$L\, U_j^n = 0 \tag{3.1}$$

As explained in the appendix A.1, L can always be expressed as a sum of step operators,

$$L \equiv \sum_{m,p} C_{mp}\, E_{mx}\, E_{pt} \tag{3.2}$$

where the coefficients C_{mp} are constants, but it is often more simply expressed as a polynomial of finite difference operators written symbolically as,

$$L \equiv P(E_x, \delta_x, \mu_x, E_t, \delta_t, \mu_t) \tag{3.3}$$

The eigenfunctions and eigenmodes of L are defined exactly as in §2.1. The finite operators all have the same eigenfunctions, $\exp[i(j\phi - n\Omega)]$. ϕ and Ω are related to the wavenumber k and frequency ω of the physical wave being modelled by,

$$\phi = k\Delta x \tag{3.4}$$

$$\Omega = \omega\Delta t \tag{3.5}$$

As shown in the appendix A.1

$$E_x \exp[i(j\phi - n\Omega)] = \exp(i\phi) \exp[i(j\phi - n\Omega)] \tag{3.6a}$$

$$\delta_x \exp[i(j\phi - n\Omega)] = 2i \sin(\phi/2) \exp[i(j\phi - n\Omega)] \tag{3.6b}$$

$$\mu_x \exp[i(j\phi - n\Omega)] = \cos(\phi/2) \exp[i(j\phi - n\Omega)] \tag{3.6c}$$

$$E_t \exp[i(j\phi - n\Omega)] = \exp(-i\Omega) \exp[i(j\phi - n\Omega)] \tag{3.6d}$$

$$\delta_t \exp[i(j\phi-n\Omega)] = -2i \sin(\Omega/2) \exp[i(j\phi-n\Omega)] \qquad (3.6e)$$

$$\mu_t \exp[i(j\phi-n\Omega)] = \cos(\Omega/2) \exp[i(j\phi-n\Omega)] \qquad (3.6f)$$

$$E_{mx} E_{pt} \exp[i(j\phi-n\Omega)] = \exp[i(m\phi-p\Omega)] \exp[i(j\phi-n\Omega)] \qquad (3.6g)$$

Thus $\exp[i(j\phi-n\Omega)]$ is an eigenfunction of E_x, δ_x, μ_x, E_t, δ_t, μ_t and $E_{mx} E_{pt}$ with eigenvalues $\exp(i\phi)$, $2i \sin(\phi/2)$, $\cos(\phi/2)$, $\exp(-i\Omega)$, $-2i \sin(\Omega/2)$, $\cos(\Omega/2)$, $\exp[i(m\phi-p\Omega)]$ respectively, and L has eigenvalue $\sum_{m,p} C_{mp} \exp[i(m\phi-p\Omega)]$ or

$$P[\exp(i\phi), 2i \sin(\phi/2), \cos(\phi/2), \exp(-i\Omega), -2i \sin(\Omega/2), \cos(\Omega/2)]$$

depending which expression for L is used.

$$U_j^n = \exp[i(j\phi-n\Omega)] \qquad (3.7)$$

is an exact solution of (3.1) provided

$$\sum_{m,p} C_{mp} \exp[i(m\phi-p\Omega)] = 0 \qquad (3.8)$$

This is the dispersion relation between ϕ and Ω.

Since $\exp[2n\pi i] = 1$ for all integers n, $\phi + 2n\pi$ is equivalent to ϕ so only solutions in the ranges

$$-\pi < \mathrm{Re}(\phi) < \pi$$

$$-\pi < \mathrm{Re}(\Omega) < \pi$$

need be considered.

If L involves P+1 time levels and M+1 spatial nodes the dispersion relation is a polynomial of degree P in $\exp(-i\Omega)$ and of degree M in $\exp(i\phi)$. Thus for a given ϕ there are P corresponding values of Ω, and for a given Ω

there are M corresponding values of ϕ.

A general solution of (3.1) for periodic boudary conditions is a superposition of eigenmodes.

$$U_j^n = \sum_{\phi} \sum_{p=1}^{P} A_p(\phi) \exp[i(j\phi - n\Omega_p)] \qquad (3.9)$$

The ϕ summation is a summation over all the values of ϕ which satisfy the periodic boundary conditions, and the p summation is over the P values of Ω corresponding to each value of ϕ.

If $A_p(\phi)$ is non-zero for all ϕ,p then a necessary and sufficient condition for U_j^n to remain bounded and not increase exponentially is that each eigenmode must remain bounded. Splitting Ω into its real and imaginary components gives,

$$\Omega = \Omega_R + i\Omega_I \qquad (3.10)$$

$$\exp[-in\Omega] = \exp[-in\Omega_R + n\Omega_I] \qquad (3.11)$$

Thus the condition that every eigenmode remain bounded , and hence a general solution remain bounded , is $\Omega_I < 0$ for all ϕ,p.

This analysis is lacking in the same three respects as the analysis of partial differential equations by eigenmode expansion in the last chapter. The analysis gives no information about the movement of an initially localised disturbance, cannot incorporate boundary conditions or analyse schemes with non-constant coefficients.

3.2 Asymptotic Amplitude Equation

In this section the coefficients C_{mp} in the definition of L (3.2) are assumed to be slowly varying functions of j. The analysis is performed in computational space with coordinates (j,n) in whicn the grid spacing is $\Delta j=1$, $\Delta n=1$. Variations in mesh spacing in physical coordinates are incorporated directly into the variable coefficients of the finite difference equations.

The theory calculates the approximate evolution of a wavetrain with waves of a constant frequency Ω and a slowly varying amplitude , so U_j^n is written as ,

$$U_j^n = A(j,n) \exp[i\Psi(j,n)] \tag{3.12}$$

where A(j,n) is the slowly varying amplitude and $\Psi(j,n)$ is the phase of the wave and is related to the frequency Ω and wavenumber ϕ by

$$\frac{\partial \Psi}{\partial n} = -\Omega \tag{3.13}$$

$$\frac{\partial \Psi}{\partial j} = \phi \tag{3.14}$$

which can be integrated to give,

$$\Psi(j,n) = \int_0^j \phi(\xi)\, d\xi \; - n\Omega \tag{3.15}$$

As in §2.2 two characteristic length scales, L_ϕ for variations in ϕ, and L_A for variations in the amplitude A and one characteristic time scale T_A for variations in the amplitude A , can be defined with numerical values being given by,

$$L_\phi = \min \left(1 \Big/ \frac{\partial \phi}{\partial j} \right) \tag{3.16a}$$

$$L_A = \min \left(A \Big/ \frac{\partial A}{\partial j} \right) \tag{3.16b}$$

$$T_A = \min \left(A \Big/ \frac{\partial A}{\partial n} \right) \tag{3.16c}$$

The asymptotic approximations are

$$L_\phi >> 1 \qquad L_A >> 1 \qquad T_A >> 1$$

which imply

$$\frac{\partial \phi}{\partial j} << 1 \qquad \frac{\partial A}{\partial j} << A \qquad \frac{\partial A}{\partial n} << A$$

A Taylor series expansion of A and Ψ about a point (j_0, n_0) gives,

$$A(j_0+m, n_0+p) = A_0 + m \frac{\partial A}{\partial j_0} + p \frac{\partial A}{\partial n_0} + \text{H.O.T} \tag{3.17}$$

$$\Psi(j_0+m, n_0+p) = \Psi_0 - p\Omega + \int_{j_0}^{j_0+m} \phi(\xi)\, d\xi$$

$$= \Psi_0 - p\Omega + \int_{j_0}^{j_0+m} \left[\phi_0 + \frac{\partial \phi}{\partial j_0} (\xi - j_0) + \text{H.O.T} \right] d\xi$$

$$= \Psi_0 - p\Omega + m\phi_0 + \frac{m^2}{2} \frac{\partial \phi}{\partial j_0} + \text{H.O.T.} \tag{3.18}$$

Subscript $_0$ denotes terms evaluated at (j_0, n_0).

The H.O.T. , higher order terms, includes terms like $\frac{\partial^2 A}{\partial j^2}$, $\frac{\partial^2 A}{\partial n^2}$, $\frac{\partial^2 \phi}{\partial j^2} A$ which are $O\{AL_A^{-2}, AT_A^{-2}, AL_\phi^{-2}\}$ and are neglected in this asymptotic approximation.

$$\exp[i\Psi(j_0+m, n_0+p)]$$

$$= \exp\left[i\Psi_0 - ip\Omega + im\phi_0 + \frac{im^2}{2} \frac{\partial \phi}{\partial j_0} + H.O.T \right]$$

$$= \exp[i\Psi_0 - ip\phi + im\phi_0] \left(1 + \frac{im^2}{2} \frac{\partial \phi}{\partial j_0} \right) + H.O.T. \quad (3.19)$$

Hence ,

$$U^{n_0+p}_{j_0+m} = \exp[i\Psi_0] \exp[i(m\phi_0 - p\Omega)] \left(1 + \frac{im^2}{2} \frac{\partial \phi}{\partial j_0} \right) \left(A_0 + m \frac{\partial A}{\partial j_0} + p \frac{\partial A}{\partial n_0} \right) + H.O.T.$$

$$= \exp[i\Psi_0] \exp[i(m\phi_0 - p\Omega)] \left(A_0 + m \frac{\partial A}{\partial j_0} + p \frac{\partial A}{\partial n_0} + \frac{im^2}{2} A_0 \frac{\partial \phi}{\partial j_0} \right) + H.O.T. \quad (3.20)$$

Hence,

$$L_j U^n_j = \exp[i\Psi] \sum_{m,p} C_{mp}(j) \exp[i(m\phi - p\Omega)] \cdot \left(A + m \frac{\partial A}{\partial j} + p \frac{\partial A}{\partial n} + \frac{im^2}{2} A \frac{\partial \phi}{\partial j} \right) + H.O.T. \quad (3.21)$$

To satisfy the homogeneous equation (3.1) the amplitude A(j,n) must satisfy,

$$a_0(\phi,\Omega,j)\, A + a_1(\phi,\Omega,j) \frac{\partial A}{\partial n} + a_2(\phi,\Omega,j) \frac{\partial A}{\partial j} + a_3(\phi,\Omega,j)\, A \frac{\partial \phi}{\partial j} = 0 + H.O.T. \quad (3.22)$$

where ,

$$a_0(\phi,\Omega,j) = \sum_{m,p} C_{mp}(j) \exp[i(m\phi - p\Omega)] \quad (3.23a)$$

$$a_1(\phi,\Omega,j) = \sum_{m,p} C_{mp}(j)\, p \exp[i(m\phi - p\Omega)] = i\left(\frac{\partial a_0}{\partial \Omega}\right)_{\phi, j \text{ const}} \quad (3.23b)$$

$$a_2(\phi,\Omega,j) = \sum_{m,p} C_{mp}(j)\ m \exp[i(m\phi - p\Omega)]$$

$$= -i\left(\frac{\partial a_0}{\partial \phi}\right)_{\Omega,j\ \text{const}} \qquad (3.23c)$$

and $$a_3(\phi,\Omega,j) = \sum_{m,p} C_{mp}(j)\ \frac{im^2}{2}\ \exp[i(m\phi - p\Omega)]$$

$$= -\frac{i}{2}\left(\frac{\partial^2 a_0}{\partial \phi^2}\right)_{\Omega,j\ \text{const}} \qquad (3.23d)$$

In the above derivation of a_0, a_1, a_2, a_3 the general shift operator expression for L_j (3.2) is used. In applications it is more convenient to use the finite operator polynomial expression (3.3). a_0 is obtained by replacing each operator with its corresponding eigenvalue and then a_1, a_2, a_3 are calculated by differentiating a_0.

Because of the asymptotic assumptions ,

$$A >> \frac{\partial A}{\partial n}\ ,\ \frac{\partial A}{\partial j}\ ,\ A\ \frac{\partial \phi}{\partial j}$$

so (3.22) can only be satisfied if

$$a_0(\phi,\Omega,j) = 0 + O\{ L_\phi^{-1}, L_A^{-1}, T_A^{-1} \} \qquad (3.24)$$

This is the asymptotic form of the dispersion relation between ϕ and Ω and will usually be satisfied by setting a_0 identically equal to zero. ϕ is now a slowly varying function of j due to the slow variation in the coefficients. The characteristic length scale L_ϕ is related to some characteristic length scale L_C for variations in the coefficients.

Neglecting the H.O.T. and dividing by a_1 gives the asymptotic amplitude equation.

$$\frac{\partial A}{\partial n} + r_g \frac{\partial A}{\partial j} = \varepsilon A \tag{3.25}$$

where $$r_g = a_2 / a_1 \tag{3.26}$$

and $$\varepsilon = - \left(a_3 \frac{\partial \phi}{\partial j} + a_0 \right) / a_1 \tag{3.27}$$

Differentiating (3.24) with Ω held constant gives,

$$da_0 = \left(\frac{\partial a_0}{\partial \phi}\right)_{\Omega, j \text{ const}} d\phi + \left(\frac{\partial a_0}{\partial j}\right)_{\phi, \Omega \text{ const}} dj$$

$$= 0 + \text{H.O.T.} \tag{3.28}$$

Hence, neglecting the higher order terms,

$$\frac{\partial \phi}{\partial j} = - \left(\frac{\partial a_0}{\partial j}\right)_{\phi, \Omega \text{ const}} \Big/ \left(\frac{\partial a_0}{\partial \phi}\right)_{\Omega, j \text{ const}}$$

$$= i \left(\frac{\partial a_0}{\partial j}\right)_{\phi, \Omega \text{ const}} / \; a_2 \tag{3.29}$$

so $$\varepsilon = -i \, a_3 \left(\frac{\partial a_0}{\partial j}\right)_{\phi, \Omega \text{ const}} / \; a_1 a_2 \tag{3.30}$$

If r_g is real the left hand side of (3.25) is a Lagrangian-type total time derivative with respect to an observer moving with velocity r_g. Thus the amplitude A is being convected with velocity r_g in computational space.

Differentiating (3.24) with j held constant gives,

$$da_0 = \left(\frac{\partial a_0}{\partial \Omega}\right)_{\phi, j \text{ const}} d\Omega + \left(\frac{\partial a_0}{\partial \phi}\right)_{\Omega, j \text{ const}} d\phi$$

$$= 0 + \text{H.O.T.} \tag{3.31}$$

Hence, neglecting the higher order terms,

$$\left(\frac{\partial \Omega}{\partial \phi}\right)_{j \text{ const}} = - \left(\frac{\partial a_0}{\partial \phi}\right)_{\Omega, j \text{ const}} \Big/ \left(\frac{\partial a_0}{\partial \Omega}\right)_{\phi, j \text{ const}}$$

$$= a_2 / a_1$$

$$= r_g \tag{3.32}$$

Substituting for Ω and ϕ using (3.5a,b),

$$r_g = \frac{\partial(\omega\Delta t)}{\partial(k\Delta x)}$$

$$= \frac{\Delta t}{\Delta x}\frac{\partial\omega}{\partial k}$$

$$= \frac{c_g \Delta t}{\Delta x} \qquad (3.33)$$

Thus r_g is the CFL number corresponding to the group velocity in physical space of the propagating numerical wave. It is the number of spatial mesh intervals which a localised disturbance travels in one time step. For the rest of the paper r_g is called the group CFL number.

3.3 Examples

The model problem which is considered is,

$$\frac{\partial u}{\partial t} + c(x) \frac{\partial u}{\partial x} = 0 \qquad -\infty < x < \infty \tag{3.34}$$

Three different methods are analysed.

3.3.1 Trapezoidal Scheme

The trapezoidal scheme is

$$\frac{1}{\Delta t} (U_j^{n+1} - U_j^n) + \frac{c_j}{4\Delta x_j} \left((U_{j+1}^{n+1} + U_{j+1}^n) - (U_{j-1}^{n+1} + U_{j-1}^n) \right) = 0 \tag{3.35}$$

which can be written using operator notation as

$$\frac{1}{\Delta t} \left(\delta_t + \frac{r_j}{2} \delta_{2x} \mu_t \right) U_j^{n+\frac{1}{2}} = 0 \tag{3.36}$$

where the CFL number r is defined as

$$r_j = \frac{c\Delta t}{\Delta x}_j \tag{3.37}$$

$$\text{and } \Delta x_j = \frac{1}{2} (x_{j+1} - x_{j-1}) \tag{3.38}$$

a_0 is obtained by replacing the operators by their eigenvalues.

$$a_0 = -2i \sin(\Omega/2) + \frac{r}{2} 2i \sin(\phi) \cos(\Omega/2) \tag{3.39}$$

The dispersion relation is

$$a_0 = 0 \tag{3.40}$$

$$\text{so} \quad \tan(\Omega/2) = \frac{r}{2} \sin(\phi) \tag{3.41}$$

a_1, a_2, a_3 are obtained by differentiating a_0

$$a_1 = i \frac{\partial a_0}{\partial \Omega}$$

$$= \cos(\Omega/2) + \frac{r}{2} \sin(\phi) \sin(\Omega/2) \qquad (3.42)$$

$$a_2 = -i \frac{\partial a_0}{\partial \phi}$$

$$= r \cos(\psi) \cos(\Omega/2) \qquad (3.43)$$

$$a_3 = -\frac{1}{2} i \frac{\partial^2 a_0}{\partial \phi^2}$$

$$= \frac{1}{2} \frac{\partial a_2}{\partial \phi}$$

$$= -\frac{r}{2} \sin(\phi) \cos(\Omega/2) \qquad (3.44)$$

Using the dispersion relation a_1 and a_3 can be simplified.

$$a_1 = \cos(\Omega/2) + \frac{r}{2} \sin(\phi) \sin(\Omega/2)$$

$$= \cos(\Omega/2) + \tan(\Omega/2) \sin(\Omega/2)$$

$$= [\cos^2(\Omega/2) + \sin^2(\Omega/2)] / \cos(\Omega/2)$$

$$= 1 / \cos(\Omega/2) \qquad (3.45)$$

$$a_3 = -\frac{r}{2} \sin(\phi) \cos(\Omega/2)$$

$$= -\tan(\Omega/2) \cos(\Omega/2)$$

$$= -\sin(\Omega/2) \qquad (3.46)$$

so $r_g = a_2 / a_1$

$$= r \cos(\phi) \cos^2(\Omega/2) \qquad (3.47)$$

and $\varepsilon = -i\, a_3 \left(\frac{\partial}{\partial j} \frac{a_0}{}\right)_{\phi, \Omega \text{ const}} / a_1 a_3 - a_0 / a_1$

$$= \frac{ir}{2} \sin(\phi) \cos(\Omega/2) \left[\frac{\partial r}{\partial j} \sin(\phi) \cos(\Omega/2) \right] / [r \cos(\phi)]$$

$$= -\frac{1}{2}\frac{\partial r}{\partial j}\sin^2(\phi)\cos^2(\Omega/2)\,/\,\cos(\phi) \qquad (3.48)$$

There are three points of interest

i) $\sin(\pi-\phi) = \sin(\phi)$

so for all Ω there are two corresponding values of ϕ given by the dispersion relation,

ϕ_1 satisfying $-\pi/2 < \mathrm{Re}(\phi_1) < \pi/2$

and $\phi_2 = \pi - \phi_1$

ii) For real Ω in the range

$0 < \Omega < 2\tan^{-1}(r/2)$

$0 < \tan(\Omega/2) < r/2$

so $0 < \sin(\phi) < 1$

Thus ϕ_1 and ϕ_2 are both real and

$0 < \phi_1 < \pi/2$

so $r_g(\phi_1) > 0$

and $\pi/2 < \phi_2 < \pi$

so $r_g(\phi_2) < 0$

Hence for every frequency in the given range there is one forward travelling wave, travelling in the same direction as the physical waves being modelled, and one backward travelling wave with wavelength less than $4\Delta x$.

iii) For real Ω in the range

$2\tan^{-1}(r/2) < \Omega < \pi$

$\sin(\phi) > 1$

so ϕ_1 and ϕ_2 are complex

Let $\phi_1 = \pi/2 + i\phi_I$

Then $\sin(\phi_1) = \cosh(\phi_I)$

so ϕ_I is real and satisfies

$$\tan(\Omega/2) = \frac{r}{2}\cosh(\phi_I)$$

$$\phi_2 = \pi - \phi_1$$

$$= \pi/2 - i\phi_I$$

These are evanescent waves. If there are boundaries at j = 0,J and the boundary conditions force a steady oscillation with a frequency in the given range one wave will decay in amplitude exponentially away from the boundary at j=0, while the other will decay exponentially away from the boundary at j=J.

3.3.2 Box Scheme

The box scheme is

$$\frac{1}{\Delta t}\left((U_j^{n+1} + U_{j+1}^{n+1}) - (U_j^n + U_{j+1}^n)\right) +$$

$$\frac{c_j}{\Delta x_j}\left((U_{j+1}^n + U_{j+1}^{n+1}) - (U_j^n + U_j^{n+1})\right) = 0 \qquad (3.49)$$

which may be written in operator notation as

$$\frac{1}{\Delta t}\left[\mu_x \delta_t + r_j \mu_t \delta_x\right] U_{j+\frac{1}{2}}^{n+\frac{1}{2}} = 0 \qquad (3.50)$$

$$a_0 = -2i \cos(\phi/2) \sin(\Omega/2) + 2ir \cos(\Omega/2) \sin(\phi/2) \qquad (3.51)$$

The dispersion relation is

$$\tan(\Omega/2) = r \tan(\phi/2) \qquad (3.52)$$

$$a_1 = i \frac{\partial a}{\partial \Omega}_0$$

$$= \cos(\phi/2) \cos(\Omega/2) + r \sin(\Omega/2) \sin(\phi/2) \qquad (3.53)$$

$$a_2 = -i \frac{\partial a_0}{\partial \phi}$$

$$= \sin(\phi/2) \sin(\Omega/2) + r \cos(\Omega/2) \cos(\phi/2) \qquad (3.54)$$

$$a_3 = \frac{i}{2} \frac{\partial^2 a_0}{\partial \phi^2}$$

$$= \frac{1}{4} \cos(\phi/2) \sin(\Omega/2) - \frac{1}{4} r \cos(\Omega/2) \sin(\phi/2)$$

$$= -\frac{i}{8} a_0$$

$$= 0 \qquad (3.55)$$

Thus $r_g = a_2 / a_1$

$$= \frac{\sin(\phi/2) \sin(\Omega/2) + r \cos(\Omega/2) \cos(\phi/2)}{\cos(\phi/2) \cos(\Omega/2) + r \sin(\Omega/2) \sin(\phi/2)}$$

$$= \frac{\tan(\phi/2)\ \tan(\Omega/2) + r}{1 + r\ \tan(\phi/2)\ \tan(\Omega/2)}$$

$$= r\ \frac{1 + \tan^2(\phi/2)}{1 + r^2\ \tan^2(\phi/2)} \qquad (3.56)$$

and $\varepsilon = 0$ (3.57)

There are two points of interest

i) For each real value of Ω there is one corresponding real value of ϕ given by the dispersion relation and the group CFL number r_g is real and positive.

ii) When $r=1$, $r_g=r$, so waves of all frequencies travel at the same velocity as the physical waves being modelled. This is because when $r=1$ the Box scheme reduces to,

$$U_{j+1}^{n+1} = U_j^n \qquad (3.58)$$

which agrees exactly with the solution of the partial differential equation,

$$u(x+ct,t) = u(x,0) \qquad (3.59)$$

3.3.3 Backward Euler Scheme

The backward Euler scheme is,

$$\frac{1}{\Delta t}\left(U_j^{n+1} - U_j^n\right) + \frac{c_j}{2\Delta x}\left(U_{j+1}^{n+1} - U_{j-1}^{n+1}\right) = 0 \qquad (3.60)$$

which may be written using operator notation as

$$\frac{1}{\Delta t}\left(\nabla_t + \frac{1}{2} r_j \delta_{2x}\right) U_j^{n+1} = 0 \qquad (3.61)$$

$$a_0 = 1 - \exp(i\Omega) + ir \sin(\phi) \qquad (3.62)$$

The dispersion relation is

$$\exp[i\Omega] - 1 = ir \sin(\phi) \qquad (3.63)$$

$$a_1 = i \frac{\partial a_0}{\partial \Omega}$$

$$= \exp(i\Omega) \qquad (3.64)$$

$$a_2 = -i \frac{\partial a_0}{\partial \phi}$$

$$= r \cos(\phi) \qquad (3.65)$$

$$a_3 = -\frac{1}{2} i \frac{\partial^2 a_0}{\partial \phi^2}$$

$$= \frac{1}{2} \frac{\partial a_2}{\partial \phi}$$

$$= -\frac{r}{2} \sin(\phi) \qquad (3.66)$$

so $r_g = a_2 / a_1$

$$= r \cos(\phi) \exp(-i\Omega) \qquad (3.67)$$

and $\varepsilon = -i\, a_3 \left(\frac{\partial a_0}{\partial j}\right)_{\phi,\Omega \text{ const}} / a_1 a_3 - a_0 / a_1$

$$= \frac{ir}{2} \sin(\phi) \left[i \frac{\partial r}{\partial j} \sin(\phi) \right] / \left[\exp(i\Omega)\, r \cos(\phi) \right]$$

$$= -\frac{1}{2} \frac{\partial r}{\partial j} \sin^2(\phi) \exp(-i\Omega) / \cos(\phi) \qquad (3.68)$$

3.4 Asymptotic Boundary Conditions

The general solution of

$$L_j U_j^n = 0 \tag{3.1}$$

is a sum of waves with different constant frequencies Ω and slowly varying wavenumber ϕ and amplitude A

$$U_j^n = \sum_{\Omega} \sum_{m=1}^{M} A_m(j,n) \exp[i(\int_0^j \phi_m(\xi)\, d\xi - n\Omega)] \tag{3.69}$$

The outer summation is over different values of Ω, and the inner summation is over the M different values of ϕ which satisfy the dispersion relation for each Ω.

For each Ω,m the amplitude A satisfies its asymptotic amplitude equation on the interior of the computational domain independent of all the other waves. All the waves of each frequency are however coupled by boundary conditions.

Suppose a finite difference boundary condition at j=J is

$$B U_J^n = F^n \tag{3.70}$$

where B is a constant finite difference operator which can be expressed in operator polynomial form as

$$B \equiv P_B(E_x, \delta_x, \mu_x, E_t, \delta_t, \mu_t) \tag{3.71}$$

and F^n is a forcing function which can be expressed as a sum of inputs of different frequencies.

$$F^n = \sum_{\Omega} f(\Omega) \exp(-i\Omega) \tag{3.72}$$

Performing exactly the same asymptotic expansion as in the derivation of the asymptotic amplitude equation the boundary condition becomes,

$$\sum_{\Omega} \sum_{m=1}^{M} \left(b\, A_m + i\, \frac{\partial b}{\partial \Omega}\, \frac{\partial A}{\partial n}m - i\, \frac{\partial b}{\partial \phi}\, \frac{\partial A}{\partial j}m - \frac{i}{2}\, \frac{\partial^2 b}{\partial \phi^2}\, A_m\, \frac{\partial \phi}{\partial j}m \right) \cdot$$

$$\exp[i(\int_0^J \phi_m\, d\xi - n\Omega)] = \sum_{\Omega} f(\Omega)\, \exp(-in\Omega) + H.O.T. \qquad (3.73)$$

where $b(\Omega,\phi_m) = P_B[\, \exp(i\phi_m), 2i\, \sin(\phi_m/2), \cos(\phi_m/2),$

$$\exp(-i\Omega), -2i\, \sin(\Omega/2), \cos(\Omega/2)\,] \qquad (3.74)$$

The coefficients of $\exp(-in\Omega)$ in (3.73) must be asymptotically equal to zero for each Ω so,

$$\sum_{m=1}^{M} \left(b\, A_m + i\, \frac{\partial b}{\partial \Omega}\, \frac{\partial A}{\partial n}m - i\, \frac{\partial b}{\partial \phi}\, \frac{\partial A}{\partial j}m - \frac{i}{2}\, \frac{\partial^2 b}{\partial \phi^2}\, A_m\, \frac{\partial \phi}{\partial j}m \right) \cdot$$

$$\exp[i \int_0^J \phi_m\, d\xi\,] = f(\Omega) + H.O.T. \qquad (3.75)$$

This paper is primarily concerned with stability and convergence rates. When analysing perturbations from a steady state or constant amplitude oscillation the boundary condition for the perturbation has

$$f(\Omega) = 0 \qquad (3.76)$$

Because the zero order terms will usually dominate the normal form of the asymptotic boundary conditions is,

$$\sum_{m=1}^{M} b(\Omega,\phi_m)\, A_m(J)\, \exp[i \int_0^J \phi_m\, d\xi\,] = 0 \qquad (3.77)$$

The first order terms $\frac{\partial A_m}{\partial n}$, $\frac{\partial A_m}{\partial j}$, A $\frac{\partial \Phi_m}{\partial j}$ are only important when,

$$b(\Omega,\phi_m) = O\ (\ T_A^{-1},\ L_A^{-1},\ L_\phi^{-1}\ )$$

As explained in §3.1 if the finite difference operator L in the interior scheme spans M+1 spatial levels there will be M values of ϕ given by the dispersion relation for a given value of Ω. If the computational domain is $0<j<J$ the interior scheme gives finite difference equations at J-M+1 nodes, so to complete the set of finite difference equations there must be M finite difference boundary conditions. Hence for each Ω the asymptotic amplitude analysis gives M independent amplitude differential equations on the interior coupled at the boundaries by M boundary conditions.

3.5 Examples

The same model problem as in §3.3 is considered,

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0 \qquad (3.34)$$

$$0 < x < x_J$$

$$c(x) > 0$$

The analytic boundary condition is,

$$u(0,t) = F(t) \qquad (3.78)$$

For perturbation analysis

$$u(0,t) = 0 \qquad (3.79)$$

The finite difference scheme using the trapezoidal or backward Euler methods on the interior requires two finite difference boundary conditions. For perturbations the boundary condition at j=0 is,

$$U_0^n = 0 \qquad (3.80)$$

The boundary condition at j=J is some form of extrapolation. Four of the most commonly used are analysed.

3.5.1 Upstream Boundary

$$U_o^n = 0 \qquad (3.80)$$

$$B = 1 \qquad (3.81)$$

so $$b = 1 \qquad (3.82)$$

Hence $$A_1(0,n) + A_2(0,n) = 0 \qquad (3.83)$$

In preparation for the theory developed in chapters 4 and 5 it is useful to define R_0, the amplitude reflection coefficient as

$$R_0 = \frac{A_1(0,n)}{A_2(0,n)} \tag{3.84}$$

so in this example

$$R_0 = -1 \tag{3.85}$$

3.5.2 Downstream Boundary : Space Extrapolation

The space extrapolation boundary condition is

$$U_J^n = U_{J-1}^n \tag{3.86}$$

$$B = 1 - E_{-x}$$

$$= E_{-\frac{1}{2}x}\ \delta_x \tag{3.87}$$

so $b = 1 - \exp(-i\phi)$

$$= 2i \exp(-i\phi/2) \sin(\phi/2) \tag{3.88}$$

Hence

$$\sum_{m=1}^{2} \sin(\phi_m/2) \exp\left[i \int_0^J \phi_m\, d\xi - i\, \phi_m(J)/2\right] A_m(J,n) = 0 \tag{3.89}$$

The amplitude reflection coefficient R_J is defined as,

$$R_J = \frac{A_2(J,n)}{A_1(J,n)} \tag{3.90}$$

so in this example

$$R_J = -\frac{\sin(\phi_1/2)}{\sin(\phi_2/2)} \exp\left(i \int_0^J (\phi_1 - \phi_2)\, d\xi - \frac{i}{2}(\phi_1(J) - \phi_2(J)) \right) \tag{3.91}$$

3.5.3 Downstream Boundary : Space-time Extrapolation

The space-time extrapolation boundary condition is

$$U_J^n = U_{J-1}^{n-1} \tag{3.92}$$

$$B = 1 - E_{-x}\, E_{-t} \tag{3.93}$$

$$\begin{aligned} b &= 1 - \exp(-i\phi + i\Omega) \\ &= 2i \exp[i(\Omega-\phi)/2] \sin[(\phi-\Omega)/2] \end{aligned} \tag{3.94}$$

Hence,

$$\sum_{m=1}^{2} \sin\left(\frac{\phi_m - \Omega}{2}\right) \exp\left(i \int_0^J \phi_m \, d\xi - i\phi_m(J)/2 \right) A_m(J,n) = 0 \tag{3.95}$$

$$R_J = - \frac{\sin\left(\frac{\phi_1 - \Omega}{2}\right)}{\sin\left(\frac{\phi_2 - \Omega}{2}\right)} \exp\left(i \int_0^J (\phi_1 - \phi_2) \, d\xi - \frac{i}{2}(\phi_1(J) - \phi_2(J)) \right) \tag{3.96}$$

3.5.4 Downstream Boundary : Box Method

In this example the Box method which was discussed in §3.3.2 as an interior scheme is now considered as a downstream boundary condition.

$$[\, \mu_x \delta_t + r\, \mu_t \delta_x \,]\, U_{J-\frac{1}{2}}^{n-\frac{1}{2}} = 0 \tag{3.97}$$

$$B = E_{-\frac{1}{2}x}\, E_{-\frac{1}{2}t}\, [\, \mu_x \delta_t + r\, \mu_t \delta_x \,] \tag{3.98}$$

$$\begin{aligned} \text{so } b = \exp[i(\Omega-\phi)/2] \{ & -2i \cos(\phi/2) \sin(\Omega/2) + \\ & 2i\, r \cos(\Omega/2) \sin(\phi/2) \} \end{aligned} \tag{3.99}$$

Hence,

$$\sum_{m=1}^{2} \left(r \sin(\phi_m/2) - \tan(\Omega/2) \cos(\phi_m/2) \right) \cdot$$

$$\exp\left[i \int_0^J \phi_m \, d\xi - i \, \phi_m(J)/2 \right] A_m(J,n) = 0 \qquad (3.100)$$

$$R_J = - \frac{r \sin(\phi_1/2) - \tan(\Omega/2) \cos(\phi_1/2)}{r \sin(\phi_2/2) - \tan(\Omega/2) \cos(\phi_2/2)} \cdot$$

$$\exp\left(i \int_0^J (\phi_1 - \phi_2) \, d\xi - \frac{i}{2}(\phi_1(J) - \phi_2(J)) \right) \qquad (3.101)$$

4. Ray Theory and Wavepacket-Particle Duality

4.1 Ray Theory

In addition to the asymptotic approximations made in chapter 3 this chapter assumes that for all real wavenumbers ϕ, the frequency Ω is real for all j and hence the group CFL number r_g is real.

$$r_g = \left(\frac{\partial \Omega}{\partial \phi}\right)_{j \text{ const}} \tag{3.32}$$

A Lagrangian-type total time derivative in computational space is defined by,

$$\frac{d}{dn} \equiv \frac{\partial}{\partial n} + r_g \frac{\partial}{\partial j} \tag{4.1}$$

so
$$\frac{dj}{dn} = \frac{\partial j}{\partial n} + r_g \frac{\partial j}{\partial j}$$

$$= r_g \tag{4.2}$$

From the asymptotic amplitude equation (3.25),

$$\frac{dA}{dn} = \frac{\partial A}{\partial n} + r_g \frac{\partial A}{\partial j}$$

$$= \varepsilon A \tag{4.3}$$

and using (3.26) and (3.29),

$$\frac{d\phi}{dn} = \frac{\partial \phi}{\partial n} + r_g \frac{\partial \phi}{\partial j}$$

$$= r_g \frac{\partial \phi}{\partial j}$$

$$= \frac{a_2}{a_1} \; i\left(\frac{\partial a_0}{\partial j}\right)_{\phi, \Omega \text{ const}} / \; a_2$$

$$= \frac{i}{a_1}\left(\frac{\partial a_0}{\partial j}\right)_{\phi, \Omega \text{ const}} \tag{4.4}$$

A general initial value problem for a wave of frequency Ω and wavenumber $\phi(\Omega,j)$ can be solved by integrating these equations (4.2)-(4.4) with initial conditions

$$j(0) = j_0 \tag{4.5a}$$

$$A(0) = A(j_0,0) \tag{4.5b}$$

$$\phi(0) = \phi(\Omega,j_0) \tag{4.5c}$$

Each value of j generates a ray and all of the rays together cover the entire domain for $n > 0$. Figure 2 shows the motion of some typical rays in computational space.

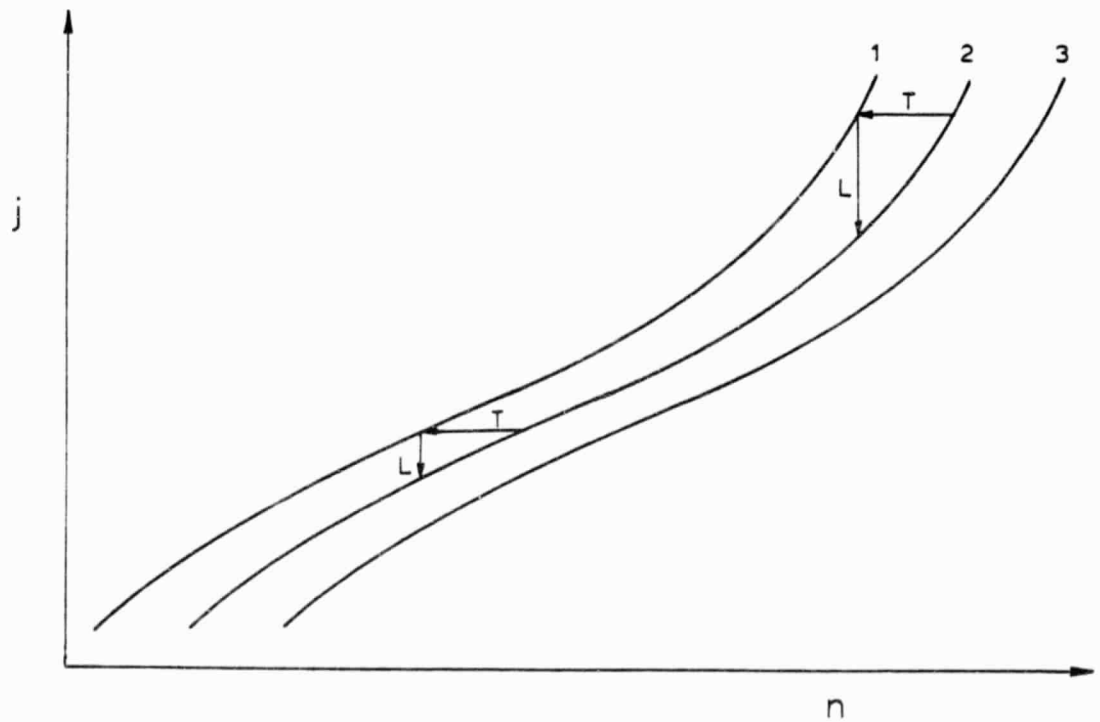


FIGURE 2. RAYS IN COMPUTATIONAL SPACE

r_g is a function of j so at a particular j all the rays have the same slope $\frac{dj}{dn}$. Hence the time separation T of two rays, illustrated in figure 2 for rays 1 and 2, remains constant but the spatial separation L varies as r_g varies.

As explained in chapter 3 if the finite difference operator spans M+1 spatial nodes then for a particular value of Ω there are M values of ϕ which satisfy the dispersion relation. Define M_+ to be the number of solutions ϕ for which the group CFL number is positive, and similarly define M_- to be the number of solutions ϕ for which the group CFL number is negative. Let the computational domain be $0 < j < J$ as usual. At j=J there are M_+ rays leaving the domain and M_- entering it. The amplitudes are related through the asymptotic boundary conditions each of which has the form,

$$\sum_{m=1}^{M} b(\Omega,\phi_m) \exp\left[i \int_0^J \phi_m \, d\xi \right] A_m(J,n) = 0 \qquad (3.37)$$

Since the M_+ amplitudes of the rays leaving the domain are known and the M_- amplitudes of the rays entering the domain are unknown there must be M_- boundary conditions to uniquely determine the amplitudes of the rays entering domain. Similarly at j=0 there must be M_+ boundary conditions to uniquely determine the amplitudes of the rays entering the domain.

4.2 Wavepacket-Particle Theory

In terms of ray theory a wavepacket is a ray tube, a group of rays, along which the amplitude is non-zero. From the discussion in the last section the time length T_A of the wavepacket remains constant but its spatial length L_A will vary whenever r_g has different values at the two ends of the wavepacket. Provided $L_A << L_\phi$ all the rays in the ray tube have approximately the same value for $\phi(\Omega,j)$ so the motion of the wavepacket is given by,

$$\frac{dj}{dn} = r_g(\Omega,\phi(j),j) \tag{4.2}$$

and

$$\frac{d\phi}{dn} = \frac{i}{a_1}\left(\frac{\partial a_0}{\partial j}\right)_{\phi,\Omega \text{ const}} \tag{4.4}$$

The energy, in physical space, of the wavepacket is defined as,

$$E(n) = \int_{x}^{x_J} |A(x,t_n)|^2 \, dx$$

$$= \int_{0}^{J} |A(j,n)|^2 \frac{dx}{dj} \, dj \tag{4.6}$$

Hence the wave energy density in computational space is defined to be,

$$\rho(j,n) = \alpha(j) \, |A(j,n)|^2 \tag{4.7}$$

where

$$\alpha(j) = \frac{dx}{dj} \tag{4.8}$$

Using (4.3) and the notation that $\bar{A}$ is the complex conjugate of A it follows that

$$\frac{\partial \rho}{\partial n} + \frac{\partial}{\partial j}(r_g \rho) = \frac{\partial}{\partial n}(\alpha A\bar{A}) + \frac{\partial}{\partial j}(r_g \alpha A\bar{A})$$

$$= \alpha\bar{A}\left(\frac{\partial A}{\partial n} + r_g \frac{\partial A}{\partial j}\right) + \alpha A\left(\frac{\partial \bar{A}}{\partial n} + r_g \frac{\partial \bar{A}}{\partial j}\right) + A\bar{A} \frac{\partial}{\partial j}(r_g \alpha)$$

$$= \alpha\bar{A} \frac{dA}{dn} + \alpha A \frac{d\bar{A}}{dn} + A\bar{A} \frac{\partial}{\partial j}(r_g \alpha)$$

$$= \alpha\bar{A}\varepsilon A + \alpha A\overline{\varepsilon A} + A\bar{A} \frac{\partial}{\partial j}(r_g \alpha)$$

$$= \left(\varepsilon + \bar{\varepsilon} + \frac{1}{\alpha} \frac{\partial}{\partial j}(r_g \alpha)\right) \rho \qquad (4.9)$$

Hence,

$$\frac{dE}{dn} = \int_0^J \frac{\partial \rho}{\partial n}\, dj$$

$$= \int_0^J - \frac{\partial}{\partial j}(r_g \rho) + \left(\varepsilon + \bar{\varepsilon} + \frac{1}{\alpha} \frac{\partial}{\partial j}(r_g \alpha)\right) \rho \quad dj$$

$$= - r_g \rho \Big|_0^J + \int_0^J \left(\varepsilon + \bar{\varepsilon} + \frac{1}{\alpha} \frac{\partial}{\partial j}(r_g \alpha)\right) \rho\, dj \qquad (4.10)$$

If the wavepacket is in the interior of the domain away from the boundaries the energy flux $r_g \rho$ at the boundaries $j=0,J$ is zero. Also assuming as before that $L_A << L_C$ then $\varepsilon + \bar{\varepsilon} + \frac{1}{\alpha} \frac{\partial}{\partial j}(r_g \alpha)$ is approximately constant over the length of the wavepacket, so

$$\frac{dE}{dn} = \left(\varepsilon + \bar{\varepsilon} + \frac{1}{\alpha} \frac{\partial}{\partial j}(r_g \alpha)\right) \int_0^J \rho\, dj$$

$$= \left(\varepsilon + \bar{\varepsilon} + \frac{1}{\alpha} \frac{\partial}{\partial j}(r_g \alpha)\right) E \qquad (4.11)$$

Thus equations (4.2), (4..., and (4.11) completely describe the motion of the wavepacket particle in the interior of the computational domain.

When the wavepacket reaches the boundary it interacts with the boundary conditions to produce one or more reflected wavepackets with the same frequency but different wavenumbers. The only case for which it is easy to incorporate boundary conditions is when M=2 and

$$r_g(\phi_1, j) > 0$$

$$r_g(\phi_2, j) < 0$$

An example of a scheme satisfying this condition is the trapezoidal method which was applied to the model convection problem in the introduction.

An additional assumption is that

$$|r_g(\phi_1, j)|, \ |r_g(\phi_2, j)| \ll J$$

so that it takes much more than one time step for a wavepacket to travel from one boundary to the other.

Suppose that initially there is one wavepacket with wavenumber ϕ_1 as in the introductory example. The wavepacket travels to the right with position and energy determined by the equations of motion previously derived (4.2), (4.4) and (4.11). When the wavepacket reaches the boundary at j=J a proportion of the energy E_1 is reflected into a left travelling wavepacket of frequency Ω, wavenumber ϕ_2 and energy E_2. Figure 1 illustrates this interaction.

Equation (4.10) is

$$\frac{dE}{dn} = - r_g \rho \Big|_0^J + \int_0^J \left(\varepsilon + \bar{\varepsilon} + \frac{1}{\alpha} \frac{\partial}{\partial j}(r_g \alpha) \right) \rho \, dj \qquad (4.10)$$

The outgoing energy flux is

$$r_g(\phi_1,J)\ \rho_1(J,n) = r_g(\phi_1,J)\ \alpha(J)\ |A_1(J,n)|^2 \qquad (4.12)$$

The incoming energy flux is

$$r_g(\phi_2,J)\ \rho_2(J,n) = r_g(\phi_2,J)\ \alpha(J)\ |A_2(J,n)|^2 \qquad (4.13)$$

The amplitude reflection coefficient R_J defined by

$$R_J = \frac{A_2(J,n)}{A_1(J,n)} \qquad (4.14)$$

is a function of Ω,ϕ_1,ϕ_2 determined by the asymptotic boundary condition.

The energy flux entering the reflected wavepacket is a factor

$$\left| \frac{r_g(\phi_2,J)}{r_g(\phi_1,J)} \right| \ |R_J|^2$$

greater, or less, than the energy flux leaving the incident wavepacket and so the total energy of the reflected wavepacket is given by,

$$E_2 = \left| \frac{r_g(\phi_2,J)}{r_g(\phi_1,J)} \right| \ |R_J|^2\ E_1 \qquad (4.15)$$

The reflected wavepacket travels left according to the equations of motion for a wave 2 wavepacket until it reaches j=0 where it is reflected into a right travelling wave 1 wavepacket. The reflected energy E_1 is given by,

$$E_1 = \left| \frac{r_g(\phi_1,0)}{r_g(\phi_2,0)} \right| \ |R_0|^2\ E_2 \qquad (4.16)$$

where $$R_0 = \frac{A_1(0,n)}{A_2(0,n)} \qquad (4.17)$$

is determined by the asymptotic boundary condition.

Figure 3 below shows the particle-like path of the wavepacket in computational space.

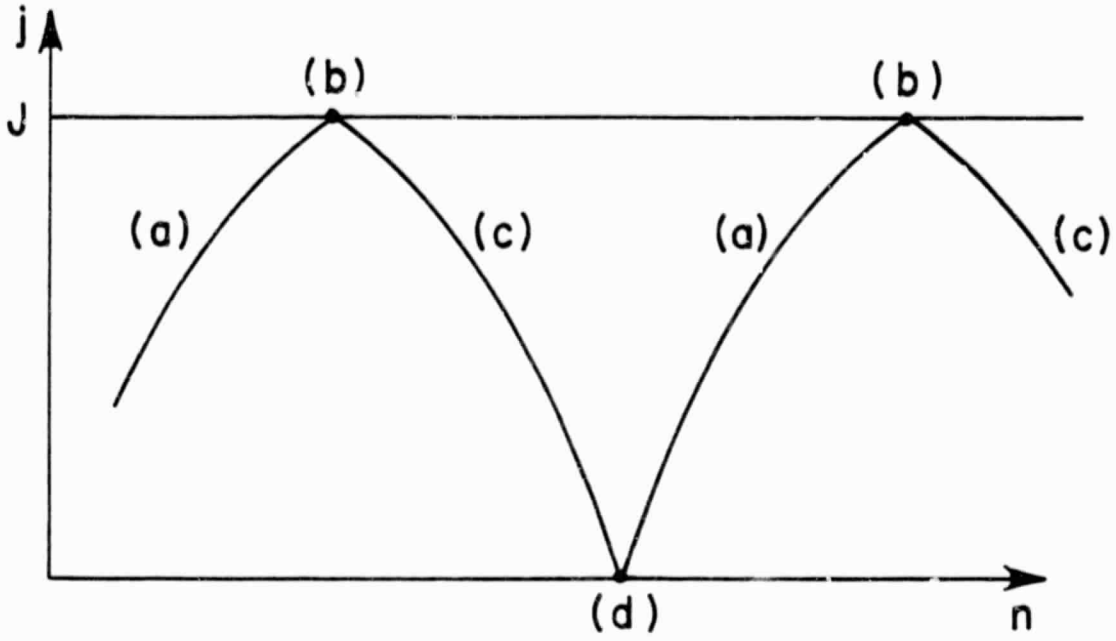


FIGURE 3. WAVEPACKET PATH IN COMPUTATIONAL SPACE

In summary the equations for the different parts of the path are,

(a) $$\frac{dj}{dn} = r_g(\phi_1, j) \qquad (4.2)$$

$$\frac{d\phi_1}{dn} = \frac{i}{a_1}\left(\frac{\partial a_0}{\partial j}\right)_{\phi,\Omega \text{ const}} \qquad (4.4)$$

$$\frac{dE_1}{dn} = \left\{ \varepsilon + \bar{\varepsilon} + \frac{1}{\alpha}\frac{\partial}{\partial j}(r_g \alpha) \right\} E_1 \qquad (4.11)$$

(b) $$E_2 = \left| \frac{r_g(\phi_2, J)}{r_g(\phi_1, J)} \right| \left| R_J \right|^2 E_1 \qquad (4.15)$$

(c) $$\frac{dj}{dn} = r_g(\phi_2, j) \qquad (4.2)$$

$$\frac{d\phi_2}{dn} = \frac{i}{a_1} \left(\frac{\partial a_0}{\partial j} \right)_{\phi, \Omega \text{ const}} \qquad (4.4)$$

$$\frac{dE_2}{dn} = \left(\varepsilon + \bar{\varepsilon} + \frac{1}{\alpha} \frac{\partial}{\partial j}(r_g \alpha) \right) E_2 \qquad (4.11)$$

(d) $$E_1 = \left| \frac{r_g(\phi_1, 0)}{r_g(\phi_2, 0)} \right| \left| R_0 \right|^2 E_2 \qquad (4.16)$$

In (a) and (c) the dispersion relation can be used as a check on the accuracy of the numerical integration of the equations or can replace the equation for the variation of ϕ.

The total number of time steps for a round trip from 0 to J and back again to 0 is

$$N = \int_0^J \left[r_g(\phi_1, j) \right]^{-1} - \left[r_g(\phi_2, j) \right]^{-1} dj \qquad (4.18)$$

The energy growth of a wavepacket travelling from 0 to J is given by

$$\frac{d}{dn} \ln(E_1) = \frac{1}{E} \frac{dE_1}{dn}$$

$$= \varepsilon + \bar{\varepsilon} + \frac{1}{\alpha} \frac{\partial}{\partial j}(r_g \alpha) \qquad (4.19)$$

so $$\frac{d}{dj} \ln(E_1) = \frac{d}{dn} \ln(E_1) \Big/ \frac{dj}{dn}$$

$$= \frac{\varepsilon + \bar{\varepsilon}}{r_g} + \frac{1}{r_g \alpha} \frac{\partial}{\partial j}(r_g \alpha)$$

$$= \frac{\varepsilon + \bar{\varepsilon}}{r_g} + \frac{\partial}{\partial j}[\ \ln(r_g \alpha)\ ] \qquad (4.20)$$

Hence,

$$\ln[E_1(J)] = \ln[E_1(0)] + \int_0^J \frac{\varepsilon + \bar{\varepsilon}}{r_g}\, dj + \ln(r_g \alpha)\Big|_0^J \qquad (4.21)$$

$$\text{so } E_1(J) = \left| \frac{r_g(\Phi_1, J)}{r_g(\Phi_1, 0)} \right| \exp \int_0^J \frac{\varepsilon + \bar{\varepsilon}}{r_g}\, dj \quad E_1(0) \qquad (4.22)$$

Similarly the energy growth of the reflected wavepacket as it travels from J to 0 is

$$\text{so } E_2(0) = \left| \frac{r_g(\Phi_2, 0)}{r_g(\Phi_2, J)} \right| \exp \int_0^J - \frac{\varepsilon + \bar{\varepsilon}}{r_g}\, dj \quad E_2(J) \qquad (4.23)$$

Combining (4.15),(4.16),(4.22) and (4.23) the round trip energy amplification factor λ is

$$\lambda = \left| \frac{r_g(\Phi_1, J)}{r_g(\Phi_1, 0)} \right| \exp \int_0^J \left(\frac{\varepsilon + \bar{\varepsilon}}{r_g} \right)_1 dj \cdot \left| \frac{r_g(\Phi_2, J)}{r_g(\Phi_1, J)} \right| |R_J|^2$$

$$\cdot \left| \frac{r_g(\Phi_2, 0)}{r_g(\Phi_2, J)} \right| \exp \int_0^J - \left(\frac{\varepsilon + \bar{\varepsilon}}{r_g} \right)_2 dj \cdot \left| \frac{r_g(\Phi_1, 0)}{r_g(\Phi_2, 0)} \right| |R_0|^2$$

$$= |R_0\ R_J|^2 \quad \exp \int_0^J \left(\frac{\varepsilon + \bar{\varepsilon}}{r_g} \right)_1 - \left(\frac{\varepsilon + \bar{\varepsilon}}{r_g} \right)_2 dj \qquad (4.24)$$

where $\left(\frac{\varepsilon + \bar{\varepsilon}}{r_g} \right)_1$ is evaluated at Φ_1, j

and $\left(\frac{\varepsilon + \bar{\varepsilon}}{r_g} \right)_2$ is evaluated at Φ_2, j

The condition for stability is

$\lambda < 1$

The equivalent average decay rate σ is

$$\sigma = -\frac{1}{N} \ln(\lambda) \qquad (4.25)$$

When M>2 and there are M_+ waves with positive r_g and M_- waves with negative r_g, one wavepacket with positive r_g reaching j=J produces M_+ reflected wavepackets with negative r_g, and one wavepacket with negative r_g reaching j=0 produces M_+ reflected wavepackets with positive r_g. Thus the total number of wavepackets increases with time exponentially. Since each wavepacket has finite length this ultimately leads to the problem of determining the effect of interference between overlapping wavepackets. In general the sum of the energies of two wavepackets is not equal to the energy of the sum of the wavepackets. If the wavepackets are identical the latter is twice the former while if the second wavepacket has opposite sign to the first the latter is zero.

5. Asymptotic Stability and Convergence Analysis I

5.1 Theory

In addition to the asymptotic approximations in chapter 3 this chapter assumes that M=2 and there is one boundary condition at both j=0 and j=J and if ϕ_1 is real at j=0 then ϕ_1 and ϕ_2 are real over the whole domain.

Examples of methods satisfying these conditions are the trapezoidal method applied to the model convective problem with variable CFL number r, and the backward Euler method applied to the model convective problem with constant CFL number. Methods which do not satisfy these conditions include Lax-Wendroff and Runge-Kutta type schemes. For these methods the general stability analysis of chapter 6 is required.

As explained in §3.1 a standard Fourier series analysis of the problem with constant coefficients and periodic boundary conditions shows that the eigenfrequencies are $\Omega(\phi)$ where ϕ is a real wavenumber satisfying the periodic boundary conditions and $\Omega(\phi)$ is the corresponding frequency given by the dispersion relation.

The common use of a Fourier series analysis to predict the stability of problems with non-periodic boundary conditions implicitly assumes that for real wavenumber ϕ, $\Omega(\phi)$ is a close approximation to an actual eigenfrequency. This chapter follows that assumption, calculates a correction Ω' to this $\Omega(\phi)$ due to the boundary conditions, and then determines the validity of the assumption for this particular class of methods based on the asymptotic errors.

For real ϕ_1, ϕ_2 and the corresponding complex Ω the general solution is,

$$U_j^n = A_1(j,n) \exp\left(i\left(\int_0^j \phi_1 \, d\xi - n\Omega\right)\right) + A_2(j,n) \exp\left(i\left(\int_0^j \phi_2 \, d\xi - n\Omega\right)\right) \tag{5.1}$$

where the amplitudes A_1 and A_2 satisfy the asymptotic amplitude equations

$$\frac{\partial A_m}{\partial n} + (r_g)_m \frac{\partial A_m}{\partial j} = \varepsilon_m A_m \qquad m=1,2 \tag{5.2}$$

and boundary conditions

$$b_1(\Omega,\phi_1)\ A_1(0,n) + b_1(\Omega,\phi_2)\ A_2(0,n) = 0 \tag{5.3}$$

$$b_2(\Omega,\phi_1) \exp\left(i \int_0^J \phi_1 \, dj\right) A_1(J,n)$$

$$+ \quad b_2(\Omega,\phi_2) \exp\left(i \int_0^J \phi_2 \, dj\right) A_2(J,n) = 0 \tag{5.4}$$

A linear system of finite difference equations with time-independent coefficients and boundary conditions has eigenmode solutions of the form,

$$U_j^n = (U_\beta)_j \exp(-in\Omega_\beta) \tag{5.5}$$

where Ω_β is the complex eigenfrequency and U_β is time-independent.

Suppose the frequency Ω in (5.1) is close to Ω_β. Define Ω' by

$$\Omega_\beta = \Omega + \Omega' \tag{5.6}$$

Thus equating (5.1) and (5.5)

$$(U_\beta)_j \exp(-in\Omega') =$$

$$A_1(j,n) \exp\left(i \int_0^j \phi_1 \, d\xi\right) + A_2(j,n) \exp\left(i \int_0^j \phi_2 \, d\xi\right) \quad (5.7)$$

$$\text{so } A_1(j,n) = \exp(-in\Omega') \, A_1(j,0) \quad (5.8)$$

$$A_2(j,n) = \exp(-in\Omega') \, A_2(j,0) \quad (5.9)$$

and

$$(U_\beta)_j = A_1(j,0) \exp\left(i \int_0^j \phi_1 \, d\xi\right) + A_2(j,0) \exp\left(i \int_0^j \phi_2 \, d\xi\right) \quad (5.10)$$

Substituting (5.8),(5.9) into (5.2) gives

$$\frac{\partial}{\partial j}[A_m(j,0)] = \left(\frac{\varepsilon + i\Omega'}{r_g}\right)_m A_m(j,o) \qquad m=1,2 \quad (5.11)$$

which can be integrated to give

$$A_m(J,0) = A_m(0,0) \exp\left(\int_0^J \left(\frac{\varepsilon + i\Omega'}{r_g}\right)_m dj\right) \quad (5.12)$$

The boundary conditions (5.3),(5.4) then become in matrix form

$$\underset{\sim}{B} \, \tilde{A} = 0 \quad (5.13)$$

$$\text{where} \quad \tilde{A} = \begin{pmatrix} A_1(0,0) \\ A_2(0,0) \end{pmatrix} \quad (5.14)$$

and

$$\underset{\sim}{B} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \quad (5.15a)$$

$$B_{11} = b_1(\Omega, \phi_1) \quad (5.15b)$$

$$B_{12} = b_1(\Omega, \Phi_2) \tag{5.15c}$$

$$B_{21} = b_2(\Omega, \Phi_1) \exp\left(i \int_0^J \Phi_1 \, dj \right) \exp\left(\int_0^J \left(\frac{\varepsilon + i\Omega'}{r_g} \right)_1 dj \right) \tag{5.15d}$$

$$B_{22} = b_2(\Omega, \Phi_2) \exp\left(i \int_0^J \Phi_2 \, dj \right) \exp\left(\int_0^J \left(\frac{\varepsilon + i\Omega'}{r_g} \right)_2 dj \right) \tag{5.15e}$$

A non-zero solution $\tilde{A}$ of (5.13) exists if, and only if,

$$\det \tilde{\tilde{B}} = 0 \tag{5.16}$$

Hence

$$\exp\left(i\Omega' \int_0^J \left[r_g(\Phi_1, j) \right]^{-1} - \left[r_g(\Phi_2, j) \right]^{-1} dj \right)$$

$$= \frac{b_2(\Omega, \Phi_2)\, b_1(\Omega, \Phi_1)}{b_2(\Omega, \Phi_1)\, b_1(\Omega, \Phi_2)} \exp\left(i \int_0^J \Phi_2 - \Phi_1 \, dj \right) \exp\left(\int_0^J \left(\frac{\varepsilon}{r_g} \right)_2 - \left(\frac{\varepsilon}{r_g} \right)_1 dj \right) \tag{5.17}$$

the right hand side of (5.17) can be expressed as its magnitude

$$\left| \frac{b_2(\Omega, \Phi_2)\, b_1(\Omega, \Phi_1)}{b_2(\Omega, \Phi_1)\, b_1(\Omega, \Phi_2)} \right| \exp\left(\mathrm{Re} \int_0^J \left(\frac{\varepsilon}{r_g} \right)_2 - \left(\frac{\varepsilon}{r_g} \right)_1 dj \right)$$

multiplied by a phase factor $\exp(i\Psi)$ where

$$\Psi = \arg \frac{b_2(\Omega, \Phi_2)\, b_1(\Omega, \Phi_1)}{b_2(\Omega, \Phi_1)\, b_1(\Omega, \Phi_2)} + \int_0^J \Phi_2 - \Phi_1 \, dj + \mathrm{Im} \left(\int_0^J \left(\frac{\varepsilon}{r_g} \right)_2 - \left(\frac{\varepsilon}{r_g} \right)_1 dj \right) \tag{5.18}$$

If Φ_1, Φ_2 are chosen so that $\Psi/2\pi$ is an integer then (5.17) reduces to

$$\Omega' = -\frac{i}{N} \ln \left| \frac{b_2(\Omega,\phi_2)\ b_1(\Omega,\phi_1)}{b_2(\Omega,\phi_1)\ b_1(\Omega,\phi_2)} \right| + \mathrm{Re}\left(\int_0^J \left(\frac{\varepsilon}{r_g}\right)_2 - \left(\frac{\varepsilon}{r_g}\right)_1 dj \right) \tag{5.19}$$

where as defined in chapter 4,

$$N = \int_0^J [\ r_g(\phi_1,j)\ ]^{-1} - [\ r_g(\phi_2,j)\ ]^{-1}\ dj \tag{4.18}$$

The stability criterion is

$$\mathrm{Im}(\Omega_\beta) = \mathrm{Im}(\Omega + \Omega') < 0 \tag{5.20}$$

Thus the frequency Ω resulting from a normal Von Neumann analysis is corrected by an amount Ω' due to boundary conditions and variable coefficients. This approach, using Ω as an initial approximation to Ω_β the actual eigenfrequency, is valid provided the asymptotic errors are small compared to Ω'.

The asymptotic error is $O(\ L_C^{-2},\ T_A^{-2}\ ) = O(\ L_C^{-2},\ \Omega'^2\ )$

Now $N = O\ \frac{J}{r_g}$

so if $r_g << J$ then $N >> 1$

and hence $\Omega' << 1$ except near frequencies for which

$$\left| \frac{b_2(\Omega,\phi_2)\ b_1(\Omega,\phi_1)}{b_2(\Omega,\phi_1)\ b_1(\Omega,\phi_2)} \right|$$

is zero, or infinite, which usually occurs at $\Omega=0$. However these frequencies are heavily damped by the boundary conditions and so an accurate estimate of their eigenfrequencies is not essential. This method gives accurate asymptotic values near the critical frequencies which are least damped and which therefore determine the

overall spectral radius of the scheme.

5.2 Example

This example is the Backward Euler method applied to the model convective problem with constant CFL number r and space extrapolation at the downstream boundary.

The dispersion relation is

$$\exp(i\Omega) - 1 = ir \sin(\phi) \tag{3.63}$$

so if ϕ_1 is real, $\phi_2 = \pi - \phi_1$ is also real

$$b_1(\Omega,\phi) = 1 \tag{3.82}$$

$$b_2(\Omega,\phi) = 2i \exp(-i\phi/2) \sin(\phi/2) \tag{3.88}$$

$$\text{so} \quad \frac{b_2(\Omega,\phi_2)\, b_1(\Omega,\phi_1)}{b_2(\Omega,\phi_1)\, b_1(\Omega,\phi_2)} = \exp[-i(\phi_2-\phi_1)/2] \frac{\sin(\phi_2/2)}{\sin(\phi_1/2)}$$

$$= \exp[-i(\phi_2-\phi_1)/2] \frac{\cos(\phi_1/2)}{\sin(\phi_1/2)}$$

$$= \exp[-i(\phi_2-\phi_1)/2] \cot(\phi_1/2) \tag{5.21}$$

Since $\frac{\partial r}{\partial j} = 0$, $\varepsilon = 0$

Hence $\Psi = -(\phi_2 - \phi_1)/2 + J(\phi_2 - \phi_1)$

$$= (J - 1/2)(\pi - 2\phi_1) \tag{5.22}$$

$\Psi/2\pi = n$, where n is an integer, implies

$$\phi_1 = \pi/2 - 2\pi n / (2J-1) \tag{5.23}$$

The group CFL numbers are

$$r_g(\phi_1) = r \cos(\phi_1) \exp(-i\Omega)$$

$$= r \cos(\phi_1) / (1 + ir \sin(\phi_1))$$

$$= r\cos(\phi_1)(1 - ir\sin(\phi_1)) \,/\, (1 + r^2\sin^2(\phi_1)) \quad (5.24)$$

$$r_g(\phi_2) = -\, r\cos(\phi_1)(1 - ir\sin(\phi_1))/(1 + r^2\sin^2(\phi_1)) \quad (5.25)$$

Hence,

$$\Omega' = -\,\frac{ir\cos(\phi_1)\,(1 - ir\sin(\phi_1))}{2J\,(\,1 + r^2\sin^2(\phi_1)\,)}\ \ln\left(\cot(\phi_1/2)\right) \quad (5.26)$$

$$\exp[-2\,\mathrm{Im}(\Omega)] = |\ \exp(i\Omega)\ |^2$$

$$= |\ 1 + ir\sin(\phi_1/2)\ |^2$$

$$= 1 + r^2\sin^2(\phi_1/2) \quad (5.27)$$

so $$\mathrm{Im}(\Omega) = -\,\frac{1}{2}\ln\,[\ 1 + r^2\sin^2(\phi_1/2)\ ] \quad (5.28)$$

Hence

$$\mathrm{Im}(\ \Omega + \Omega'\ ) = -\,\frac{1}{2}\ln\,[\ 1+ r^2\sin^2(\phi_1/2)\ ] -$$

$$\frac{r\cos(\phi_1/2)}{2J(1 + r^2\sin^2(\phi_1/2)}\ \ln\,[\ \cot(\phi_1/2)\ ] \quad (5.29)$$

Define the decay rate σ to be

$$\sigma = -\,\mathrm{Im}(\ \Omega + \Omega'\ ) \quad (5.30)$$

For small $\phi \ll 1$

$$\sigma \approx \frac{1}{2}\,r^2\phi^2 + \frac{r}{2J}\ln(2/\phi) \quad (5.31)$$

$$\frac{d\sigma}{d\phi} \approx r^2\phi - \frac{r}{2J}\,\phi^{-1} \quad (5.32)$$

Thus $$\frac{d\sigma}{d\phi} = 0 \quad \text{at } \phi = (2rJ)^{-\frac{1}{2}} \quad (5.33)$$

so $$\min\,\sigma(\phi) = \frac{r}{4J} + \frac{r}{4J}\ln(8rJ) \quad (5.34)$$

The spectral radius λ is

$$\lambda = \max \exp(-\sigma)$$

$$= \max (1-\sigma)$$

$$= 1 - \frac{r}{4J} [\ 1 + \ln(8rJ)\ ] \qquad (5.35)$$

§6.2 continues this example proving rigorously that every eigenmode is stable, and deriving an asymptotic expression for the decay rate of the eigenmodes near $\phi_1=0$, showing as expected that the decay rate is greater than the minimum decay rate obtained above and so the above analysis is valid in calculating the spectral radius.

6. Asymptotic Stability and Convergence Analysis II

6.1 Theory

This chapter continues the analysis of chapter 3 without any additional assumptions or approximations. The eigenmodes of a linear system of finite difference equations with time-independent coefficients and boundary conditions vary exponentially with time so a general eigenmode can be written as

$$U_j^n = \exp(-in\Omega) \sum_{m=1}^{M} A_m(j) \exp\left(i \int_0^j \phi_m \, d\xi \right) \tag{6.1}$$

Note that the amplitudes A_m are independent of n. The time evolution of the eigenmode is contained solely in the term $\exp(-in\Omega)$. The complex amplitudes A_m each satisfy their asymptotic amplitude equation,

$$r_g \frac{\partial A_m}{\partial j} = \varepsilon A_m \tag{6.2}$$

$$\text{so} \qquad A_m(J) = A_m(0) \exp\left(\int_0^J \left(\frac{\varepsilon}{r_g} \right)_m dj \right) \tag{6.3}$$

$\left(\frac{\varepsilon}{r_g} \right)_m$ denotes $\frac{\varepsilon}{r_g}$ evaluated at ϕ_m, j

Now that $A_m(J)$ is related to $A_m(0)$ the M asymptotic boundary condions can be written in tensor summation form as

$$B_{km} A_m = 0 \tag{6.4}$$

where $$A_m = A_m(0) \tag{6.5}$$

and

$$B_{km} = \begin{cases} b_k(\Omega,\phi_m) & \text{if the } k^{th} \text{ boundary condition is applied at } j=0 \\ b_k(\Omega,\phi_m)\ \exp\left(\int_0^J \left(\frac{\varepsilon}{r_g}\right)_m + i\phi_m \ dj \right) & \text{if the } k^{th} \text{ boundary condition is applied at } j=J \end{cases} \tag{6.6}$$

The term $b_k(\Omega,\phi_m)\ \exp\left(\int_0^J i\phi_m \ dj \right)$ comes from (3.77) and the term $\exp\left(\int_0^J \left(\frac{\varepsilon}{r_g}\right)_m \ dj \right)$ is the factor relating $A_m(J)$ to $A_m(0)$ in (6.3).

A non-zero solution to (6.4) exists if, and only if,

$$\det(B) = 0 \tag{6.7}$$

Since all the elements of B are implicit functions of Ω this is the equation which determines the eigenfrequencies.

If the coefficients are constant all of the A_m are constant and so there are no asymptotic errors. The theory is then exact and is identical to the P-stability analysis of Beam, Warming and Yee [5]. If the coefficients are variable the asymptotic error is of order $O(L_c^{-2}, J^{-1}L_c^{-1})$. The $O(L_c^{-2})$ comes from neglecting second derivatives of A_m and ϕ in the asymptotic amplitude equations. The $O(J^{-1}L_c^{-1})$ comes from neglecting the first derivatives of A_m and ϕ_m in the asymptotic boundary conditions.

For all but the very simplest problems it will be

impossible to solve (6.7) analytically to obtain the eigenfrequencies. Three possible numerical approaches are outlined below.

6.1.1 Iterative Solution

Suppose Ω^n is a good approximation to an eigenfrequency Ω and

$$\Omega^{n+1} = \Omega^n + \Delta\Omega^n \tag{6.8}$$

is to be a better approximation.

The terms in the definition of B_{km} which change most rapidly with variations in Ω are $\exp \int_0^J i\phi_m \, dj$ since these are oscillatory functions because ϕ_m usually has a real component of order O(1).

$$\frac{\partial \phi}{\partial \Omega}_m = \left(\frac{\partial \Omega}{\partial \phi}\right)_m^{-1}$$

$$= \left(r_g\right)_m^{-1} \tag{6.9}$$

Hence $$\phi_m^{n+1} = \phi_m^n + \left(r_g\right)_m^{-1} \Delta\Omega \tag{6.10}$$

Subscript m means evaluated at ϕ_m, j

Superscript n means evaluated at Ω^n

Thus

$$\exp\left(\int_0^J i\phi_m^{n+1} \, dj\right) = \exp\left(\int_0^J i\phi_m^n \, dj\right) \exp\left(\int_0^J i\Delta\Omega \left(r_g\right)^{-1} \, dj\right) \tag{6.11}$$

Define

$$B^n_{km} = \begin{cases} b_k(\Omega^n, \phi^n_m) & \text{if the } k^{th} \text{ boundary condition is applied at } j=0 \\ b_k(\Omega^n, \phi^n_m)\ \exp\left(\int_0^J \left(\frac{\varepsilon}{r_g}\right)^n_m + i\phi^n_m\ dj\right) \exp\left(i\Delta\Omega \int_0^J \left(r_g\right)^n_m{}^{-1}\ dj\right) & \text{if the } k^{th} \text{ boundary condition is at } j=J \end{cases} \qquad (6.12)$$

Then choose $\Delta\Omega$ to be the smallest root of

$$\det(B^n) = 0 \qquad (6.13)$$

The method in chapter 5 performs one step in this iterative procedure. As explained in chapter 5 the asymptotic error remaining after a correction $\Delta\Omega$ is $O(\Delta\Omega^{-2}, L_c^{-2})$. For constant coefficients this gives quadratic convergence to the true eigenfrequency. For variable coefficients it will converge to a frequency which differs from the true eigenfrequency by an asymptotic error of order $O(L_c^{-2})$.

The solution procedure will fail as described in chapter 5 near frequencies for which $b_k(\Omega, \phi_m)$ is zero since the fractional variation in $b_k(\Omega, \phi_m)$ is comparable to, or may exceed, the fractional variation in $\exp\left(\int_0^J i\phi_m\ dj\right)$.

6.1.2 Newton-Raphson Solution

The Newton-Raphson solution procedure is

$$\Omega^{n+1} = \Omega^{n} - \frac{\det[B(\Omega^{n})]}{\frac{\partial}{\partial\Omega}\det[B(\Omega^{n})]} \qquad (6.14)$$

Although this seems straightforward there will be problems in practice because det(B)=0 has many roots and so $\frac{\partial}{\partial\Omega}\det[B(\Omega^{n})]$ has many zeros near which the Newton-Raphson iterative procedure is badly behaved.

6.1.3 Stability Domain Method

If z is defined as

$$z = \exp(i\Omega) \qquad (6.15)$$

then the condition for stability is that det(B) has no zeros in $|z| < 1$. In very simple cases this criterion can be tested analytically (an example is given in §6.2). In more complicated cases because det[B(z)] is an analytic function of z the 'Principle of the Argument' method outlined in appendix A.2 can be used to find whether det[B(z)] has any zeros in $|z| < 1$. To find the spectral radius λ the same method can be used to find the largest λ for which there are no zeros in $|z| < \lambda$.

6.2 Example

The example is the Trapezoidal method applied to the model convective problem with constant CFL number r and space extrapolation at the downstream boundary.

The dispersion relation is

$$\tan(\Omega/2) = \frac{r}{2}\sin(\phi) \tag{3.41}$$

Define $\kappa = \exp(i\phi)$ (6.16)

$z = \exp(-i\Omega)$ (6.17)

Then
$$\tan(\Omega/2) = \frac{1}{i}\frac{\exp(i\Omega/2) - \exp(-i\Omega/2)}{\exp(i\Omega/2) + \exp(-i\Omega/2)}$$

$$= \frac{1}{i}\frac{1 - \exp(-i\Omega)}{1 + \exp(-i\Omega)}$$

$$= \frac{1}{i}\frac{1-z}{1+z} \tag{6.18}$$

$$\sin(\phi) = \frac{1}{2i}\left[\exp(i\phi) - \exp(-i\phi)\right]$$

$$= \frac{1}{2i}\left(\kappa - \kappa^{-1}\right) \tag{6.19}$$

Thus the dispersion relation becomes

$$\frac{1-z}{1+z} = \frac{r}{4}\left(\kappa - \kappa^{-1}\right) \tag{6.20}$$

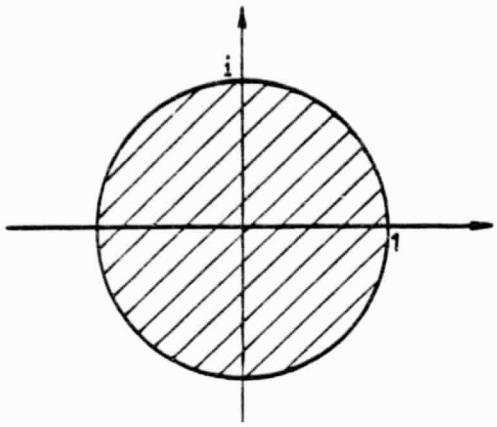
Note that if κ_1 is one solution then $\kappa_2 = -\kappa_1^{-1}$ is the other solution. The condition for stability is that $|z| < 1$ for all eigenmodes.

Let $z = R\exp(i\theta)$ $\quad R,\theta$ real, $R > 0$ (6.21)

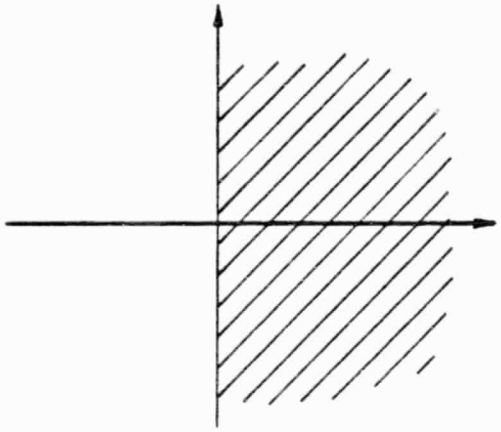
Then
$$\frac{1-z}{1+z} = \frac{1 - R\exp(i\theta)}{1 + R\exp(i\theta)}$$

$$= \frac{[1 - R\exp(i\theta)]\,[1 + R\exp(-i\theta)]}{[1 + R\exp(i\theta)]\,[1 + R\exp(-i\theta)]}$$

i) Stability domain for z



ii) Stability domain for $\frac{1-z}{1+z}$



iii) Stability domain for κ

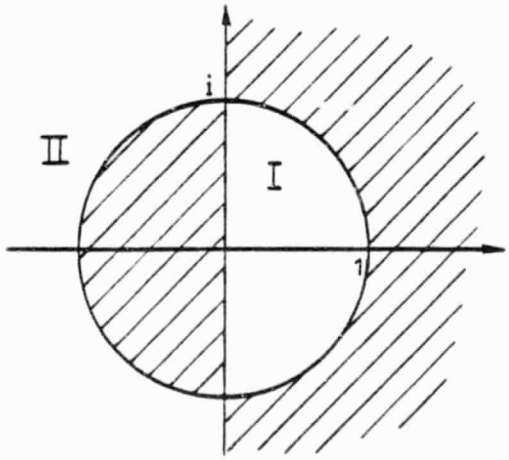


FIGURE 4. STABILITY DOMAINS : SHADED REGIONS ARE STABLE

$$= \frac{1 - 2iR\sin(\theta) - R^2}{1 + 2R\cos(\theta) + R^2}$$

$$= \frac{(1 - R^2) - 2iR\sin(\theta)}{[1 + R\cos(\theta)]^2 + [R\sin(\theta)]^2} \qquad (6.22)$$

Hence $R < 1 \implies \mathrm{Re}\ \frac{1-z}{1+z} > 0$

$R > 1 \implies \mathrm{Re}\ \frac{1-z}{1+z} < 0$

Thus if r is positive then $R \equiv |z| < 1$ if, and only if, $\mathrm{Re}\ \frac{1-z}{1+z} > 0 \iff \mathrm{Re}[\kappa - \kappa^{-1}] > 0$

Now let $\kappa = R\exp(i\theta)$ (6.23)

$$\kappa - \kappa^{-1} = R\cos(\theta) + iR\sin(\theta) - R^{-1}\cos(\theta) - iR\sin(\theta)$$

$$= (R - R^{-1})\cos(\theta) + i(R + R^{-1})\sin(\theta) \qquad (6.24)$$

so $|z| < 1 \iff \mathrm{Re}[\kappa - \kappa^{-1}] > 0$

$\iff$ either $R > 1$, $-\pi/2 < \theta < \pi/2$
or $R < 1$, $\pi/2 < \theta < 3\pi/2$

Figure 4 shows the stability domains for the different variables.

The upstream boundary condition has

$$b_1 = 1 \qquad (3.82)$$

The downstream space extrapolation has

$$b_2 = 1 - \exp(-i\Omega) \qquad (3.88)$$

Hence $B = \begin{pmatrix} 1 & 1 \\ \exp(iJ\phi_1)[1-\exp(-i\phi_1)] & \exp(iJ\phi_1)[1-\exp(-i\phi_1)] \end{pmatrix}$

$$= \begin{pmatrix} 1 & 1 \\ \kappa^J(1 - \kappa^{-1}) & (-\kappa^{-1})^J(1 + \kappa) \end{pmatrix} \qquad (6.25)$$

$\det(B) = 0$ implies

$$\kappa^{2J-1}(\kappa - 1) - (-1)^J(\kappa + 1) = 0 \qquad (6.26)$$

$$\Longrightarrow \quad \frac{\kappa+1}{\kappa-1} = (-1)^J \kappa^{2J-1} \qquad (6.27)$$

Consider the two instability regions I and II marked in figure 4 iii).

In region I $|\kappa + 1| > |\kappa - 1|$ and $|\kappa| > 1$

so $\left| \frac{\kappa+1}{\kappa-1} \right| > \left|(-1)^J \kappa^{2J-1}\right|$

In region II $|\kappa + 1| < |\kappa - 1|$ and $|\kappa| < 1$

so $\left| \frac{\kappa+1}{\kappa-1} \right| < \left|(-1)^J \kappa^{2J-1}\right|$

Hence there are no possible solutions of (6.32) in the two regions of instability. Every eigenmode is stable and so the method is stable. This example is a special example taken from a more general result, proved by Beam, Warming and Yee [5], that the class of A-stable Beam-Warming multistep methods with q^{th} order space extrapolation at the downstream boundary is stable.

Continuing this example asymptotic decay rates for small ϕ_1 can be derived.

$$\phi_1 = 0 \quad \Rightarrow \quad \kappa = 1$$

Let $\kappa = 1 + \delta \qquad (6.28)$

Equation (6.31) becomes

$$(1 + \delta)^{2J-1}\delta - (-1)^J(2 + \delta) = 0 \qquad (6.29)$$

Now $$\lim_{J\to\infty} \left(1 + \frac{x}{J}\right)^J = \exp(x) \qquad (6.30)$$

so provided $J >> 1 >> \delta$

$$(1 + \delta)^{2J} = \exp(2J\delta) \tag{6.31}$$

Hence $$\exp(2J\delta) = 2(-1)^J \delta^{-1} (2 + \delta)(1 + \delta)$$

$$= 2(-1)^J \delta^{-1} \tag{6.32}$$

If J is even this has a solution for real δ

$$2J\delta = \ln(2/\delta) \tag{6.33}$$

so $$\delta = \frac{1}{2J} \ln(2/\delta)$$

$$= \frac{1}{2J} \ln \frac{4J}{\ln(2/\delta)}$$

$$= \frac{\ln(J)}{2J} + \text{terms of order } O\left(\frac{\ln[\ln(J)]}{J}\right) \tag{6.34}$$

Since $\kappa = \exp(i\phi_1)$

$$= 1 + i\phi_1 \tag{6.35}$$

$$\phi_1 = -i\delta$$

$$= -\frac{i \ln(J)}{2J} \tag{6.36}$$

Linearising the dispersion relation about $\phi = 0$ gives

$$\Omega = r\,\phi_1$$

$$= -\frac{ir \ln(J)}{2J} \tag{6.37}$$

The asymptotic decay rate is

$$\sigma = \frac{r \ln(J)}{2J} \tag{6.38}$$

If J is odd the smallest δ solutions are

$$\delta = \frac{r \ln(J)}{2J} \pm \frac{\pi i}{2J} \tag{6.39}$$

and the decay rate is the same.

7. Assorted Examples and Further Developments

7.1 Instability of Backward Euler with Spacetime Extrapolation

Consider the Backward Euler method applied to the model convective problem with constant CFL number r on domain $0 < j < J$ with space-time extrapolation at the downstream boundary and $J >> r >> 1$.

The dispersion relation is

$$\exp(i\Omega) - 1 = ir \sin(\phi) \tag{3.63}$$

The upstream boundary has

$$b_1 = 1 \tag{3.82}$$

The downstream boundary with space-time extrapolation has

$$b_2 = 1 - \exp[i(\Omega-\phi)] \tag{3.94}$$

The eigenfrequency equation, $\det(B)=0$ reduces to

$$\exp(iJ\phi_1)\{1 - \exp[i(\Omega-\phi_1)]\} = \exp(iJ\phi_2)\{1 - \exp[i(\Omega-\phi_2)]\} \tag{7.1}$$

Because $\phi_2 = \pi-\phi_1$ this can be written as

$$\exp(2iJ\phi_1)\{1 - \exp[i(\Omega-\phi_1)]\} - (-1)^J \{1 + \exp[i(\Omega+\phi_1)]\} = 0 \tag{7.2}$$

This and the dispersion relation form two equations in the two unknowns Ω and ϕ_1. Considering only the case in which J is eve one eigenmode is given by,

$$\Omega = \pi + \Omega' \qquad |\Omega'| < \pi \tag{7.3}$$

$$\text{so } ir \sin(\phi_1) = -2 - i\Omega' + O(\Omega'^2) \tag{7.4}$$

$$\phi_1 = \frac{2i}{r} - \frac{\Omega'}{r} + \text{H.O.T.} \qquad (7.5)$$

Substituting into (7.3) with $\Omega'=0$ gives

$$\exp(-4J/r)\{1 + \exp(2/r)\} - 1 + \exp(-2/r) = 0 + \text{H.O.T.} \qquad (7.6)$$

Expanding the exponentials under the approximation $r >> 1$ this reduces to,

$$J = \frac{r}{4}\ln(r) + \text{H.O.T.} \qquad (7.7)$$

For a particular even value of J the value of r satisfying (7.7) makes this eigenmode asymptotically neutrally stable. To find whether increasing r makes it unstable or not (7.2) is differentiated by r.

$$2iJ\frac{d\phi_1}{dr}\exp(2iJ\phi_1)\{1 - \exp[i(\Omega - \phi_1)]\}$$

$$+ \exp(2iJ\phi_1)\{i\frac{d\phi_1}{dr} - i\frac{d\Omega}{dr}\}\exp[i(\Omega - \phi_1)]$$

$$- \{i\frac{d\phi_1}{dr} + i\frac{\Omega}{dr}\}\exp[i(\Omega + \phi_1)] = 0 \qquad (7.8)$$

This is evaluated at $\Omega'=0$, $J = \frac{r}{4}\ln(r)$ so

$$\ln(r)\frac{d\phi_1}{dr} - \frac{1}{r}\{\frac{d\phi_1}{dr} - \frac{d\Omega}{dr}\} + \{\frac{d\phi_1}{dr} + \frac{d\Omega}{dr}\} = 0 + \text{H.O.T.} \qquad (7.9)$$

Hence $$\frac{d\Omega}{dr} = \ln(r)\frac{d\phi_1}{dr} + \text{H.O.T.}$$

$$= -\frac{2i\ln(r)}{r^2} + \text{H.O.T.} \qquad (7.10)$$

As r increases from the neutrally stable value, $\text{Im}(\Omega)$ becomes negative so the eigenmode becomes unstable. Thus this eigenmode is stable only if

$$\frac{r}{4}\ln(r) < J$$

In numerical experiments it is found that this is

the least stable eigenmode and so this condition is both necessary and sufficient. The condition is asymptotically equal to an exact stability condition derived by Beam, Warming and Yee [5].

7.2 Optimum CFL Number For Trapezoidal Method

Consider the Trapezoidal method applied to the model convective problem with constant CFL number r on domain $0 < j < J$ with space extrapolation at the downstream boundary.

The dispersion relation is

$$\tan(\Omega/2) = \frac{r}{2} \sin(\phi) \tag{3.52}$$

The upstream boundary condition has

$$b_1 = 1 \tag{3.82}$$

The downstream boundary has

$$b_2 = 2i \exp(-i\phi/2) \sin(\phi/2) \tag{3.88}$$

Following the stability analysis of chapter 5, $\varepsilon=0$ since r is constant, and the group CFL numbers are given by,

$$r_{g_1} = r \cos(\phi_1) \cos^2(\Omega/2) \tag{3.47}$$

$$r_{g_1} = r \cos(\phi_2) \cos^2(\Omega/2)$$

$$= - r \cos(\phi_1) \cos^2(\Omega/2) \tag{7.11}$$

$$\text{so } \Psi = \arg \frac{b_2(\Omega,\phi_2)\, b_1(\Omega,\phi_1)}{b_2(\Omega,\phi_2)\, b_1(\Omega,\phi_2)} + \int_0^J \phi_2 - \phi_1 \, dj$$

$$+ \operatorname{Im}\left(\int_0^J \left(\frac{\varepsilon}{r_g}\right)_2 - \left(\frac{\varepsilon}{r_g}\right)_1 dj \right) \tag{5.18}$$

$$= (J-1/2)\,(\,\phi_2 - \phi_1\,)$$

$$= (J-1/2)\,(\,\pi - 2\phi_1\,) \tag{7.12}$$

$\Psi/2\pi = n$ where n is an integer implies

$$\phi_1 = \frac{\pi}{2} - \frac{2\pi n}{2J-1} \tag{7.13}$$

$$\Omega' = -\frac{i}{N}\left(\ln\left| \frac{b_2(\Omega,\phi_2)\ b_1(\Omega,\phi_1)}{b_2(\Omega,\phi_2)\ b_1(\Omega,\phi_2)} \right| + \mathrm{Re}\left(\int_0^J \left(\frac{\varepsilon}{r_g}\right)_2 - \left(\frac{\varepsilon}{r_g}\right)_1 dj \right)\right)$$

$$= -\frac{i}{N} \ln[\ \sin(\phi_2/2) \ / \ \sin(\phi_1/2)\ ]$$

$$= -\frac{i}{N} \ln[\ \cot(\phi_1/2)\ ] \tag{7.14}$$

where $$N = \int_0^J [r_g(\phi_1,j)]^{-1} - [r_g(\phi_2,j)]^{-1}\ dj \tag{5.20}$$

$$= \frac{2J}{r\ \cos(\phi_1)\ \cos^2(\Omega/2)} \tag{7.15}$$

The decay rate σ was defined as

$$\sigma = -\ \mathrm{Im}(\ \Omega + \Omega'\ ) \tag{5.30}$$

$$\sigma = \frac{r}{2J} \cos(\phi_1)\ \cos^2(\Omega/2)\ \ln[\cot(\phi_1/2)]$$

$$= \frac{r\ \cos(\phi_1)\ \ln[\ \cot(\phi_1/2)\ ]}{2J\ [\ 1 + \tan^2(\Omega/2)\ ]}$$

$$= \frac{2r\ \cos(\phi_1)\ \ln[\ \cot(\phi_1/2)\ ]}{J\ [\ 4 + r^2\sin^2(\phi_1)\ ]} \tag{7.16}$$

Figure 5 shows the dacay rate σ as a function of r for various values of ϕ.

When a system of finite difference equations is solved using a time-independent approach to a steady state solution, usually the initial conditions and the final solution are smooth so the initial error is smooth i.e. the error is predominantly in the low wavenumber, long wavelength eigenmodes. For this problem suppose that the

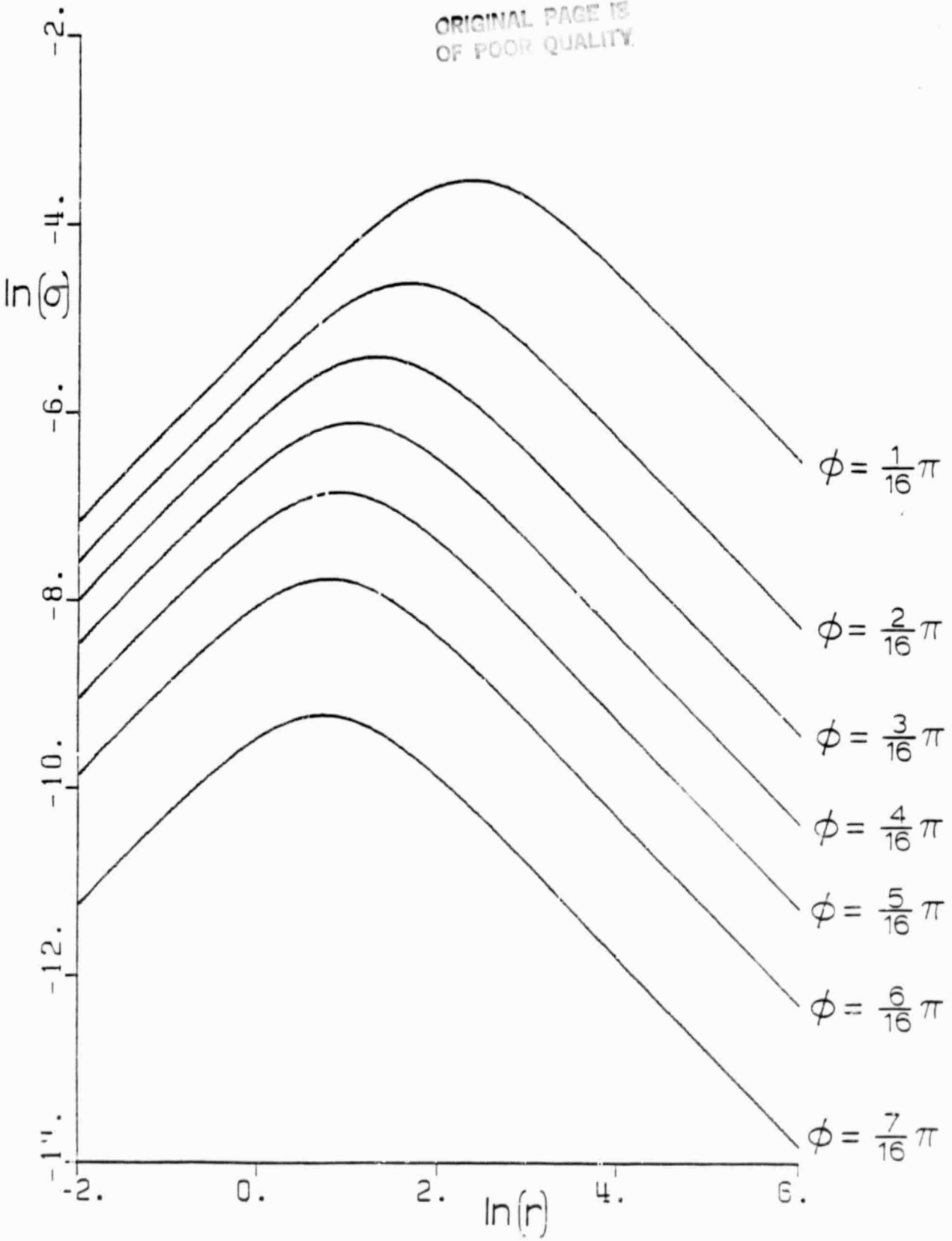


FIGURE 5. LOG-LOG PLOT OF DECAY RATE AGAINST CFL NO.

initial error is in the wavenumber range $0 < \phi < \phi_0$ where ϕ_0 is a constant, $\phi_0 < \pi/2$

$$\frac{\partial \sigma}{\partial \phi} = -\frac{2r}{J}\left(\left(\frac{\sin(\phi)}{4 + r^2\sin^2(\phi)} + \frac{2r^2\sin(\phi)\cos^2(\phi)}{[4 + r^2\sin^2(\phi)]^2}\right)\ln[\cot(\phi/2)] + \frac{\cos(\phi)}{\sin(\phi)[4 + r^2\sin^2(\phi)]}\right) < 0 \qquad (7.17)$$

so over the range given above σ has a minimum at ϕ_0. To maximise the overall rate of convergence the CFL number r can be changed by altering Δt.

$$\frac{\partial \sigma}{\partial r} = \frac{\cos(\phi)\,[\,4 - r^2\sin^2(\phi)\,]}{2J\,[\,4 + r^2\sin^2(\phi)\,]}\ln[\cot(\phi/2)] \qquad (7.18)$$

$$> 0 \quad \text{for} \quad 0 < r < 2/\sin(\phi)$$

$$< 0 \quad \text{for} \quad 2/\sin(\phi) < r$$

so $\sigma(\phi_0)$ is maximised by choosing

$$r = \frac{2}{\sin(\phi_0)} \qquad (7.19)$$

The smoother the initial error the lower the value of ϕ_0 and the higher the optimum CFL number. If the initial error is not at all smooth with ϕ_0 approaching $\pi/2$, the optimum CFL number drops down towards 2.

7.3 Discontinuous CFL Number In Trapezoidal Method

Consider the following Trapezoidal method applied to the model convective problem.

$$\left[\delta_t + \frac{r_{j-\frac{1}{2}}}{2} \mu_t \nabla_x + \frac{r_{j+\frac{1}{2}}}{2} \mu_t \Delta_x \right] U_j^{n+\frac{1}{2}} = 0 \qquad (7.20)$$

Suppose that r is discontinuous

$$r_j = \begin{cases} r_- & , \quad j < 0 \\ r_+ & , \quad j > 0 \end{cases}$$

On the two sides of the discontinuity a general solution of frequency Ω can be written as,

$$U_j^n = \begin{cases} A_1(j,n)\exp[i(j\phi_1 - n\Omega)] + A_2(j,n)\exp[i(j\phi_2 - n\Omega)] & j < 0 \\ \\ A_3(j,n)\exp[i(j\phi_3 - n\Omega)] + A_4(j,n)\exp[i(j\phi_4 - n\Omega)] & j > 0 \end{cases} \qquad (7.21)$$

ϕ_1 and ϕ_2 satisfy the dispersion relation

$$\tan(\Omega/2) = \frac{r_-}{2} \sin(\phi) \qquad (3.52)$$

with $\phi_2 = \pi - \phi_1$

and $0 < \mathrm{Re}(\phi_1) < \pi/2$

ϕ_3 and ϕ_4 satisfy the dispersion relation

$$\tan(\Omega/2) = \frac{r_+}{2} \sin(\phi) \qquad (3.52)$$

with $\phi_4 = \pi - \phi_3$

and $0 < \mathrm{Re}(\phi_3) < \pi/2$

The amplitudes A_1, A_2 are related to A_3, A_4 at j=0 by two equations. The first comes from the requirement that the two expressions for U_j^n are equivalent at j=0.

Thus $$A_1 + A_2 = A_3 + A_4 \qquad (7.22)$$

The second requirement is that

$$[\ \delta_t + \frac{1}{2}\, r_{j-\frac{1}{2}}\, \mu_t \nabla_x + \frac{1}{2}\, r_{j+\frac{1}{2}}\, \mu_t \Delta_x\ ]\ U_0^{n+\frac{1}{2}} = 0 \qquad (7.23)$$

which implies

$$[\ \delta_t + r_-\, \mu_t \nabla_x]\ U_0^{n+\frac{1}{2}} = -\ [\ \delta_t + r_+\, \mu_t \Delta_x]\ U_0^{n+\frac{1}{2}} \qquad (7.24)$$

The left hand side involves U_j^n at j=0,-1 so the first expression for U_j^n is used. The right hand side involves U_j^n at j=0,1 so the second expression is used. Neglecting derivatives of the amplitudes the resulting equation is,

$$\begin{aligned} & A_1\{\ -2i\ \sin(\Omega/2) + r_-\cos(\Omega/2)[1-\exp(-i\phi_1)]\ \} \\ & + A_2\{\ -2i\ \sin(\Omega/2) + r_-\cos(\Omega/2)[1-\exp(-i\phi_2)]\ \} \\ = & - A_3\{\ -2i\ \sin(\Omega/2) + r_+\cos(\Omega/2)[1-\exp(-i\phi_3)]\ \} \\ & - A_4\{\ -2i\ \sin(\Omega/2) + r_+\cos(\Omega/2)[1-\exp(-i\phi_4)]\ \} \end{aligned} \qquad (7.25)$$

Substituting for $\tan(\Omega/2)$ using the dispersion relations and using

$$\exp(i\phi) = \cos(\phi) + i\,\sin(\phi) \qquad (7.26)$$

this reduces to

$$\begin{aligned} & r_-\{\ A_1[1-\cos(\phi_1)] + A_2[1-\cos(\phi_2)]\ \} \\ & = r_+\{\ A_3[1-\cos(\phi_3)] + A_4[1-\cos(\phi_4)]\ \} \end{aligned} \qquad (7.27)$$

In the general problem the two equations relate the amplitudes on either side of the discontinuity at j=0. Now consider the particular problem in which ϕ_1, ϕ_2, ϕ_3, ϕ_4 are all real and A_4 is zero. This is the situation when a wavepacket with wavenumber ϕ_1, and positive group CFL number travels from j<0 to the interface at j=0 producing a transmitted wavepacket with wavenumber ϕ_3 and positive group CFL number and a reflected wavepacket with wavenumber ϕ_2 and negative group CFL number. The reflection coefficient R and transmission coefficient T are defined by,

$$R = A_2 \; / \; A_1 \tag{7.28}$$

$$T = A_3 \; / \; A_1 \tag{7.29}$$

Since $A_4=0$

$$A_1 + A_2 = A_3 \tag{7.30}$$

so $$r_-\{A_1[1-\cos(\phi_1)] + A_2[1-\cos(\phi_2)]\} = r_+[A_1+A_2][1-\cos(\phi_3)] \tag{7.31}$$

Hence $$R = -\frac{r_-[1-\cos(\phi_1)] - r_+[1-\cos(\phi_3)]}{r_-[1-\cos(\phi_2)] - r_+[1-\cos(\phi_3)]} \tag{7.32}$$

and $$T = 1 + R$$

$$= \frac{r_-[1-\cos(\phi_2)] - r_-[1-\cos(\phi_1)]}{r_-[1-\cos(\phi_2)] - r_+[1-\cos(\phi_3)]} \tag{7.33}$$

8. Assorted Loose Ends

8.1 Degeneracy

So far in this paper it has been implicitly assumed that no two eigenmodes have the same frequency Ω and wavenumber Ω. This section considers the degenerate case in which this happens. An example is the Trapezoidal method. The dispersion relation is

$$\tan(\Omega/2) = \frac{r}{2}\sin(\phi) \tag{3.41}$$

which has wavenumber solutions ϕ_1 and $\phi_2 = \pi - \phi_1$. These are identical when

$$\phi_1' = \phi_2' = \pi/2 \tag{8.1}$$

which occurs at the degenerate frequency Ω' given by,

$$\tan(\Omega/2) = \frac{r}{2} \tag{8.2}$$

Note that $r_g' = \left(\frac{\partial\Omega}{\partial\phi}\right)_j'$ const

$$= r\cos(\phi')\cos^2(\Omega'/2)$$

$$= 0 \tag{8.3}$$

This is characteristic of degeneracy because in the neighbourhood of the double zero

$$\Omega - \Omega' = a(\phi - \phi')^2 \qquad , a = \text{constant} \tag{8.4}$$

so $$r_g = \frac{\partial\Omega}{\partial\phi}$$

$$= 2a(\phi - \phi')$$

$$= 0 \qquad \text{at } \phi = \phi' \tag{8.5}$$

For $\Omega \neq \Omega'$ the general solution for constant r is

$$U_j^n = A_1 \exp[i(j\phi_1 - n\Omega)] + A_2 \exp[i(j\phi_2 - n\Omega)] \qquad (8.6)$$

Consider the limit as Ω approaches Ω'

$$\phi - \phi' \approx \pm \left(\frac{\Omega - \Omega'}{a}\right)^{\frac{1}{2}} \qquad (8.7)$$

$$\text{so} \quad \phi_1 - \phi' \approx -(\phi_2 - \phi') \qquad (8.8)$$

Hence $A_1 \exp[ij\phi_1] + A_2 [ij\phi_2]$

$$= \exp[ij\phi'] \{ A_1 \exp[ij(\phi_1 - \phi')] + A_2 \exp[ij(\phi_2 - \phi')] \}$$

$$\approx \exp[ij\phi'] \{ A_1 \exp[ij(\phi_1 - \phi')] + A_2 \exp[ij(\phi' - \phi_1)] \}$$

$$\approx \exp[ij\phi'] \{ (A_1 + A_2)\cos[j(\phi_1 - \phi')] + i(A_1 - A_2)\sin[j(\phi_1 - \phi')] \}$$

$$\approx \exp[ij\phi'] \{ A_3 + A_4 j \} \qquad (8.9)$$

where $A_3 = A_1 + A_2$ (8.10)

and $A_4 = i(A_1 - A_2)(\phi_1 - \phi')$ (8.11)

If A_3 and A_4 are now considered to be constants in the limiting process the general form of the degenerate eigenfunction is,

$$U_j^n = (A_3 + A_4 j) \exp[i(j\phi' - n\Omega')] \qquad (8.12)$$

Another way of deriving this result is through the asymptotic amplitude equation. With the amplitude A time-independent, r constant, and r_g equal to zero the largest term is the second order derivative of A so,

$$\frac{\partial^2 A}{\partial j^2} = 0 \qquad (8.13)$$

which implies $A = A_3 + A_4 j$ (8.14)

In all the examples I have analysed I have not yet

found a degenerate eigenfunction which satisfies the asymptotic boundary conditions and so is a degenerate eigenmode. The degenerate frequency does however satisfy the determinant condition

$$\det(B) = 0 \tag{6.7}$$

because two of the wavenumbers are equal so their columns are identical and B is singular. This is not a problem in the stability analysis because the separation of eigenfrequencies, using the result from chapter 5 , is approximately,

$$\frac{2\pi}{N} = 2\pi \left(\int_0^J [r_g(\phi_1, j)]^{-1} - [r_g(\phi_2, j)]^{-1} \, dj \right)^{-1}$$

$$\to 0 \quad \text{as} \quad \Omega \to \Omega'$$

so there is usually a true eigenfrequency which differs from the degenerate frequency by less than the overall asymptotic error.

8.2 Eigenmodes and Eigenfrequencies

In chapters 5 and 6 it was stated that the eigenmodes of a linear system of finite difference equations with time-independent coefficients and boundary conditions vary exponentially with time. This section outlines the proof.

Let $U_j^n = z^n V_j$ (8.15)

If the domain is $0 < j < J$ then there must be J+1 finite difference equations so V_j satisfy

$$\tilde{\tilde{C}}\,\tilde{V} = 0 \tag{8.16}$$

where $\tilde{V}$ is the J+1 vector of V_j elements and $\tilde{\tilde{C}}$ is a $(J+1)^2$ matrix whose elements are polynomials of z.

For there to be a non-zero solution requires

$$\det(\tilde{\tilde{C}}) = 0 \tag{8.17}$$

This is the equation that determines the eigenvalues z of the eigenmodes. Apart from the problem of possible degeneracy the only remaining difficulty is to show that the number of eigenmodes equals the number of independent initial conditions needed to start a numerical solution. For the Backward Euler method with space extrapolation at the downstream boundary there are J-1 independent initial conditions since

$$U_0^0 = 0 \tag{8.18}$$

and $$U_J^0 = U_{J-1}^0 \tag{8.19}$$

$$\underset{\approx}{C} = \begin{pmatrix} 1 & & & & & & \\ -r_1 z/2 & z-1 & r_1 z/2 & & & & \\ & -r_2 z/2 & z-1 & r_2 z/2 & & & \\ & & -r_3 z/2 & z-1 & r_3 z/2 & & \\ & & & \cdots\cdots\cdots\cdots\cdots & & & \\ & & & & -r_{J-1} z/2 & z-1 & r_{J-1} z/2 \\ & & & & & 1 & -1 \end{pmatrix} \tag{8.20}$$

so $\det(\underset{\approx}{C})$ is a polynomial in z of degree J-1 giving J-1 eigenmodes. Thus any solution can be expressed as a sum of eigenmodes which vary exponentially with time.

8.3 Other Asymptotic Approaches

One approach which can be used when the variation in coefficients is extremely small is to let,

$$U_j^n = A(j,n) \exp[i(j\phi - n\Omega)] \tag{8.21}$$

where Ω and ϕ are both constant and satisfy

$$a_0(\Omega,\phi,j) = 0 \tag{8.22}$$

at some point j.

The asymptotic amplitude equation is then

$$a_0 A + a_1 \frac{\partial A}{\partial n} + a_2 \frac{\partial A}{\partial j} = 0 \tag{8.23}$$

where a_0, a_1 and a_2 are defined and calculated as before.

$$L_A = \min\left(A \; / \; \frac{\partial A}{\partial j} \right)$$

$$= O\,(a_2 \; / \; a_0) \tag{8.24}$$

so the fractional error using this method is $O\,(a_0/a_2)^2$. If the variations in the coefficients are small this is fine but if the variations are O(1) the fractional error is O(1) i.e. the method fails to give accurate asymptotic approximations. The asymptotic approach used in this paper allows total variations in the coefficients of O(1) and only requires that the length scale of those variations is much greater than 1.

Another approach is to set

$$U_j^n = A(j,n) \exp[i\Psi(j,n)] \tag{8.25}$$

with $$\frac{\partial \Psi}{\partial j} = \phi \tag{8.26}$$

and $\frac{\partial \Psi}{\partial n} = - \Omega$ (8.27)

and A, Ω, ϕ are all real and slowly varying. This approach is used extensively in the analysis of water waves and other partial differential equations with dispersion and very little dissipation. This approach applied to finite difference equations would give poor results because dissipation over one time step is O(1) so if Ω is real A reflects this dissipation and so T_A = O(1). The method presented in this paper is able to use constant complex Ω rather than variable real Ω as in the above method because the eigenmodes have constant complex eigenfrequencies provided the finite difference equations are time-independent.

9. Wavepacket Test Program

9.1 Program Description

The program solves the model convective problem,

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0 \tag{9.1}$$

using a choice of Box or Trapezoidal methods on the domain,

$$0 < j < 200 \tag{9.2}$$

and time step range,

$$0 < n < 400 \tag{9.3}$$

The CFL number

$$r_j = \frac{c_j \Delta t}{\Delta x_j} \tag{9.4}$$

is specified by the user at j=0,200 and the program interpolates for intermediate values by fitting an exponential curve through r_0 and r_{200}.

$$r(j) = r_0 \exp\left(\frac{j}{200} \ln(r_{200}/r_0)\right) \tag{9.5}$$

Methods 1-3 are different Trapezoidal methods which are identical if r is constant. For these methods the program offers a choice of space extrapolation, space-time extrapolation or box condition as the downstream boundary condition. Method 4 is the Box method. The wavepacket theory for each of these methods is derived in the next section.

The upstream boundary condition is,

$$U_0^n = 0 \tag{9.6}$$

The initial conditions are given by,

$$U_j^n = A(j,n)\ \exp[i\Psi(j,n)] \tag{9.7}$$

where $$A(j,n) = \exp[-(j-100)^2/200] \tag{9.8}$$

and $$\Psi(j,n) = \sum_{k=0}^{j} \phi_1(\Omega,j) \tag{9.9}$$

This produces a wavepacket about 40 mesh points long. The user specifies Ω by specifying the critical CFL number which is defined as,

$$r_{crit} = 2\ \tan(\Omega/2) \tag{9.10}$$

so $$\Omega = 2\ \tan^{-1}(r_{crit}/2) \tag{9.11}$$

The significance of the critical CFL number is that when,

$$r = r_{crit} \tag{9.12}$$

the group CFL number for the Trapezoidal methods is zero. Wave propagation can only occur for $r > r_{crit}$.

At each time step the program calculates 'experimental' values for the position X(n) and energy E(n) of the wavepacket. If U was continuous X and E would be defined as,

$$E(n) = \int_{x_0}^{x_{200}} |U(x,t_n)|^2\ dx$$

$$= \int_{0}^{200} |U(j,n)|^2\ \frac{dx}{dj}\ dj \tag{9.13}$$

$$\text{and} \quad X(n) = \frac{1}{E(n)} \int_{x_0}^{x_{200}} |U(x,t_n)|^2 \, x \, dx$$

$$= \frac{1}{E(n)} \int_{0}^{200} |U(j,n)|^2 \, x(j) \, \frac{dx}{dj} \, dj \qquad (9.14)$$

so since U is discrete X(n) and E(n) are defined as,

$$E(n) = \sum_{j=0}^{200} |U_j^n|^2 \, \alpha_j \qquad (9.15)$$

$$X(n) = \frac{1}{E(n)} \sum_{j=0}^{200} |U_j^n|^2 \, x_j \, \alpha_j \qquad (9.16)$$

$$\text{where} \quad \alpha_j = \frac{1}{2}(x_{j+1} - x_{j-1}) \qquad (9.17)$$

In the program c is taken to be constant so variations in r are due to variations in Δx.

$$\Delta x = \frac{c\Delta t}{r} \qquad (9.18)$$

$$\text{so} \quad x_{j+1} - x_{j-1} \propto \left(r_{j+\frac{1}{2}}\right)^{-1}$$

A physical domain $0<x<1$ is used so,

$$x_j = \sum_{k=1}^{j} \left(r_{j-\frac{1}{2}}\right)^{-1} \Bigg/ \sum_{k=1}^{200} \left(r_{j-\frac{1}{2}}\right)^{-1} \qquad (9.19)$$

After completing the 400 time steps the program calculates predicted values for X(n) and E(n) using the wavepacket equations derived in the next section with the experimental values at n=1 as initial conditions.

9.2 Theory

Method 1 is a Trapezoidal method which is second order accurate in Δx only when Δx is constant.

$$\left[\delta_t + \frac{r_j}{2} \mu_t \delta_{2x} \right] U_j^{n+\frac{1}{2}} = 0 \qquad (9.20)$$

$$r_j = \frac{2c\Delta t}{x_{j+1} - x_{j-1}} \qquad (9.21)$$

Following the analyses of chapters 3 and 4,

$$a_0 = -2i \sin(\Omega/2) + ir \cos(\Omega/2) \sin(\phi) \qquad (9.22)$$

$$a_1 = \cos(\Omega/2) + \frac{r}{2} \sin(\Omega/2) \sin(\phi) \qquad (9.23)$$

$$a_2 = r \cos(\Omega/2) \cos(\phi) \qquad (9.24)$$

$$a_3 = -\frac{r}{2} \cos(\Omega/2) \sin(\phi) \qquad (9.25)$$

The dispersion relation is

$$\tan(\Omega/2) = \frac{r}{2} \sin(\phi) \qquad (9.26)$$

so $a_0 = 0$ (9.27)

and $a_1 = \cos(\Omega/2) + \sin(\Omega/2) \tan(\Omega/2)$

$$= (\cos^2(\Omega/2) + \sin^2(\Omega/2)) / \cos(\Omega/2)$$

$$= \sec(\Omega/2) \qquad (9.28)$$

$$r_g = a_2 / a_1 \qquad (3.26)$$

$$= r \cos(\phi) \cos^2(\Omega/2) \qquad (9.29)$$

$$\frac{d\phi}{dn} = i \left(\frac{\partial a_0}{\partial j} \right)_{\Omega, \phi \text{ const}} / a_1 \qquad (4.4)$$

$$= - \frac{\partial r}{\partial j} \sin(\phi) \cos^2(\Omega/2) \qquad (9.30)$$

$$\varepsilon = -ia_3 \left(\frac{\partial a_0}{\partial j}\right)_{\Omega,\phi \text{ const}} / a_1 a_2 \quad - a_0 / a_1 \qquad (3.30)$$

$$= - \frac{1}{2} \frac{\partial r}{\partial j} \cos^2(\Omega/2) \sin(\phi) \tan(\phi) \qquad (9.31)$$

Since α is proportional to r^{-1},

$$\frac{1}{\alpha} \frac{\partial \alpha}{\partial j} = - \frac{1}{r} \frac{\partial r}{\partial j} \qquad (9.32)$$

so $$\frac{1}{\alpha} \frac{\partial}{\partial j} (r_g \alpha) = r \frac{\partial}{\partial j} (r_g / r)$$

$$= - r \sin(\phi) \cos^2(\Omega/2) \frac{\partial \phi}{\partial j}$$

$$= - r \sin(\phi) \cos^2(\Omega/2) \; i \frac{\partial a_0}{\partial j} \quad / a_2$$

$$= \frac{\partial r}{\partial j} \cos^2(\Omega/2) \sin(\phi) \tan(\phi) \qquad (9.33)$$

Hence the equations of motion for the wavepacket for real Ω and ϕ are,

$$\frac{dj}{dn} = r \cos(\phi) \cos^2(\Omega/2) \qquad (9.34)$$

$$\frac{d\phi}{dn} = - \frac{\partial r}{\partial j} \sin(\phi) \cos^2(\Omega/2) \qquad (9.30)$$

$$\frac{d}{dn} \ln(E) = 0 \qquad (9.35)$$

Method 2 is also a Trapezoidal method which is second order accurate in Δx only when Δx is constant.

$$\left[\delta_t + \frac{r_{j+\frac{1}{2}}}{2} \mu_t \Delta_x + \frac{r_{j-\frac{1}{2}}}{2} \mu_t \nabla_x \right] U_j^{n+\frac{1}{2}} = 0 \qquad (9.36)$$

$$r_{j+\frac{1}{2}} = \frac{c\Delta t}{x_{j+1}-x_j} \tag{9.37}$$

$$a_0 = -2i \sin(\Omega/2) + \frac{1}{2}\cos(\Omega/2)\left(r_{j+\frac{1}{2}}[\exp(i\phi)-1] + r_{j-\frac{1}{2}}[1-\exp(-i\phi)] \right)$$

$$= -2i \sin(\Omega/2) + \frac{1}{2}\cos(\Omega/2)\left((r_j + \frac{1}{2}\frac{\partial r}{\partial j})[\exp(i\phi)-1] + (r_j - \frac{1}{2}\frac{\partial r}{\partial j})[1-\exp(-i\phi)] \right) + \text{H.O.T.}$$

$$= -2i \sin(\Omega/2) + ir\cos(\Omega/2)\sin(\phi) + \frac{1}{2}\frac{\partial r}{\partial j}\cos(\Omega/2)[\cos(\phi)-1] + \text{H.O.T.} \tag{9.38}$$

The H.O.T. are neglected because they are of the same order of magnitude as other terms already neglected in the derivation of the asymptotic amplitude equation. To this same level of asymptotic accuracy a_1, a_2 and a_3 are exactly the same as in the analysis of method 1.

The dispersion relation remains,

$$\tan(\Omega/2) = \frac{r}{2}\sin(\phi) \tag{9.26}$$

so now $$a_0 = \frac{1}{2}\frac{\partial r}{\partial j}\cos(\Omega/2)\,[\cos(\phi)-1] \tag{9.39}$$

After some algebra it follows that the equations of motion are,

$$\frac{dj}{dn} = r\cos(\phi)\cos^2(\Omega/2) \tag{9.40}$$

$$\frac{d\phi}{dn} = -\frac{1}{2}\frac{\partial r}{\partial j}\sin(\phi)\cos^2(\Omega/2) \tag{9.41}$$

$$\frac{d}{dn}\ln(E) = \frac{\partial r}{\partial j}[1-\cos(\phi)]\cos^2(\Omega/2) \tag{9.42}$$

Method 3 is a Trapezoidal method which is second

order accurate in Δx for cases with smoothly varying Δx.

$$\left[\delta_t + \frac{\left(r_{j+\frac{1}{2}}\right)^2}{r_{j+\frac{1}{2}}+r_{j-\frac{1}{2}}} \mu_t \Delta_x + \frac{\left(r_{j-\frac{1}{2}}\right)^2}{r_{j+\frac{1}{2}}+r_{j-\frac{1}{2}}} \mu_t \nabla_x \right] U_j^{n+\frac{1}{2}} = 0 \qquad (9.43)$$

The dispersion relation is the same as for methods 1 and 2 so after some algebra,

$$a_0 = \frac{\partial j}{\partial j} \cos(\Omega/2) \, [\cos(\phi)-1] \qquad (9.44)$$

and the equations of motion are,

$$\frac{dj}{dn} = r \cos(\phi) \cos^2(\Omega/2) \qquad (9.45)$$

$$\frac{d\phi}{dn} = - \frac{\partial r}{\partial j} \sin(\phi) \cos^2(\Omega/2) \qquad (9.46)$$

$$\frac{d}{dn} \ln(E) = 2 \frac{\partial r}{\partial j} [1-\cos(\phi)] \cos^2(\Omega/2) \qquad (9.47)$$

When the wavepackets of methods 1-3 reach the downstream boudary at j=200 they are reflected into backward travelling wavepackets. The energy E_2 and the wavenumber ϕ_2 of the reflected wavepacket are related to the energy E_1 and wavenumber ϕ_1 of the incident wavepacket by the equations,

$$\phi_2 = \pi - \phi_1 \qquad (9.48)$$

$$\ln(E_2) = \ln(E_1) + 2 \ln|R_J| \qquad (9.49)$$

where R_J is the amplitude reflection coefficient.

For space extrapolation (see §3.5.2)

$$|R_J| = \frac{\sin(\phi_1/2)}{\sin(\phi_2/2)}$$

$$= \tan(\phi_1/2) \qquad (9.50)$$

For space-time extrapolation (see §3.5.3)

$$|R_J| = \left| \frac{\sin[(\phi_1 - \Omega)/2]}{\sin[(\phi_2 - \Omega)/2]} \right| . \qquad (9.51)$$

For the box boundary condition (see §3.5.4)

$$|R_J| = \left| \frac{\cos(\phi_1/2)\ \tan(\Omega/2) - r\ \sin(\phi_1/2)}{\cos(\phi_2/2)\ \tan(\Omega/2) - r\ \sin(\phi_2/2)} \right|$$

$$= \tan^3(\phi_1/2) \qquad (9.52)$$

after substituting for $\tan(\Omega/2)$ using the dispersion relation and replacing ϕ_2 by $\pi - \phi_1$.

The reflection relations at the upstream boundary are,

$$\phi_1 = \pi - \phi_2 \qquad (9.53)$$

$$\ln(E_1) = \ln(E_2) \qquad (9.54)$$

since $R_0 = -1$ (see §3.5.1)

Method 4 is the Box method discussed in §3.3.2

$$[\ \mu_x \delta_t + r_j\ \mu_t \delta_x\ ]\ U_{j+\frac{1}{2}}^{n+\frac{1}{2}} = 0 \qquad (9.55)$$

The dispersion relation is

$$\tan(\Omega/2) = r\ \tan(\phi/2) \qquad (9.56)$$

and the equations of motion for the wavepacket are,

$$\frac{dj}{dn} = r\ [\ 1 + \tan^2(\phi/2)\ ]\ \cos^2(\Omega/2) \qquad (9.57)$$

$$\frac{d\phi}{dn} = -2\ \frac{\partial r}{\partial j}\ \tan(\phi/2)\ \cos^2(\Omega/2) \qquad (9.58)$$

$$\frac{d}{dn} \ln(E) = -2 \frac{\partial r}{\partial j} \tan^2(\phi/2) \cos^2(\Omega/2) \qquad (9.59)$$

9.3 Numerical Results

9.3.1 Trapezoidal Method with Variable CFL Number

This example uses,

Method type = 2 ; one of the Trapezoidal methods
Boundary type = 1 ; space extrapolation

$r_0 = 0.05$ $r_{200} = 0.2$ $r_{crit} = 0.04$

Figure 6 shows X(n) and ln[E(n)] both predicted and experimental. This example shows the movement of a wavepacket and the change in its energy due to the variation in the CFL number. The agreement between the predicted and experimental values is excellent. The energy of the analytic solution is constant so the wavepacket theory has successfully predicted almost all of change in the numerical energy due to variable Δx.

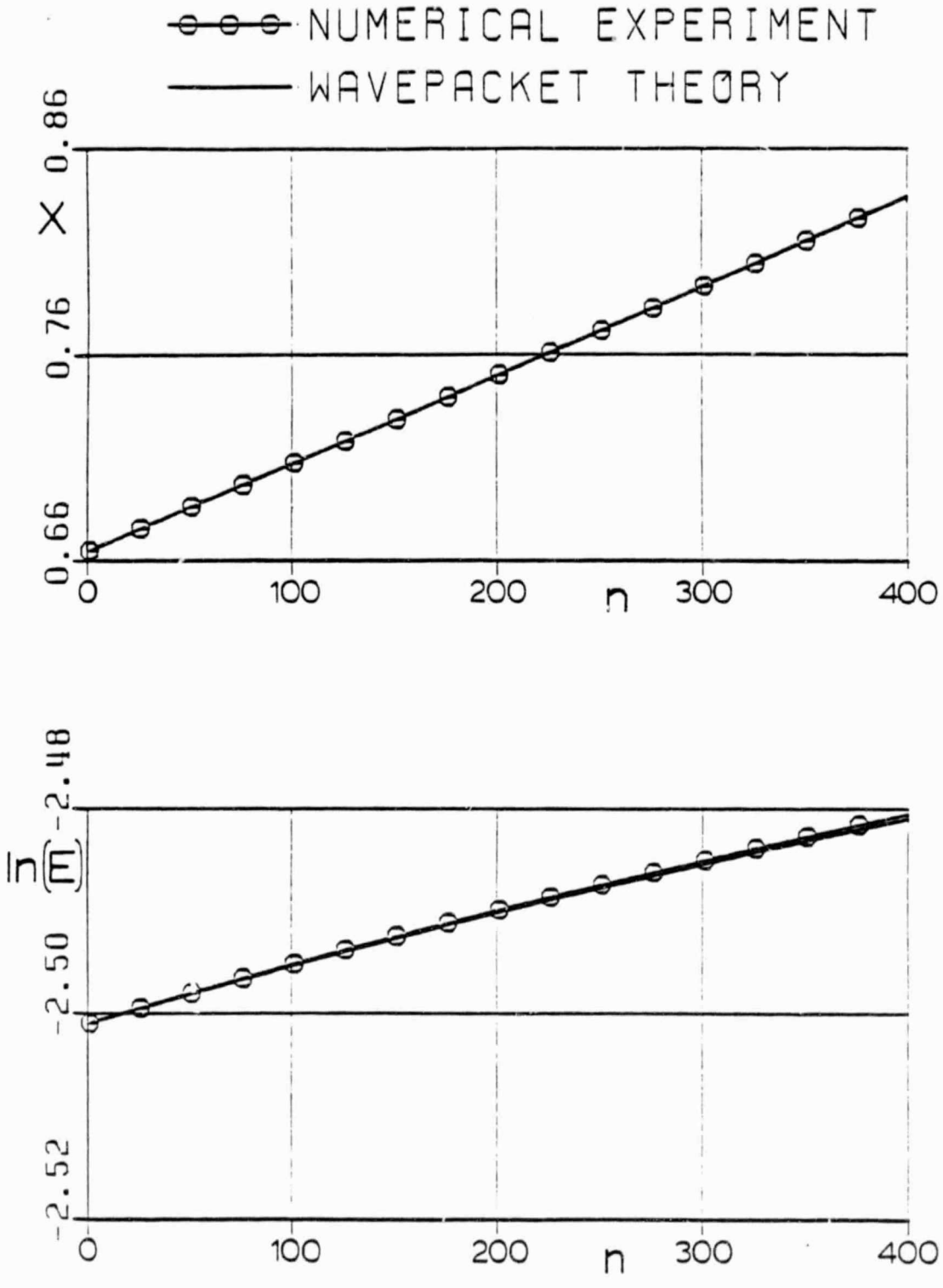


FIGURE 6. EXAMPLE 9.3.1 TRAPEZOIDAL METHOD 2

9.3.2 Box Method with Variable CFL Number

This example uses,

Method type = 4 ; Box method

$r_0 = 0.05$ $r_{200} = 0.2$ $r_{crit} = 0.04$

Figure 7 shows X(n) and ln[E(n)]. As in §9.3.1 the agreement between experiment and theory is excellent.

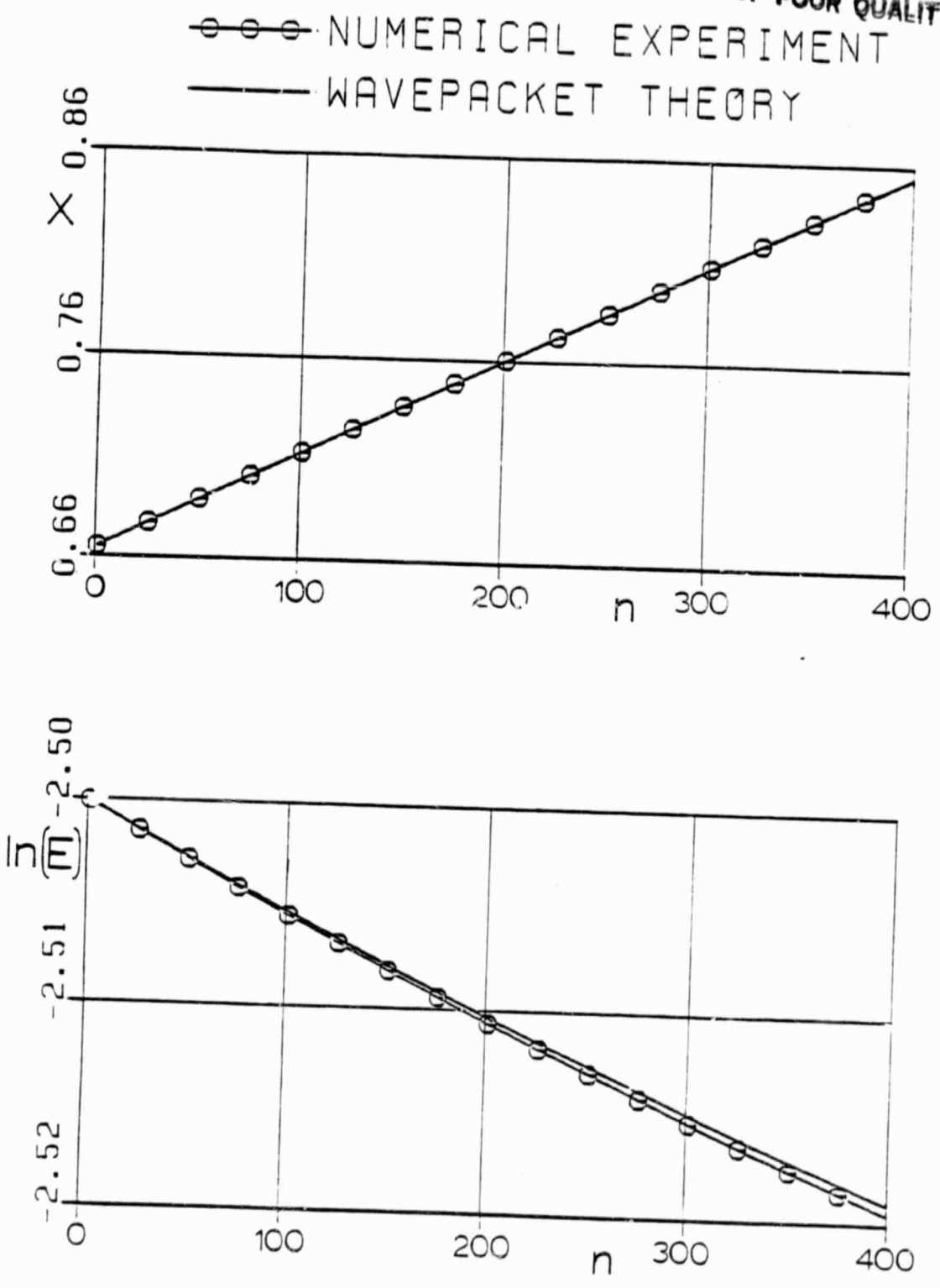


FIGURE 7. EXAMPLE 9.3.2 BOX METHOD

9.3.3 Trapezoidal Method with Space Extrapolation

This example uses,

Method type = 1 ; one of the Trapezoidal methods
Boundary type = 1 ; space extrapolation

$r_0 = 1.0$ $r_{200} = 1.0$ $r_{crit} = 0.3$

Figure 8 shows X(n) and ln[E(n)]. This example illustrates the effect of the downstream boundary reflecting a wavepacket with reduced energy. Because of the finite length of the wavepacket the drop in energy is smeared and X(n) does not quite reach 1.0 .

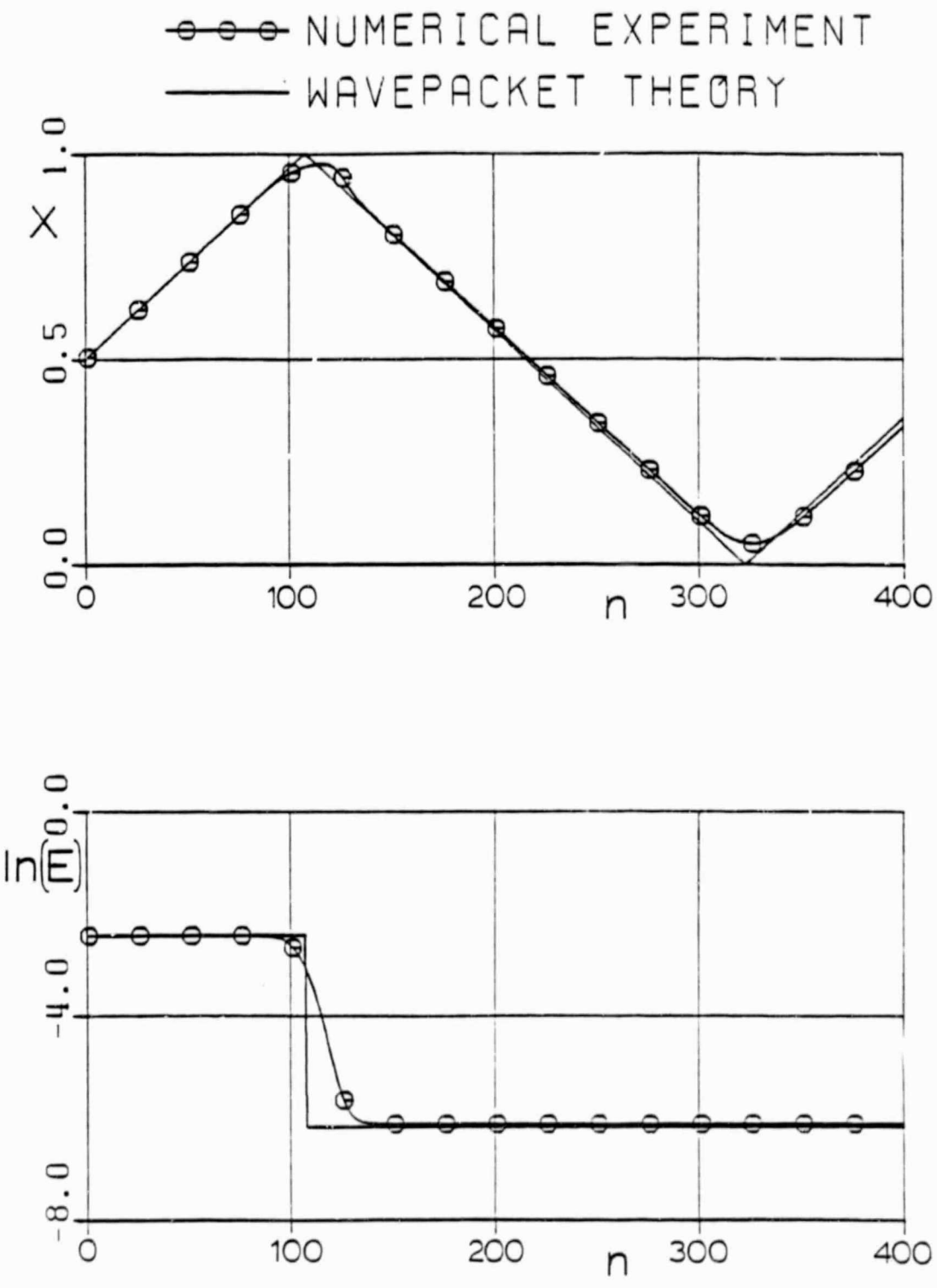


FIGURE 8. EXAMPLE 9.3.3 TRAPEZOIDAL METHOD 1 WITH SPACE EXTRAPOLATION

9.3.4 Trapezoidal Method with Space-time Extrapolation

This example is the same as §9.3.3 except that

Boundary type = 2 ; space-time extrapolation

Figure 9 shows X(n) and ln[E(n)]. The energy of the reflected wavepacket is less than in §9.3.3. As a consequence the first order terms which are neglected in the asymptotic boundary conditions are more significant and so the energy decrease is more smeared and there is a larger discrepancy between the experimental and predicted values.

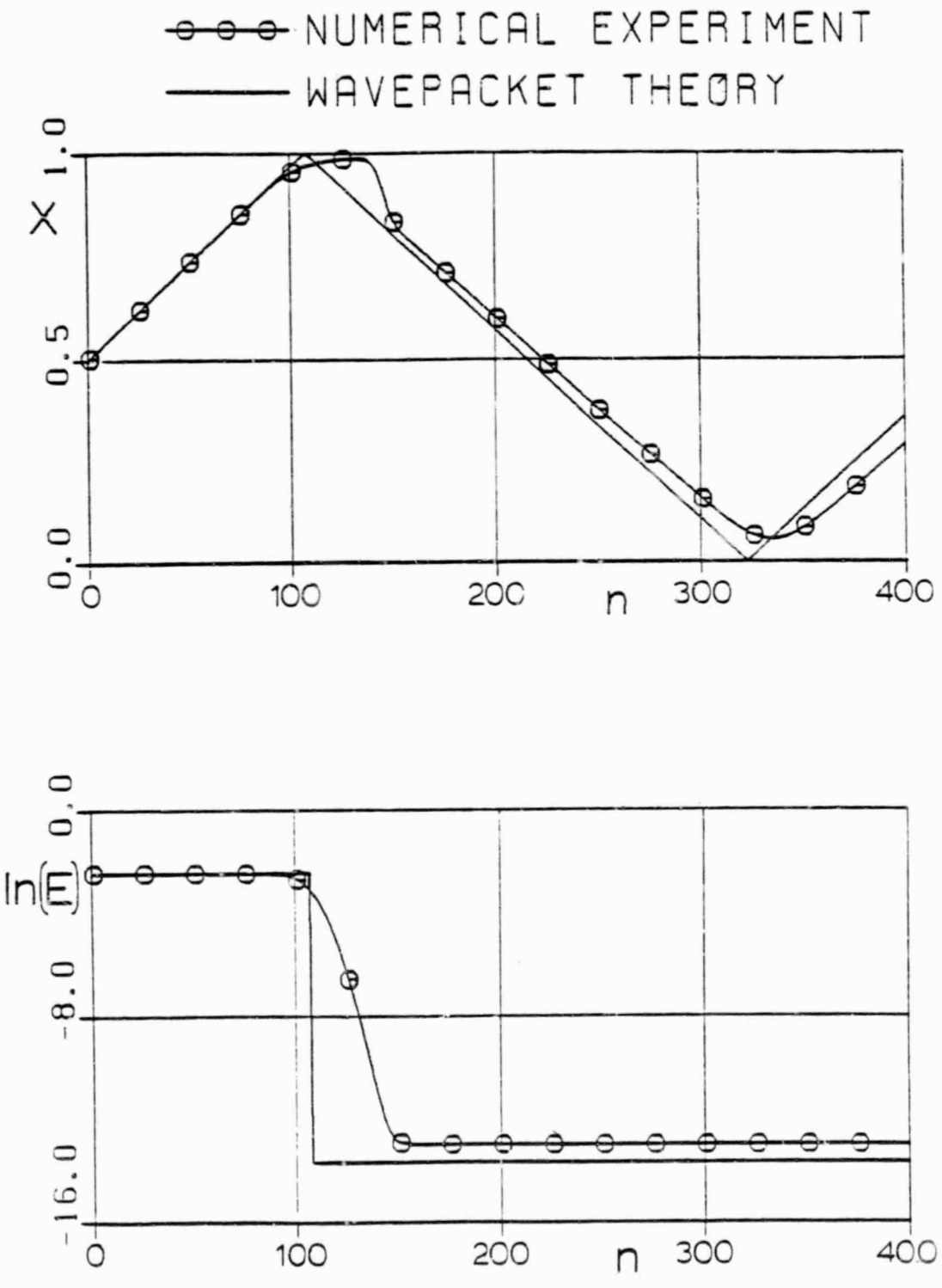


FIGURE 9. EXAMPLE 9.3.4 TRAPEZOIDAL METHOD 1 WITH SPACE-TIME EXTRAPOLATION

9.3.5 Trapezoidal Method with Box Boundary Condition

This example is the same as §9.3.3 except that,

Boundary type = 3 ; box boundary condition

Figure 10 shows X(n) and ln[E(n)]. The energy drop in this example is three times that in §9.3.3 because,

$$\ln(E_2) = \ln(E_1) + 2\ln(|R_J|) \tag{9.49}$$

and

$$|R_J| = \begin{cases} \tan(\phi_1/2) & \text{space extrapolation} \quad (9.50) \\ \tan^3(\phi_1/2) & \text{space-time extrapolation} \quad (9.52) \end{cases}$$

Thus the box boundary condition increases the overall convergence rate by factor 3 with minimal extra computational effort.

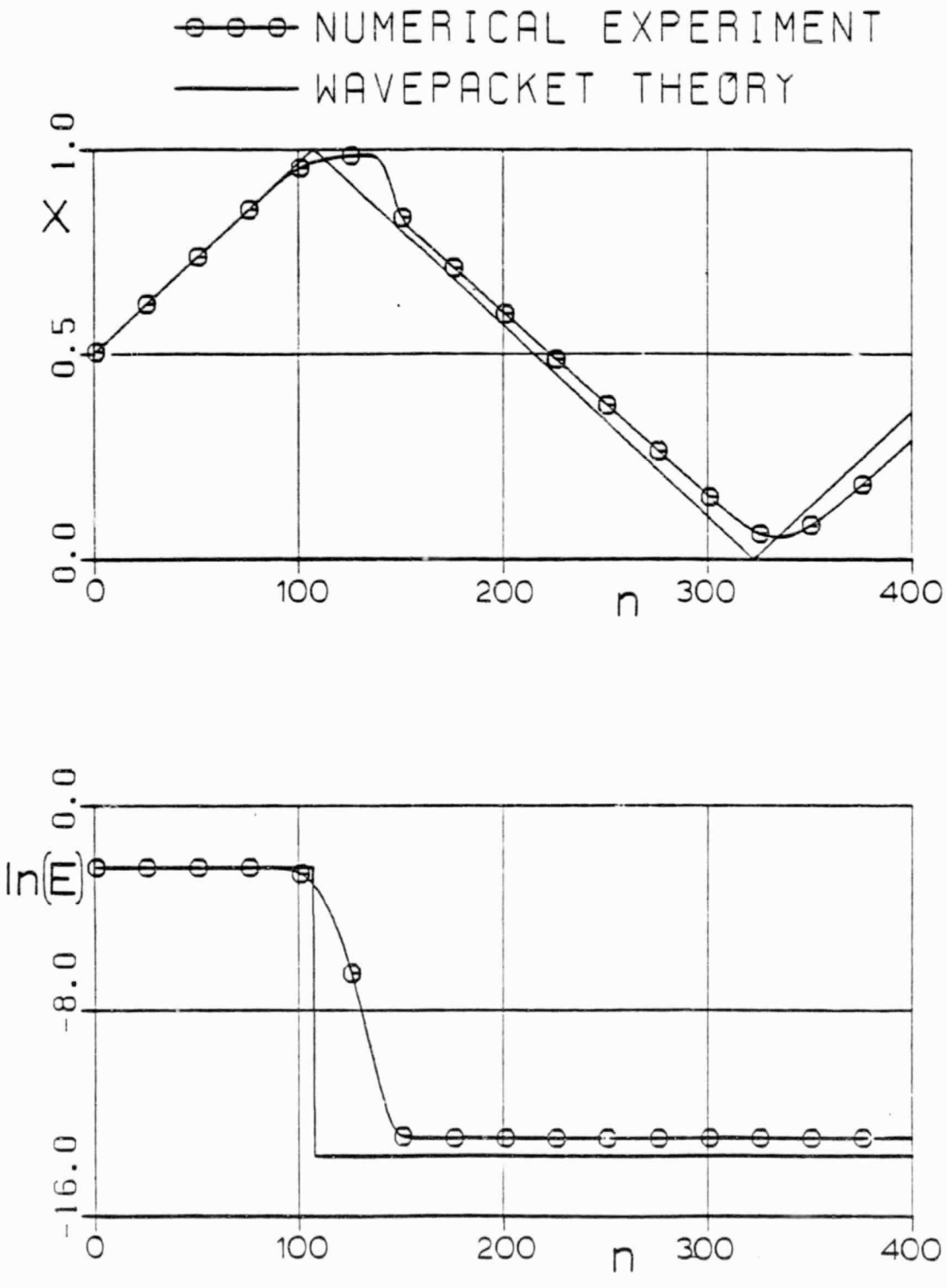


FIGURE 10. EXAMPLE 9.3.5 TRAPEZOIDAL METHOD 1 WITH BOX BOUNDARY CONDITION

9.3.6 Wavepacket Outflow in Box Method

This example uses,

Method type = 4 ; Box method

$r_0 = 0.4$ $r_{200} = 0.4$ $r_{crit} = 0.2$

Figure 11 shows X(n) and ln[E(n)]. Note that when the wavepacket reaches the downstream boundary the experimental value for X(n) remains near 1.0 and ln(E) decreases rapidly towards $-\infty$.

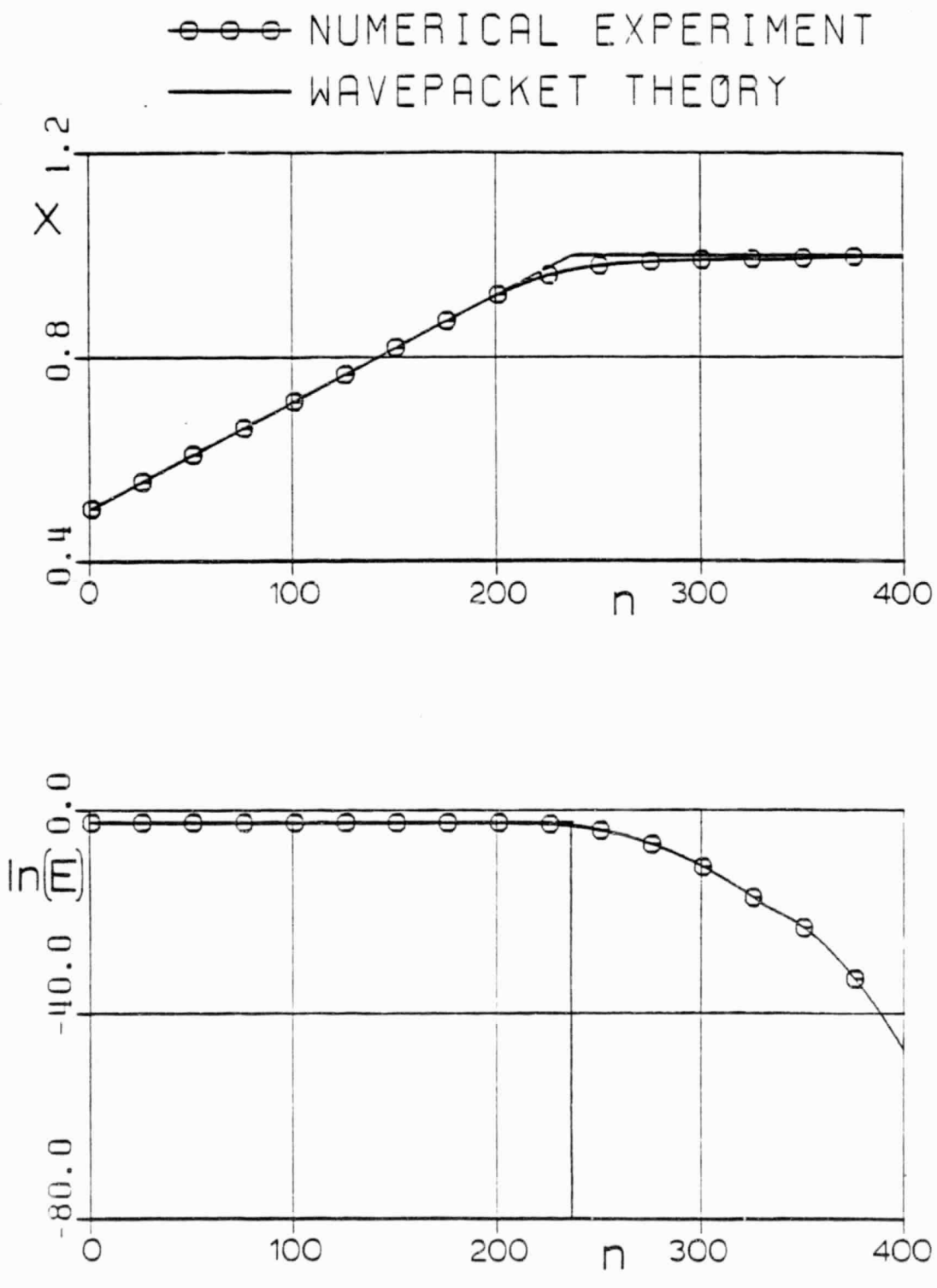


FIGURE 11. EXAMPLE 9.3.6 BOX METHOD

9.3.7 Instability of Trapezoidal Method with Space-time Extrapolation

This example uses,

Method type = 1 ; one of the Trapezoidal methods
Boundary type = 2 ; space-time extrapolation

$r_0 = 3.0$ $r_{200} = 3.0$ $r_{crit} = 2.1$

Figure 12 shows X(n) and ln[E(n)]. Both the theoretical prediction and the numerical result show that the energy increases every time the wavepacket reflects off the downstream boundary and so the numerical scheme is unstable. The CFL stability condition for the case with constant CFL number is obtained by considering the amplitude reflection coefficient.

$$|R_J| = \left| \frac{\sin\left(\frac{\phi_1 - \Omega}{2}\right)}{\sin\left(\frac{\phi_2 - \Omega}{2}\right)} \right| \qquad (9.52)$$

The dispersion relation is

$$\tan(\Omega/2) = \frac{r}{2} \sin(\phi) \qquad (9.26)$$

Now consider the two cases r<2 and r>2 .

a) $r < 2$

$0 < \phi_1 < \pi/2 \quad ==> \quad \tan(\Omega/2) < 1$

$==> \quad \Omega < \pi/2$

$==> \quad (\pi/2 - \Omega) > 0 \quad \text{and} \quad (\pi/2 - \phi_1) > 0$

$==> \quad \pi > |\, (\pi/2-\Omega) + (\pi/2-\phi_1)\, | > |\, (\pi/2-\Omega) - (\pi/2-\phi_1)\, | > 0$

$$==> \quad \left| \sin\left(\frac{\phi_2 - \Omega}{2}\right) \right| > \left| \sin\left(\frac{\phi_1 - \Omega}{2}\right) \right|$$

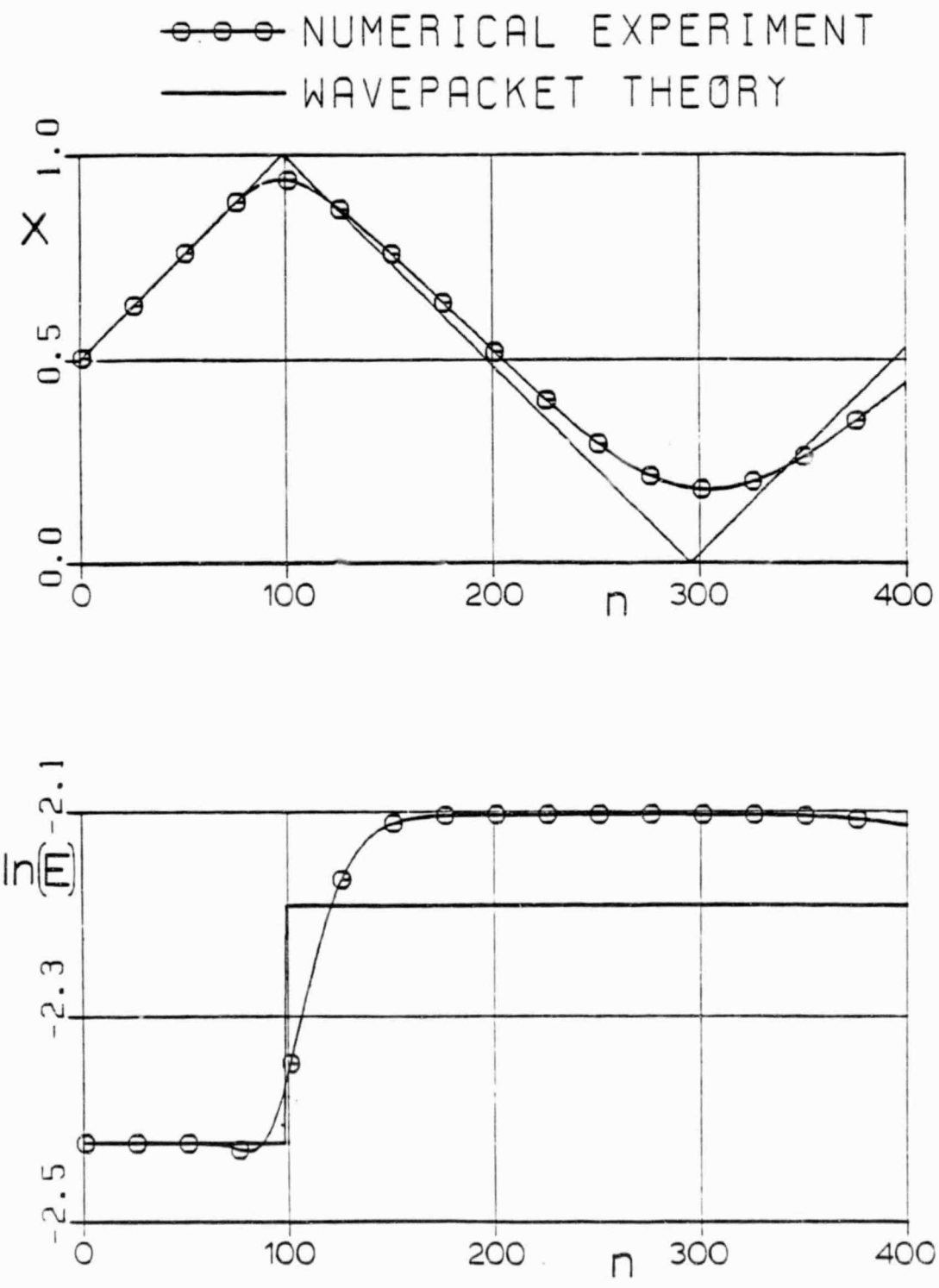


FIGURE 12. EXAMPLE 9.3.7 TRAPEZOIDAL METHOD 1 WITH SPACE-TIME EXTRAPOLATION

$$\Longrightarrow \quad |R_J| < 1$$

b) $\underline{r > 2}$

For $\phi=\pi/2$, $\tan(\Omega/2) > 1 \quad \Longrightarrow \quad \Omega > \pi/2 \quad \Longrightarrow \quad \Omega > \phi$

For $\phi=\pi$, $\tan(\Omega/2) = 0 \quad \Longrightarrow \quad \Omega < \phi$

As ϕ_2 varies from $\pi/2$ to π, $\Omega-\phi_2$ varies continuously so at some intermediate value,

$$\Omega = \phi_2$$

and hence $|R_J| = \infty$

The numerical scheme is thus stable if, and only if, the CFL number is less than 2. This stability condition has previously been derived by Beam, Warming and Yee [5].

In a numerical experiment an infinite amplitude reflection coefficient does not occur because the first order derivatives of the amplitude which are neglected in the asymptotic boundary conditions become significant. In fact in all the unstable cases I have tried the agreement between experiment and prediction is poor because of the neglected first order terms in the boundary condition and the neglected second order terms in the amplitude equation. The example given is one of the best. The qualitative effects of these neglected terms can be understood as follows;

Since $\frac{\partial A}{\partial n} = r_g \frac{\partial A}{\partial j}$ all the second order terms in the amplitude equation $\frac{\partial^2 A}{\partial n^2}$, $\frac{\partial^2 A}{\partial n \partial j}$, $\frac{\partial^2 A}{\partial j^2}$ can be expressed in terms of $\frac{\partial^2 A}{\partial j^2}$. Thus the amplitude equation including second order terms has the form,

$$\frac{\partial A}{\partial n} + r_g \frac{\partial A}{\partial j} = \nu \frac{\partial^2 A}{\partial j^2} \qquad (9.61)$$

where ν is a function of Ω,ϕ and corresponds to an artificial viscosity. The effect of this artificial viscosity is to smear the wavepacket increasing its length and decreasing its maximum amplitude. This has the largest effect on X(n) since the longer the wavepacket the further X(n) must be from the ends of the domain. X(n) still oscillates approximately in phase with the predictions but the amplitude of the oscillations decreases steadily. The effect on the energy is much smaller.

The downstream boundary condition including first order terms can be written as,

$$A_2(j,n) + \tau_2 \frac{\partial A_2}{\partial n}(J,n) = R_J \left[A_1(J,n) + \tau_1 \frac{\partial A_1}{\partial n}(J,n)\right] \qquad (9.62)$$

where τ_1,τ_2 are functions of Ω,ϕ. If τ_1 and τ_2 are both small compared to T_A then (9.62) is approximately equal to

$$A_2(J,n+\tau_2) = R_J\, A_1(j,n+\tau_1) \qquad (9.63)$$

Thus the amplitude is reflected with a delay of $\tau_2-\tau_1$. This explains the fact that in almost all the examples in this chapter the reflected wavepacket lags behind the position predicted by wavepacket theory.

9.3.8 Instability of Trapezoidal Method Due to Varying Mesh

This example uses,

Method type = 3 ; one of the Trapezoidal methods
Boundary type = 1 ; space extrapolation

$r_0 = 1.0 \qquad r_{200} = 10.0 \qquad r_{crit} = 0.8$

Figure 13 shows X(n) and ln[E(n)]. The agreement between experiment and prediction is good for the energy but as in §9.3.7 the agreement is poor for X(n) because of the effect of the second order terms which are neglected in the asymptotic amplitude equation. The significance of this example is that this numerical scheme is stable for uniform meshes which give constant CFL number r but if the mesh, and hence r, varies sufficiently as in this example the scheme becomes unstable. This instability is best understood by expanding the finite difference equation in computational space.

$$\left[\delta_t + \frac{\left(r_{j+\frac{1}{2}}\right)^2}{r_{j+\frac{1}{2}}+r_{j-\frac{1}{2}}} \mu_t \Delta_x + \frac{\left(r_{j-\frac{1}{2}}\right)^2}{r_{j+\frac{1}{2}}+r_{j-\frac{1}{2}}} \mu_t \nabla_x \right] U_j^{n+\frac{1}{2}} = 0 \qquad (9.43)$$

$$\text{so } \left[\delta_t + \frac{r_j}{2} \mu_t \delta_{2x} + \frac{\partial r}{\partial j} \mu_t \delta_x^2 + \text{H.O.T.} \right] U_j^{n+\frac{1}{2}} = 0 \qquad (9.64)$$

The term $\frac{\partial r}{\partial j} \mu_t \delta_x^2$ corresponds to a viscous term in computational space. If $\frac{\partial r}{\partial j}$ is positive it corresponds to negative artificial viscosity and so causes the instability in the above example.

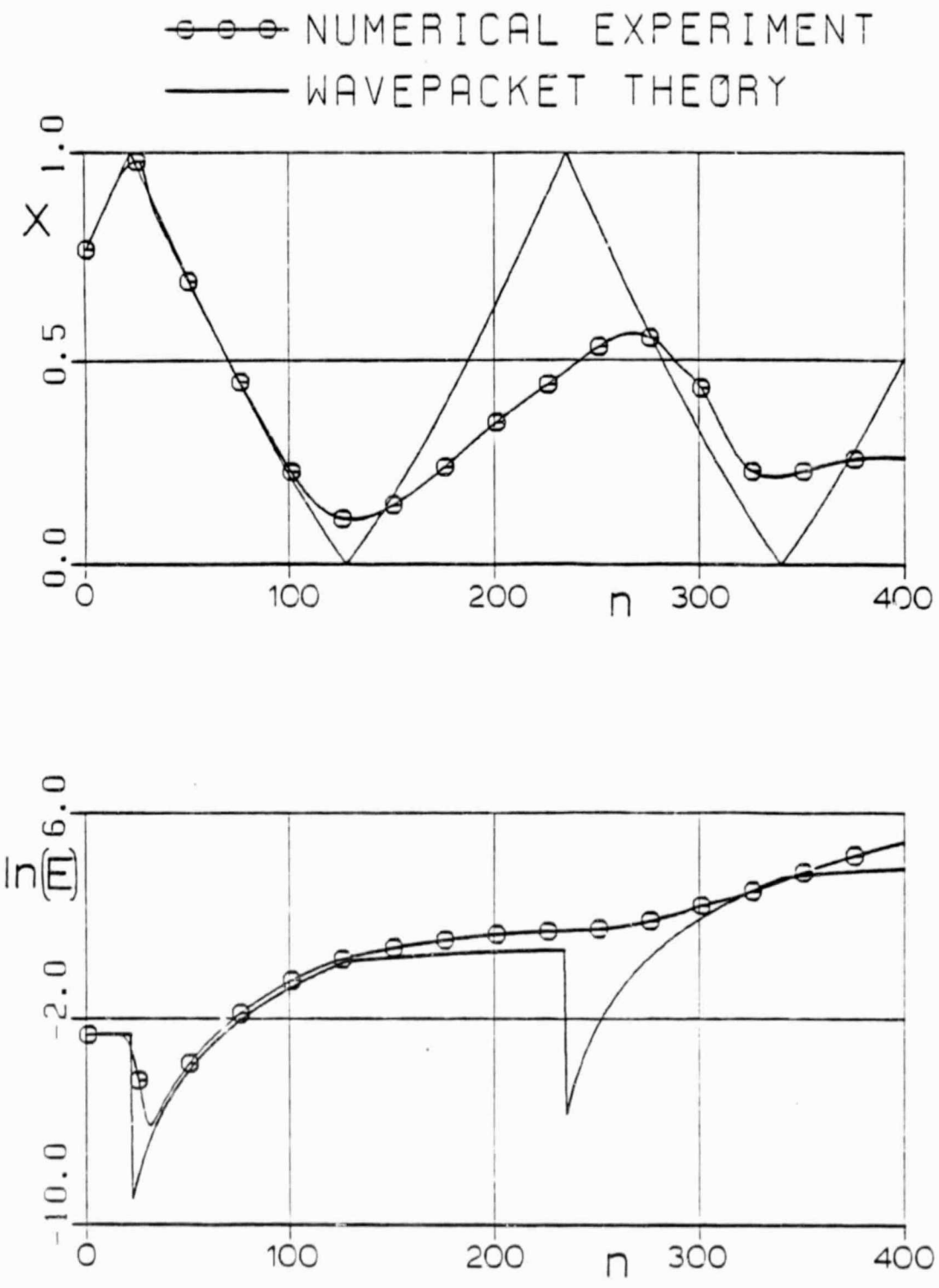


FIGURE 13. EXAMPLE 9.3.8 TRAPEZOIDAL METHOD 3 WITH SPACE EXTRAPOLATION

9.4 Program Listing

The program is written in FORTRAN 4 PLUS to be run on a PDP 11-70 with Versatec graphics subroutines.

```
C ***** PROGRAM WAVE

      DIMENSION U(0:202),V(0:202),A(0:200,3),W(0:200),XJ(0:202)
      DIMENSION EX(402),EE(402),TX(402),TE(402),T(402)

      EQUIVALENCE (TX(1),U(0)),(TX(204),V(0)),(TE(1),A(0,1)),
     1                  (T(1),A(0,3),(T(202),W(0))

      EXTERNAL R,J
      REAL K,J
      COMMON /RCONST/R0,R1,MT
      COMMON /X/X(0:200)

C ***** Input parameters

      TYPE *,'INPUT TERMINAL TYPE'
      TYPE *,'0 VERSATEC'
      TYPE *,'3 VT100 WITH GRAPHICS'
      TYPE *,'4 CHROMATICS'
      TYPE *,'5 VISUAL 500'
      ACCEPT *,NT

      TYPE *,' '
      TYPE *,'INPUT METHOD TYPE'
      TYPE *,'1-3 DIFFERENT TRAPEZOIDAL METHODS - SEE NOTES'
      TYPE *,'4 BOX METHOD'
      ACCEPT *,MT

      IF (MT.EQ.4) GOTO 1
      TYPE *,' '
      TYPE *,'INPUT DOWNSTREAM BOUNDARY TYPE'
      TYPE *,'1 SPACE EXTRAPOLATION'
      TYPE *,'2 SPACE-TIME EXTRAPOLATION'
      TYPE *,'3 BOX METHOD'
      ACCEPT *,MDB

    1 TYPE *,' '
      TYPE *,'INPUT CFL NUMBERS R(0),R(200),RCRIT'
      ACCEPT *,R0,R1,RC

C ***** Omega definition explained in notes

      OMEGA=2.*ATAN(0.5*RC)
```

```
      PI=3.14159
      JMAX=200
      M=400
      XJ(0)=0.

      DO 2 J1=1,JMAX
      DX=1./R(FLOAT(J1)-0.5)
      XJ(J1)=XJ(J1-1)+DX
      U(J1)=0.
2     V(J1)=0.

      DO 3 J1=1,JMAX
      XJ(J1)=XJ(J1)/XJ(JMAX)
3     X(J1)=XJ(J1)

C ***** Initialise U(j,0)+iV(j,0)

      PSI=0.
      J2=JMAX/2
      DO 4 J1=J2-40,J2+40
      IF(MT.LE.3) PHI=ASIN(RC/R(FLOAT(J1)))
      IF(MT.EQ.4) PHI=2.*ATAN(0.5*RC/R(FLOAT(J1)))
      PSI=PSI+PHI
      AMP=EXP(-(J1-J2)**2/200.)
      U(J1)=AMP*COS(PSI)
4     V(J1)=AMP*SIN(PSI)

      W(0)=0.
      KOUNT=0

9     TYPE *,'NO. OF STEPS TILL NEXT PLOT OF U?'
      ACCEPT *,NSTEP
      IF (NSTEP.LE.0) GOTO 5

      DO 6 KOUNT2=1,NSTEP
      KOUNT=KOUNT+1
      IF (KOUNT.GT.M) GOTO 7

C ***** Calculate new U+iV

      CALL METHOD(U,A,W,JMAX,R,MT,MDB)
      CALL METHOD(V,A,W,JMAX,R,MT,MDB)

C ***** Calculate new X,Log(E) of wavepacket

      S=0.
      SJ=0.
      XJ(JMAX+1)=XJ(JMAX)
```

```
      DO 8 J1=1,JMAX
      P=(U(J1)**2+V(J1)**2)*(XJ(J1+1)-XJ(J1-1))/2.
      S=S+P
    8 SJ=SJ+P*XJ(J1)

      EX(KOUNT)=SJ/S
    6 EE(KOUNT)=LOG(S)

    5 CALL OUTPT1(XJ  ,203,'X   ','U   ',NT)
      GOTO 9

C ***** Calculate predicted X(n),Log(E(n))
C ***** Initial valu  j(1) is passed to prediction subroutine PRED
C ***** as TX(1)

    7 TX(1)=J(EX(1))
      TE(1)=EE(1)
      CALL PRED(T,TX,TE,JMAX,M,RC,R,MT,MDB)

C ***** Plot results

   12 TYPE *,'PLOT E,X?'
      ACCEPT 1000,C
      IF(C.EQ.'N   ') GOTO 10
      IF(C.EQ.'E   ') CALL OUTPT2(T,EE,TE,402,'N   ','LN E',NT)
      IF(C.EQ.'X   ') CALL OUTPT2(T,EX,TX,402,'N   ','X   ',NT)
      IF(C.EQ.'Y   ') GOTO 11
      GOTO 12

   11 CALL OUTPT3(T,EX,TX,EE,TE,402,'N   ','X   ','LN E',NT)
   10 CALL PLOT(0.,0.,999)
      STOP

 1000 FORMAT(A4)
      END



C ***** Function calculates R(j)

      FUNCTION R(J)
      COMMON /RCONST/R0,R1,MT
      REAL J

      R=R0*EXP(LOG(R1/R0)*J/200.)
      RETURN

      END
```

```
C ***** Function calculates X(j)

      FUNCTION X(J)
      COMMON /X/XJ(0:200)
      REAL J

      IF(J.GT.200.) J=200.
      IF(J.LT.0.) J=0.
      J1=INT(J)
      X=XJ(J1)+(J-FLOAT(J1))*(XJ(J1+1)-XJ(J1))
      RETURN
      END


C ***** Function calculates j(X)

      FUNCTION J(X)
      COMMON /X/XJ(0:200)
      REAL J

      IF(X.GT.1.) X=1.
      IF(X.LT.0.) X=0.
      J1=0
    1 J1=J1+1
      IF(XJ(J1).LT.X) GOTO 1

      J=FLOAT(J1)-(XJ(J1)-X)/(XJ(J1)-XJ(J1-1))
      RETURN
      END


      SUBROUTINE METHOD(U,A,W,JMAX,R,MT,MDB)

C ***** METHOD sets up the coefficients of the tridiagonal equations
C       for the calculation of the new U using method MT and downstream
C       boundary type MDB, if needed.  The tridiagonal equations are
C       solved by TRID and the new values of U are returned to T.

      DIMENSION U(0:JMAX),A(0:JMAX,3),W(0:JMAX)

      IF(MT.EQ.4) GOTO 1

C ***** Set up coefficients for Trapezoidal and Backward Euler interior
C ***** schemes
```

```
      DO 2 J=1,JMAX-1
      GOTO(3,4,5) MT
    3 C1=-0.25*R(FLOAT(J))
      C2=-C1
      GOTO 6
    4 C1=-0.25*R(FLOAT(J)-0.5)
      C2=0.25*R(FLOAT(J)+0.5)
      GOTO 6
    5 R1=R(FLOAT(J)-0.5)
      R2=R(FLOAT(J)+0.5)
      C1=-0.5*R1**2/(R1+R2)
      C2=0.5*R2**2/(R1+R2)
    6 A(J,1)=C1
      A(J,2)=1.-C1-C2
      A(J,3)=C2
    2 W(J)=U(J)+C1*(U(J)-U(J-1))+C2*(U(J)-U(J+1))

      GOTO(7,8,1) MDB

C ***** Set up coefficients for space and space-time extrapolation at
C        downstream boundary

    7 A(JMAX,1)=-1.
      A(JMAX,2)=1.
      W(JMAX)=0.
      GOTO 9
    8 A(JMAX,1)=0.
      A(JMAX,2)=1.
      W(JMAX)=U(JMAX-1)
      GOTO 9

C ***** Set up coefficients for box method on interior or at downstream
C        boundary as appropriate

    1 JMIN=JMAX
      IF(MT.EQ.4)JMIN=1
      DO 10 J=JMIN,JMAX
      C=R(FLOAT(J)-0.5)
      A(J,1)=1.-C
      A(J,2)=1.+C
      A(J,3)=0.
   10 W(J)=U(J-1)*A(J,2)+U(J)*A(J,1)


C ***** Set up coefficients for upstream boundary

    9 A(0,2)=1.
      A(0,3)=0.
      W(0)=0.
```

```
 14   CALL TRID(U,A,W,JMAX)

      RETURN
      END



      SUBROUTINE TRID(U,A,W,JMAX)

      DIMENSION U(0:JMAX),A(0:JMAX,3),W(0:JMAX)

      DO 1 J=1,JMAX
      C=A(J,1)/A(J-1,2)
      A(J,2)=A(J,2)-A(J-1,3)*C
 1    W(J)=W(J)-W(J-1)*C
      U(JMAX)=W(JMAX)/A(JMAX,2)
      DO 2 J=JMAX-1,0,-1
 2    U(J)=(W(J)-U(J+1)*A(J,3))/A(J,2)
      RETURN
      END



      SUBROUTINE PRED(T,TX,TE,JMAX,M,RC,R,MT,MDB)

      DIMENSION T(M),TX(M),TE(M)
      REAL J,K
      EXTERNAL R,X

      DR(J)=100.*(R(J+0.005)-R(J-0.005))

      PI=3.14159
      J=TX(1)
      TX(1)=X(J)
      T(1)=1.

      IF(MT.EQ.4) GOTO 1

C ***** Prediction for trapezoidal schemes

      K=ASIN(RC/R(J))
      IF(RC.LT.0.) K=PI-K
      OM=ATAN(RC/2.)*2.
      C1=1./(1.+0.25*RC**2)

      DO 2 KOUNT=2,M
      T(KOUNT)=FLOAT(KOUNT)
      DJ=R(J)*COS(K)*C1
```

```
      DK=-DR(J)*SIN(K)*C1
      J=J+0.5*(DJ+R(J+DJ)*COS(K+DK)*C1)
      K=K+0.5*(DK-DR(J+DJ)*SIN(K+DK)*C1)
      TX(KOUNT)=X(J)

      DE=DR(J+0.5*DJ)*C1*(1.-COS(K-0.5*DK))
      IF(MT.EQ.1) TE(KOUNT)=TE(KOUNT-1)
      IF(MT.EQ.2) TE(KOUNT)=TE(KOUNT-1)+DE
      IF(MT.EQ.3) TE(KOUNT)=TE(KOUNT-1)+2.*DE

      IF(J.GT.0.) GOTO 3
      J=-J
      K=ASIN(RC/R(J))
      TX(KOUNT)=X(J)

    3 IF(J.LT.FLOAT(JMAX)) GOTO 2
      J=2.*FLOAT(JMAX)-J
      K=PI-ASIN(RC/R(J))
      TX(KOUNT)=X(J)
      GOTO(6,7,8) MDB
    6 TE(KOUNT)=TE(KOUNT)-2.*LOG(TAN(K/2.))
      GOTO 2
    7 TE(KOUNT)=TE(KOUNT)-2.*LOG(ABS(SIN((K-OM)/2.)/SIN((PI-K-OM)/2.)))
      GOTO 2
    8 TE(KOUNT)=TE(KOUNT)-6.*LOG(TAN(K/2.))


    2 CONTINUE
      RETURN

C ***** Prediction for box scheme

    1 T1=0.25*RC*RC
      T2=1./(1.+T1)
      T3=0.5*RC*T2
      K=2.*ATAN(0.5*RC/R(J))

      DO 9 KOUNT=2,M
      T(KOUNT)=FLOAT(KOUNT)
      IF(J.GE.FLOAT(JMAX)) GOTO 10
      R1=R(J)
      DJ=(R1+T1/R1)*T2
      DK=-2.*DR(J)*T3/R1
      R2=R(J+DJ)
      J=J+0.5*(DJ+(R2+T1/R2)*T2)
      K=K+0.5*(DK+DR(J+DJ)*T3/R2)
      TX(KOUNT)=X(J)
      TE(KOUNT)=TE(KOUNT-1)-2.*T1*T2*DR(J+0.5*DJ)/(R1*R2)
      GOTO 9
```

```
10    TX(KOUNT)=1.
      TE(KOUNT)=TE(KOUNT-1)

 9    CONTINUE

      RETURN

      END


      SUBROUTINE OUTPT1(X,Y,NPLUS2,C1,C2,NT)

      DIMENSION X(NPLUS2),Y(NPLUS2),XD(40)
      N=NPLUS2-2
      NPLUS1=N+1
      XL=5.
      YL=4.
      CALL PLOTS(0,0,NT)

      CALL SCALE(X,XL,N,1)
      CALL SCALE(Y,YL,N,1)

      DO 1 I=1,N/8
 1    XD(I)=(X(8*I+1)-X(8*I-7))/X(NPLUS2)

      IF(NT.EQ.0) GOTO 2
      CALL FACTOR(1.8)
      CALL PLOT(1.,1.,-3)
      CALL AXIS(0.,0.,C1,-4,XL,0.,X(NPLUS1),X(NPLUS2))
      CALL AXIS(0.,0.,C2,4,YL,90.,Y(NPLUS1),Y(NPLUS2))
      CALL LINE(X,Y,N,1,1,3)
      CALL GRID(0.,0.,1000+N/8,XD,-1,YL,-1)
      GOTO 3

 2    CALL FACTOR(1.4)
      CALL PLOT(5.,1.,-3)
      CALL AXIS(0.,0.,C1,-4,XL,90.,X(NPLUS1),X(NPLUS2))
      CALL AXIS(0.,0.,C2,4,YL,180.,Y(NPLUS1),Y(NPLUS2))
      Y(NPLUS2)=-Y(NPLUS2)
      CALL NEWPEN(2)
      CALL LINE(Y,X,N,1,1,3)
      CALL NEWPEN(1)
      CALL GRID(0.,0.,-1,-YL,1000+N/8,XD,-1)

 3    CALL PLOT(0.,0.,-999)
      RETURN
      END
```

```
      SUBROUTINE OUTPT2(X,Y1,Y2,NPLUS2,C1,C2,NT)

      DIMENSION X(NPLUS2),Y1(NPLUS2),Y2(NPLUS2),Y(4)
      N=NPLUS2-2
      NPLUS1=N+1
      XL=5.
      YL=4.

      Y(1)=1.E10
      Y(2)=-1.E10
      DO 1 J=1,N
      Y(1)=AMIN1(Y(1),Y1(J),Y2(J))
    1 Y(2)=AMAX1(Y(2),Y1(J),Y2(J))

      CALL PLOTS(0,0,NT)
      CALL FACTOR(2.0)
      CALL PLOT(1.,0.75,-3)
      CALL NEWPEN(2)
      CALL SCALE(X,XL,N,1)
      CALL SCALE(Y,YL,2,1)
      CALL AXIS(0.,0.,C1,-4,XL,0.,X(NPLUS1),X(NPLUS2))
      CALL AXIS(0.,0.,C2,4,YL,90.,Y(3),Y(4))
      Y1(NPLUS1)=Y(3)
      Y1(NPLUS2)=Y(4)
      Y2(NPLUS1)=Y(3)
      Y2(NPLUS2)=Y(4)
      CALL NEWPEN(3)
      CALL LINE(X,Y1,N,1,25,1)
      CALL LINE(X,Y2,N,1,0,0)
      CALL PLOT(0.,0.,-999)
      RETURN
      END


      SUBROUTINE OUTPT3(X,Y1,Y2,Z1,Z2,NPLUS2,CX,CY,CZ,NT)

      DIMENSION X(NPLUS2),Y1(NPLUS2),Y2(NPLUS2),Y(4)
      DIMENSION Z1(NPLUS2),Z2(NPLUS2),Z(4)
      EQUIVALENCE (Y(1),Z(1))

      N=NPLUS2-2
      NPLUS1=N+1
      XL=4.
      YL=2.
      ZL=2.
```

```
      CALL PLOTS(0,0,NT)
      CALL FACTOR(1.9)
      CALL SCALE(X,XL,N,1)

      Y(1)=1.E10
      Y(2)=-1.E10
      DO 1 J=1,N
      Y(1)=AMIN1(Y(1),Y1(J),Y2(J))
    1 Y(2)=AMAX1(Y(2),Y1(J),Y2(J))
      CALL PLOT(3.5,0.5,-3)
      CALL SCALE(Y,YL,2,1)
      CALL NEWPEN(2)
      CALL AXIS(0.,0.,CX,-4,XL,90.,X(NPLUS1),X(NPLUS2))
      CALL AXIS(0.,0.,CY,4,YL,180.,Y(3),Y(4))
      Y1(NPLUS1)=Y(3)
      Y1(NPLUS2)=-Y(4)
      Y2(NPLUS1)=Y(3)
      Y2(NPLUS2)=-Y(4)
      CALL NEWPEN(3)
      CALL LINE(Y1,X,N,1,25,1)
      CALL LINE(Y2,X,N,1,0,0)
      CALL NEWPEN(1)
      CALL GRID(0.,0.,2,-1.,4,1.,-21846)
      CALL NEWPEN(2)
      CALL PLOT(-2.3,0.4,3)
      CALL PLOT(-2.3,1.0,2)
      CALL SYMBOL(-2.23,1.1,0.14,'WAVEPACKET THEORY',90.,17)
      CALL PLOT(-2.6,0.4,3)
      CALL SYMBOL(-2.6,0.5,0.08,1,0.,-2)
      CALL SYMBOL(-2.6,0.7,0.08,1,0.,-2)
      CALL SYMBOL(-2.6,0.9,0.08,1,0.,-2)
      CALL PLOT(-2.6,1.0,2)
      CALL SYMBOL(-2.53,1.1,0.14,'NUMERICAL EXPERIMENT',90.,20)

      Z(1)=1.E10
      Z(2)=-1.E10
      DO 2 J=1,N
      Z(1)=AMIN1(Z(1),Z1(J),Z2(J))
    2 Z(2)=AMAX1(Z(2),Z1(J),Z2(J))
      CALL PLOT(3.2,0.,-3)
      CALL SCALE(Z,ZL,2,1)
      CALL AXIS(0.,0.,CX,-4,XL,90.,X(NPLUS1),X(NPLUS2))
      CALL AXIS(0.,0.,CZ,4,ZL,180.,Z(3),Z(4))
      Z1(NPLUS1)=Z(3)
      Z1(NPLUS2)=-Z(4)
      Z2(NPLUS1)=Z(3)
      Z2(NPLUS2)=-Z(4)
      CALL NEWPEN(3)
      CALL LINE(Z1,X,N,1,25,1)
```

```
CALL LINE(Z2,X,N,1,0,0)
CALL NEWPEN(1)
CALL GRID(0.,0.,2,-1.,4,1.,-21846)

RETURN

END
```

10. Conclusions

The validity of the asymptotic approach developed in this paper is demonstrated by the numerical results in chapter 9. The limitations of the wavepacket theory are due to the asymptotic approximations involved in treating the wavepacket as a particle. The stability analyses in chapters 5 and 6 use fewer approximations and so the asymptotic errors will be substantially smaller. In particular when the coefficients are constant the analysis in chapter 6 reduces to the P-stability analysis of Beam, Warming and Yee [5].

The calculation of the asymptotic amplitude equation and asymptotic boundary conditions for a particular case is no more difficult than a normal Von Neumann analysis. For applicable cases the wavepacket theory and the stability analysis of chapter 5 are straightforward. The general stability analysis of chapter 6 will usually require numerical computation. In the more complex cases the main benefit from this theory will be the insight given by the asymptotic amplitude equation and boundary conditions. The amplitude equation gives the group velocities of the different wavenumbers and the effect of varying coefficients, which is of great interest since in 2-D cascade geometries cell lengths can vary by factors of up to 100 in inviscid calculations and 1000 in viscous calculations. The asymptotic boundary conditions give the amplitude reflection coefficients which provide a practical criterion for choosing the best numerical boundary conditions.

There are various possibilities for future research

in this area. Further applications to relatively simple problems can be done to gain insight into understanding harder problems and improving boundary conditions. Numerical procedures, such as those suggested in §6.1.1-6.1.3 , can be developed to solve the equations given by the stability analysis in chapter 6. Finally the asymptotic amplitude equation and boundary conditions can be extended to 2-D and 3-D.

References

[1] G.B. Whitham , Linear And Nonlinear Waves , chapter 11 John Wiley & Sons (1974)

[2] J. Lighthill , Waves In Fluids , pp 237-260 , Cambridge Univerity Press (1978)

[3] C.P. Kentzer , 'Group Velocity And Propagation Of Numerical Errors' , AIAA Paper No. 72-153

[4] R. Vichnevetsky and J. Bowles , Fourier Analysis of Numerical Approximations of Hyperbolic Equations , SIAM Studies In Applied Mathematics (1982)

[5] R.M. Beam , R.F. Warming and H.C. Yee , 'Numerical Boundary Conditions For Unconditionally Stable Implicit Difference Approximations Of Hyperbolic Problems' , Paper presented at SIAM 1981 National Meeting, Troy, New York, June 8-10 1981

[6] M.R. Spiegel , Complex Variables , p127-8 , McGraw-Hill (1964)

Appendix

A.1 Finite Difference Operator Notation

An operator notation for finite difference equations simplifies the analysis of finite difference schemes and is a necessity for making any general statements and proving them.

The principal operators are δ_x, central difference, μ_x, central averaging E_x, shift operator, Δ_x, forward difference and ∇_x, backward difference. Their definitions are;

$$\delta_{mx} U_j = U_{j+m/2} - U_{j-m/2} \qquad (A.1(a))$$

$$\mu_{mx} U_j = \frac{1}{2} (U_{j+m/2} + U_{j-m/2}) \qquad (A.1(b))$$

$$E_{mx} U_j = U_{j+m} \qquad (A.1(c))$$

$$\Delta_{mx} U_j = U_{j+m} - U_j \qquad (A.1(d))$$

$$\nabla_{mx} U_j = U_j - U_{j-m} \qquad (A.1(e))$$

Usually these definitions will be used with m=1. The main exceptions are δ_{2x} which is a node-centered central difference,

$$\delta_{2x} U_j = U_{j+1} - U_{j-1} \qquad (A.2)$$

and E_{mx} which can be used to define a general linear operator,

$$L_j U_j = \sum_{m=1}^{M} a_{jm} U_{j+m} \qquad (A.3)$$

$$\text{so} \qquad L_j = \sum_{m=1}^{M} a_{jm} E_{mx} \tag{A.4}$$

When there are several independent variables the subscript on the finite operator denotes the direction of the shift, differencing or averaging. For example if,

$$U_j^n = u(x_j, t_n) \tag{A.5}$$

$$\text{then} \quad \delta_x U_{j+\frac{1}{2}}^n = U_{j+1}^n - U_j^n \tag{A.6}$$

$$\text{and} \quad \delta_t U_j^{n+\frac{1}{2}} = U_j^{n+1} - U_j^n \tag{A.7}$$

The general shift operator expression for a finite operator in 2-D is

$$L_j = \sum_{m,p} C_{mp}(j) E_{mx} E_{pt} \tag{A.8}$$

In applications however this expression can be very complicated and it is usually simpler to express L as a polynomial in the finite operators. As an example the operator in §3.3.2 has the polynomial form

$$L_j = \mu_x \delta_t + r_j \mu_t \delta_x \tag{A.9}$$

but in the shift operator form it is,

$$L_j = \frac{1+r}{2} E_{x/2} E_{t/2} + \frac{1-r}{2} E_{-x/2} E_{t/2} - \frac{1-r}{2} E_{x/2} E_{-t/2} - \frac{1+r}{2} E_{-x/2} E_{-t/2} \tag{A.10}$$

Part of the advantage in using operator notation when analysing finite difference schemes arises because all of the finite operators have the same eigenfunction which in 2-D is $\exp[i(j\phi - n\Omega)]$.

$$\delta_x \exp[i(j\phi-n\Omega)] = \exp[i((j+\tfrac{1}{2})\phi-n\Omega)] - \exp[i((j-\tfrac{1}{2})\phi-n\Omega)]$$

$$= 2i \sin(\phi/2) \exp[i(j\phi-n\Omega)] \qquad (A.11(a))$$

$$\mu_x \exp[i(j\phi-n\Omega)] = \frac{1}{2}\left(\exp[i((j+\tfrac{1}{2})\phi-n\Omega)] + \exp[i((j-\tfrac{1}{2})\phi-n\Omega)] \right)$$

$$= \cos(\phi/2) \exp[i(j\phi-n\Omega)] \qquad (A.11(b))$$

$$E_x \exp[i(j\phi-n\Omega)] = \exp[i((j+1)\phi-n\Omega)]$$

$$= \exp(i\phi) \exp[i(j\phi-n\Omega)] \qquad (A.11(c))$$

$$\Delta_x \exp[i(j\phi-n\Omega)] = \exp[i((j+1)\phi-n\Omega)] - \exp[i(j\phi-n\Omega)]$$

$$= \{ \exp(i\phi) - 1 \} \exp[i(j\phi-n\Omega)] \qquad (A.11(d))$$

$$\nabla_x \exp[i(j\phi-n\Omega)] = \exp[i(j\phi-n\Omega)] - \exp[i((j-1)\phi-n\Omega)]$$

$$= \{ 1 - \exp(-i\phi) \} \exp[i(j\phi-n\Omega)] \qquad (A.11(e))$$

and similarly,

$$\delta_t \exp[i(j\phi-n\Omega)] = -2i \sin(\Omega/2) \exp[i(j\phi-n\Omega)] \qquad (A.11(f))$$

$$\mu_t \exp[i(j\phi-n\Omega)] = \cos(\Omega/2) \exp[i(j\phi-n\Omega)] \qquad (A.11(g))$$

$$E_t \exp[i(j\phi-n\Omega)] = \exp(-i\Omega) \exp[i(j\phi-n\Omega)] \qquad (A.11(h))$$

$$\Delta_t \exp[i(j\phi-n\Omega)] = \{ \exp(-i\Omega) - 1 \} \exp[i(j\phi-n\Omega)] \qquad (A.11(i))$$

$$\nabla_t \exp[i(j\phi-n\Omega)] = \{ 1 - \exp(i\Omega) \} \exp[i(j\phi-n\Omega)] \qquad (A.11(j))$$

A.2 Principle of the Argument

Let f(z) be an analytic complex function with simple poles in a region of the complex z-plane and let C be a closed curve in the region. Then the number of zeros of F minus the number of poles of f lying inside C is equal to,

$$\frac{1}{2\pi i} \int_C \frac{f'(z)}{f(z)}\, dz$$

$$= \frac{1}{2\pi i} \left[\ln(f) \right]_C$$

$$= \frac{1}{2\pi} \left[\arg(f) \right]_C \qquad (A.12)$$

$[\ ]_C$ denotes the change as z goes round C anticlockwise.

arg(f) is defined by,

$$f(z) = R \exp(i\theta) \qquad R,\theta \text{ real} \quad R>0 \qquad (A.13)$$

$$\arg(f) = \theta \qquad (A.14)$$

with the restriction that θ must vary continuously as z goes round C.

The proof is given in many standard texts on complex analysis, e.g. [6]. This provides a very simple test when considering stability problems in which it is sufficient to know whether there are any zeros in a critical region without knowing their exact position. This is the basis of the Nyquist criterion in control theory stability analysis. The test can also be performed numerically relatively easily. The step size Δz in going round C is decreased, if

necessary, until $|\Delta\arg(f)| < \varepsilon$. Since $\frac{1}{2\pi}\,[\,\arg(f)\,]_C$ is an integer there is no rounding error. The only possible error is if the magnitude of $\Delta\arg(f)$ over one step lies in the range $2n\pi-\varepsilon < \Delta\arg(f) < 2n\pi+\varepsilon$ for some integer n other than zero. Decreasing ε reduces the chance of an error at the expense of increased computation. $\varepsilon=\pi/6$ should be adequate in most cases.

A.3 Definitions of Norms and Stability

The norms used in this paper are generalised L_2 norms. For a continuous function $u(x,t)$ defined on $0<x<X$ the norm $||u(t)||$ is defined by,

$$||u(t)||^2 = \int_0^X |u(x,t)|^2 \, \alpha(x) \, dx \qquad (A.15)$$

where $\alpha(x)$ is a positive non-zero function.

For a discrete function U_j^n defined on $0<j<J$ the norm $||U^n||$ which is a function of n is defined by,

$$||U^n||^2 = \sum_0^J |U_j^n|^2 \, \alpha_j \qquad (A.16)$$

where α_j is a positive non-zero function.

The stability used in this paper is asymptotic Liapounov stability which is defined as,

Given $\delta>0$ there exists $\varepsilon>0$ such that

$||u(0)|| < \varepsilon$ ==> a) $||u(t)|| < \delta$

and b) $||u(t)|| \to 0$ as $t \to \infty$

Condition a) is Liapounov stability which limits how large an initially small disturbance can become. Condition b) is asymptotic stability which specifies that a sufficiently small initial disturbance must ultimately tend to zero.

For linear systems of equations an equivalent definition is

a) There exists $M>0$ such that $||u(t)|| < M \, ||u(0)||$

and b) $||u(t)|| \to 0$ as $t \to \infty$

The corresponding definitions for a discrete function are,

Given $\delta > 0$ there exists $\varepsilon > 0$ such that

$||U^0|| < \varepsilon \implies$ a) $||U^n|| < \delta$

and b) $||U^n|| \to 0$ as $n \to \infty$

and for linear discrete systems,

a) There exists $M > 0$ such that $||U^n|| < M ||U^0||$

and b) $||U^n|| \to 0$ as $n \to \infty$