Proceedings of the Edinburgh Mathematical Society (1982) 25, 1–18 ©

# DERIVATION OF BOUNDARY CONDITIONS FOR THE ARTIFICIAL BOUNDARIES ASSOCIATED WITH THE SOLUTION OF CERTAIN TIME DEPENDENT PROBLEMS BY LAX–WENDROFF TYPE DIFFERENCE SCHEMES

*by* JOHN C. WILSON

(Received 13th December 1979)

## 1. Introduction

Many problems involving the solution of partial differential equations require the solution over a finite region with fixed boundaries on which conditions are prescribed. It is a well known fact that the numerical solution of many such problems requires additional conditions on these boundaries and these conditions must be chosen to ensure stability. This problem has been considered by, amongst others, Kreiss [**11**, **12**, **13**], Osher [**16**, **17**], Gustafsson et al. [**9**] Gottlieb and Tarkel [**7**] and Burns [**1**]

It is also necessary to solve partial differential equations over infinite domains, this type of problem occurring in transonic flows, seismology and meteorology. A numerical solution must be over a finite domain and one method of limiting the area of computation is to use artificial boundaries on which suitable conditions must be obtained. These conditions must be such that the solution of the modified problem is close to that of the original one over their common domain. If there is exact correspondence between the solutions the boundary condition becomes non-reflecting.

Engquist and Majda [**2**, **3**] examined such a problem, using the theory of pseudo-differential operators to construct well-posed boundary conditions for wave and other differential equations. Although the ideal non-reflecting boundary conditions are non-local in both time and space, practical computing considerations required them to use conditions which are local in both time and space. For the numerical problem these local conditions were approximated in a stable manner and it was shown that the resulting reflection at the artificial boundaries was small. In the second paper Engquist and Majda also considered the construction of radiation boundary conditions for the difference equation approximating the differential equation. For a finite difference approximation to the wave equation they obtained the symbol of the theoretical discrete-radiation boundary condition from the symbol of the approximation to the differential equation. Practical stable finite difference boundary conditions involving a parameter were given such that their symbol approximates closely that of the theoretical boundary condition. It is suggested that the parameter be chosen to minimise the truncation error of the boundary approximation and an estimate of the reflection coefficient for the boundary is given.

Gustafsson and Kreiss [**10**] examine the problem of artificial boundaries in a different manner though some of the boundary approximations they derive are equivalent to the 'no reflection' principle of Engquist and Majda. Initially a scalar problem on the half line $0 \leqq x < \infty$ was considered and approximated by a similar problem over a finite interval $0 \leqq x \leqq a$ so that a boundary condition is required at $x = a$. They sought conditions for the Lax–Wendroff scheme such that convergence occurred and they extended their results to systems with constant coefficients. The variable coefficient problem was also considered and the structure for conditions on the artificial boundary obtained. However only for certain cases can they be represented in a simple manner. In their analysis Gustafsson and Kreiss used the correspondence at $x = a$ between the solutions for $0 \leqq x \leqq a$ and $a \leqq x < \infty$. A somewhat similar approach is used in this paper.

The equation

$$\frac{\partial \boldsymbol{u}}{\partial t} + \frac{\partial \boldsymbol{F}(\boldsymbol{u})}{\partial x} = 0 \tag{1.1}$$

can, for example, be used to represent, in conservation form, hydrodynamic flows in one space variable, $\boldsymbol{u}$ being a three vector for this with components of density, momentum and energy and $\boldsymbol{F}(\boldsymbol{u})$ is a vector function of $\boldsymbol{u}$. Consider two problems involving the solution of an equation of the type given by (1.1), firstly on the half line $0 \leqq x < \infty$, $t \geqq 0$ with initial conditions $\boldsymbol{u}(x, 0) = \boldsymbol{f}(x)$ and some boundary condition on $x = 0$ and secondly on the line $-\infty < x < \infty$ with $\boldsymbol{u}(x, 0) = 0$ for $x < 0$ and $\boldsymbol{u}(x, 0) = \boldsymbol{f}(x)$ for $x \geqq 0$. The problems are discretised with a rectangular mesh, of sides $\Delta x$ in the space direction at $\Delta t$ in the time direction, imposed on the solution domain such that $x = 0$ is a mesh line. Suppose (1.1) is approximated by the one space variable Lax–Wendroff scheme. With regard to the second problem implementation of this scheme at mesh points in the region $x < 0$ shows that $\boldsymbol{U}_0^n$ can be evaluated in terms of $\boldsymbol{U}_j^N$, $N < n$ and $j = 0$ or 1, $\boldsymbol{U}_j^n$ denoting the value of $\boldsymbol{u}$ at the point $(j\,\Delta x, n\,\Delta t)$.

The solution of the first problem using the Lax–Wendroff scheme requires a boundary condition at $x = 0$ and matching the solutions of the two problems at $x = 0$ suggests a suitable boundary condition to be of the form

$$\boldsymbol{U}_0^{n+1} = \sum_{k=1}^{n} C_k \boldsymbol{U}_1^{n-k+1}. \tag{1.2}$$

This condition is local in space but non-local in time and the $C_k$ sequence has to be determined.

In Section 2 a generalisation of (1.1) and (1.2) to any finite number of space variables is considered and a recurrence relation obtained for the corresponding $C_k$ sequence. The stable boundary conditions obtained are non-local in both time and space and it was found impossible to implement them because of storage requirements. However usable stable boundary conditions for the one space variable case are obtained. Initially the conditions are obtained as a special case of the multidimensional problem and involve some storage of past data. They are referred to as 'time dependent' boundary conditions. 'Space dependent' conditions, which are local in time but not in space, are obtained in Section 3 for the one space variable case and involve no storage of past information.

Finally, in Section 4, the use of the one dimensional boundary conditions for higher dimensional problems, with a Strang scheme used at the interior points, is examined.

Gourlay and Morris [**5**, **6**] and Morris and McGuire [**14**] considered the use of Strang schemes for multidimensional problems with fixed boundaries and values prescribed on some of them, that is they considered the situation where the characteristic conoid from a point on the boundary cuts this boundary at only that point. Morris and McGuire [**15**] also examined this problem and developed a method for feeding in the boundary data such that stability was not upset.

A more difficult situation arises when artificial boundaries have to be used and the characteristic conoid from a point on such a boundary cuts that boundary at more than one point. The 'time dependent' one space variable boundary conditions are modified for use with the Strang schemes and it is shown that stable boundary conditions of the required accuracy are obtained. A similar result is obtained by modifying the 'space dependent' conditions for use with the Strang schemes.

## 2. The problem with artificial boundaries

### 2.1. Time dependent boundary conditions for quarter space problems

The equation

$$\frac{\partial \boldsymbol{u}}{\partial t}+\sum_{p=1}^{m}\frac{\partial \boldsymbol{F}_p(\boldsymbol{u})}{\partial x_p}=0 \tag{2.1}$$

in $m$ ($\geqq 1$) space variables $x_p$, $-\infty < x_p < \infty$, $t \geqq 0$, $p = 1, 2, \ldots, m$, with $\boldsymbol{u}$ given at $t=0$ and $\boldsymbol{u}$ asymptotically constant as any $x_p \to \pm\infty$ describes an initial value problem. (2.1) is taken to be hyperbolic and in it $\boldsymbol{F}_p(\boldsymbol{u})$ is a vector of the same dimension as the vector $\boldsymbol{u}$. To enable this problem to be solved numerically two artificial boundaries are required in each space direction. These boundaries are taken normal to the co-ordinate axes and without loss of generality can be considered to be situated at $x_p = 0$ and $x_p = a_p > 0$, $p = 1, 2, \ldots, m$.

Suitable conditions must be derived for each of these boundaries. To obtain these conditions the problem may be considered to be the intersection of $2m$ quarter space problems, two in each space direction. Typical are the problems associated with the boundaries $x_P = 0$, $x_P = a_P$, $P$ a value of $p$. The first, a right quarter space problem, requires an expression to be determined for $\boldsymbol{u}$ at $x = 0$ for the problem defined by (2.1), $0 \leqq x_P < \infty$, $t \geqq 0$ with $\boldsymbol{u}$ given at $t = 0$ and the second, a left quarter space problem, requires an expression for $\boldsymbol{u}$ at $x_P = a_P$ for $-\infty < x_P \leqq a_P$, $t \geqq 0$. In both problems $-\infty < x_p < \infty$, $p = 1, 2, \ldots, m$, $p \neq P$. By virtue of the similarity of the problems in each space direction analysis need only be carried out for the artificial boundaries associated with one of the space variables, namely $x_P$.

Provided the boundary is sufficiently far removed from any source of disturbance and the restriction imposed on $\boldsymbol{u}$ in (2.1) applies the linearised form of (2.1) can be used in the vicinity of the boundary. Therefore for the purpose of the boundary analysis (2.1) is considered to be replaced by

$$\frac{\partial \boldsymbol{u}}{\partial t}+\sum_{p=1}^{m}E_p\frac{\partial \boldsymbol{u}}{\partial x}=0, \tag{2.2}$$

where $E_p(\boldsymbol{u})$ is the Jacobian of $\boldsymbol{F}_p(\boldsymbol{u})$ and $E_p = E_p(\boldsymbol{u})$ is taken to be constant. The problem is discretised and (2.2) approximated by the stable difference scheme

$$\boldsymbol{U}_j^{n+1} = \alpha_1 \boldsymbol{U}_{j+1}^n + \alpha_0 \boldsymbol{U}_j^n + \alpha_{-1} \boldsymbol{U}_{j-1}^n \tag{2.3}$$

where $\boldsymbol{U}_j^n$ is a vector containing all the nodal values at the time station $t = n\,\Delta t$ and the space station $x = j\,\Delta x$. $\boldsymbol{U}$ is required at all the mesh points in the solution domain and it is assumed that the vector space remains finite. The $\alpha$'s are square matrix operators which can only be simultaneously reduced to diagonal form in the one space variable case.

The discrete forms of the initial and boundary value problems are now considered and without loss of generality $P$ is taken to be one in the analysis.

**Problem 1.** (2.3) is to be solved for $-\infty < j < \infty$ with initial conditions $\boldsymbol{U}_j^0$ given for all $j \geqq 0$ and $\boldsymbol{U}_j^0 = 0$ for $j < 0$.

**Problem 2.** (2.3) is to be solved in the quarter space $0 \leqq x_1 < \infty$, $-\infty < x_p < \infty$, $t \geqq 0$ and $p = 2, 3, \ldots, m$. $\boldsymbol{U}_j^{n+1}$ for $j > 0$ is obtained from (2.3) provided suitable initial and boundary conditions are prescribed. The initial conditions are those for Problem 1 for $j \geqq 0$ and $\boldsymbol{U}_0^n$ is obtained from the boundary conditions

$$\sum_{k=1}^{n} A_k \boldsymbol{U}_0^{n-k+1} = \sum_{k=1}^{n} B_k \boldsymbol{U}_1^{n-k}, \tag{2.4}$$

$A_k$ and $B_k$ being constant matrices of the same order as the $\alpha$'s and have to be determined.

These matrices are determined by demanding exact correspondence between the solutions of Problems 1 and 2 at $x_1 = 0$. A corresponding left quarter space problem can be defined for $-\infty < x_1 \leqq N\,\Delta x_1$, $-\infty < x_p < \infty$, $p = 2, 3, \ldots, m$, $t \geqq 0$ where $N\,\Delta x_1 = a_1$, $N$ a natural number. The boundary condition to be applied at $x = N\,\Delta x_1$ is

$$\sum_{k=1}^{n} \hat{A}_k \boldsymbol{U}_N^{n-k+1} = \sum_{k=1}^{n} \hat{B}_k \boldsymbol{U}_{N-1}^{n-k}.$$

The matrices in this are constant and the method of determining them is similar to that for $A_k$ and $B_k$ in (2.4) so that the analysis will be confined to that case.

Problems 1 and 2 are solved by discrete Laplace Transforms. The transform is defined by

$$\tilde{\boldsymbol{U}}_j(s) = \sum_{k^1=0}^{\infty} \boldsymbol{U}_j(k^1) \exp(-sk^1\,\Delta t), \quad \text{Re}(s) \geqq c,$$

with $c$ large enough to ensure the uniform convergence of the series. Multiply (2.3) by $\exp(-sn\,\Delta t)$, sum $n = 0$ to $\infty$ and simplify to give

$$\alpha_1 \tilde{\boldsymbol{U}}_{j+1} + (\alpha_0 - \exp(s\,\Delta t)I)\tilde{\boldsymbol{U}}_j + \alpha_{-1}\tilde{\boldsymbol{U}}_{j-1} = \boldsymbol{f}_j, \tag{2.5}$$

where

$$\boldsymbol{f}_j = -\exp(s\,\Delta t)\boldsymbol{U}_j^0 \quad \text{and} \quad \boldsymbol{f}_j = 0 \quad \text{for} \quad j < 0.$$

https://doi.org/10.1017/S0013091500004053 Published online by Cambridge University Press

Let

$$\tilde{A}=\sum_{k=1}^{\infty} A_k \xi^{k-1}, \quad \tilde{B}=\sum_{k=1}^{\infty} B_k \xi^k, \tag{2.6}$$

where $\xi=\exp(-s\,\Delta t)$ so that transforming (2.4) in a manner similar to (2.3) gives

$$\tilde{A}\tilde{\boldsymbol{U}}_0=\tilde{B}\tilde{\boldsymbol{U}}_1. \tag{2.7}$$

The homogeneous form of (2.5) is

$$\alpha_1 \mathbf{V}_{j+1}+(\alpha_0-\exp(s\,\Delta t)I)\mathbf{V}_j+\alpha_{-1}\mathbf{V}_{j-1}=0. \tag{2.8}$$

A solution of the form $\boldsymbol{V}_j=\lambda^j\boldsymbol{X}$ is considered for this and for $\boldsymbol{X}\neq 0$ this requires

$$\operatorname{Det}[\lambda^2\alpha_1+\lambda(\alpha_0-\exp(s\,\Delta t)I)+\alpha_{-1}]=0.$$

Assuming the $\alpha$'s are of order $b\times b$ this is a polynomial of degree $2b$ in $\lambda$. In general, if $\alpha_1$ and $\alpha_{-1}$ are non-singular with linear divisors and Re $(s)$ is large, the roots separate asymptotically into two sets of $b$ roots. This gives

(i) $\lambda^-\sim\exp(s\,\Delta t)/\lambda(\alpha_1)$ with $\boldsymbol{X}^-$ the eigenvector corresponding to the eigenvalue $\lambda(\alpha_1)$.

(ii) $\lambda^+\sim\exp(-s\,\Delta t)\lambda(\alpha_{-1})$ with $\boldsymbol{X}^+$ the eigenvector corresponding to the eigenvalue $\lambda(\alpha_{-1})$.

The sets can be analytically continued for smaller Re $(s)$ and the corresponding $\alpha$'s and $\boldsymbol{X}$'s are written $\lambda^-(s)$, $\boldsymbol{X}^-(s)$, $\lambda^+(s)$ and $\boldsymbol{X}^+(s)$. Writing $X^*=[X_1^*X_2^*\dots X_b^*]$, * representing either + or −, defining $P_+(s)$, $P_-(s)$ by

$$P_+(s)=X^+\operatorname{diag}(\lambda_1^+,\lambda_2^+,\dots,\lambda_b^+)(X^+)^{-1}=X^+\Delta^+(X^+)^{-1} \tag{2.9a}$$

$$P_-(s)=X^-\operatorname{diag}(\lambda_1^-,\lambda_2^-,\dots,\lambda_b^-)(X^-)^{-1}=X^-\Delta^-(X^-)^{-1} \tag{2.9b}$$

and substituting the expressions for $\boldsymbol{V}_j$ into (2.8) leads to

$$\alpha_1 P_+^2+(\alpha_0-\exp(s\,\Delta t)I)P_++\alpha_{-1}=0 \tag{2.10}$$

$$\alpha_1 P_-^2+(\alpha_0-\exp(s\,\Delta t)I)P_-+\alpha_{-1}=0. \tag{2.11}$$

Thus $P_+$, $P_-$ satisfy (2.8) so that its general solution is, with $\boldsymbol{C}_+$, $\boldsymbol{C}_-$ independent of $j$,

$$\mathbf{V}_j(s)=(P_+)^j\boldsymbol{C}_++(P_-)^j\boldsymbol{C}_-.$$

For Re $(s)\to\infty$ (2.9) gives $\|P_+\|\to 0$, $\|P_-\|\to\infty$ so that for large Re $(s)$ $\|P_+\|<1$, $\|P_-\|>1$. Thus, for large Re $(s)$, $(P_+)^\beta\to 0$ as $\beta\to\infty$ and $(P_-)^\beta\to 0$ as $\beta\to-\infty$. The solutions to Problems 1 and 2 can thus be given in terms of $P_+$ and $P_-$. For Problem 1, $\tilde{\boldsymbol{U}}_j$ is taken such that

$$\tilde{\boldsymbol{U}}_j=\sum_{i=1}^{j-1} P_+^{j-i}K\boldsymbol{f}_i+\sum_{i=j}^{\infty} P_-^{j-i}K\boldsymbol{f}_i. \tag{2.12}$$

Substituting (2.12) into (2.5) and equating coefficients of $\boldsymbol{f}_j$ leads to

$$K=[\alpha_1 P_++(\alpha_0-\exp(s\,\Delta t)I)+\alpha_{-1}P_-^{-1}]^{-1}. \tag{2.13}$$

The expression $\alpha_1 P_+ + \alpha_0 - \exp(s\,\Delta t)I + \alpha_{-1}P_-^{-1}$ and consequently $U_j(s)$ may be singular for specific values of $s$. Such singularities give a contribution proportional to $\exp(sk\,\Delta t)$ in the calculation of $U_j^k$. For $\mathrm{Re}\,(s) > 0$ the stability of (2.3) ensures no such points exist. For Problem 2, $\tilde{U}_j$ is taken such that

$$\tilde{U}_j = \sum_{i=1}^{j-1} P_+^{j-i} L_i f_i + \sum_{i=j}^{\infty} (P_+^{j-i} M_i + P_-^{j-i} N_i) f_i. \tag{2.14}$$

Substituting this into (2.5) and (2.7) with the coefficients of $f_j$ equated gives the following relations for $L_i$, $M_i$ and $N_i$:

$$\begin{aligned} \alpha_1 P_+ L_i + (\alpha_0 - \exp(s\,\Delta t) I)(M_i + N_i) + \alpha_{-1}(P_+^{-1} M_i + P_-^{-1} N_i) = I \\ \tilde{A}(P_+^{-i} M_i + P_-^{-i} N_i) - \tilde{B}(P_+^{1-i} M_i + P_-^{1-i} N_i) = 0. \end{aligned} \tag{2.15}$$

When $L_i = K = N_i$ and $M_i = 0$ there will be exact correspondence between the solutions to Problems 1 and 2 and it immediately follows that

$$\tilde{A} = \tilde{B} P_-. \tag{2.16}$$

The 'exact' boundary condition for the right quarter space problem is obtained from this. A similar result holds for the left quarter space problem, the analysis being similar with $\alpha_1$ and $\alpha_{-1}$ interchanged.

### 2.2. 'Exact' boundary conditions for the quarter space problems

The right quarter space problem is considered first, 'exact' conditions being obtained from (2.16) by taking $\tilde{A} = I$, $\tilde{B} = P_-^{-1}$. From (2.5) $B_k$ is the coefficient of $\xi^k$ in the expansion of $B$ in powers of $\xi$. For large $\mathrm{Re}\,(s)$, (2.11) gives $P_- \sim \xi^{-1}\alpha_1^{-1}$ so that $P_-^{-1} \sim \xi\alpha_1$ is an initial approximation for $P_-^{-1}$. Equation (2.11) can be rearranged to give

$$P_-^{-1} = \xi(\alpha_1 + \alpha_0 P_-^{-1} + \alpha_{-1} P_-^{-2}). \tag{2.17}$$

Successively better approximations to $P_-^{-1}$ and hence $B_k$ are obtained from this, (2.17) producing from $P_- \sim \xi^{-1}\alpha_1^{-1}$ the sequence

$$\begin{aligned} &\xi\alpha_1 + \xi^2\alpha_0\alpha_1;\ \xi\alpha_1 + \xi^2\alpha_0\alpha_1 + \xi^3(\alpha_0^2\alpha_1 + \alpha_{-1}\alpha_1^2);\ \xi\alpha_1 + \xi^2\alpha_0\alpha_1 \\ &\qquad + \xi^3(\alpha_0^2\alpha_1 + \alpha_{-1}\alpha_1^2) + \xi^4(\alpha_0^3\alpha_1 + \alpha_0\alpha_{-1}\alpha_1^2 + \alpha_{-1}\alpha_0\alpha_1^2 + \alpha_{-1}\alpha_1\alpha_0\alpha_1); \ldots \end{aligned}$$

from which, since $B_k$ is the coefficient of $\xi^k$, follow the results

$$\begin{aligned} B_1 &= \alpha_1 \\ B_2 &= \alpha_0\alpha_1 \\ B_k &= \alpha_0 B_{k-1} + \alpha_{-1}\sum_{l'=1}^{k-2} B_{k-1-l'} B_{l'}, \quad k \geqq 3. \end{aligned} \tag{2.18}$$

Since the only boundary under consideration is $x_p = 0$ the simplest way of implementing (2.18) is to regard $U_j^n$ in (2.3) as the fourier transform of all the independent variables other than $t$ and $x_p$. Thus if $G(\eta_1, \ldots, \eta_m)$ is the amplication matrix for a component proportional to $\exp\left(i\hat{\Sigma}\dfrac{\eta_q x_q}{\Delta x_q}\right)$, where $\hat{\Sigma}$ means $\sum_{q=1, q\neq p}^{m}$, then

$$G(\eta_1, \ldots, \eta_m) = \alpha_1 e^{i\eta_q} + \alpha_0 + \alpha_{-1} e^{-i\eta_q},$$

the $\alpha$'s being functions of $\exp(i\gamma)$, where $\gamma = \hat{\Sigma}\eta_q$. There is a contribution to each of the $\alpha$'s from each mesh point at the appropriate space station. Let $2t_q+1$ be the width of support in the direction $x_q$, $q \neq p$. Denote by $\Sigma_f = \Sigma_{d_f=-t_f}^{t_f} \sum_{1\leq f\leq m, f\neq p}$ the summation over such mesh points and let $\Sigma_k = \Sigma_{d_f=-kt_f}^{kt_f} \sum_{1\leq f\leq m, f\neq p}$ denote the summation over mesh points contributing to $B_k$, it being noted that the domain of dependence for these extends back to the initial hyperplane. Then, for $\delta = -1, 0, 1$, with

$$\boldsymbol{d} = d_1, \ldots, d_{p-1}, d_{p+1}, \ldots, d_m, \quad \alpha_\delta = \Sigma_f \prod_{\substack{q=1\\ q\neq p}}^{m} \exp(id_q\eta_q)\alpha_{\delta;\boldsymbol{d}}$$

and

$$B_k = \Sigma_k \prod_{\substack{q=1\\ q\neq p}}^{m} \exp(id_q\eta_q) B_{k,\boldsymbol{d}}$$

with $B_{k,\boldsymbol{d}}$ interpreted as zero for $|d_f| > kt_f$.

With $\boldsymbol{j}_l = j_1, \ldots, j_{p-1}, l, j_{p+1}, \ldots j_m$ equation (2.4) becomes

$$\boldsymbol{U}_{\boldsymbol{j}_0}^n = \sum_{k=1}^{n} \Sigma_k B_{k,\boldsymbol{d}} \boldsymbol{U}_{\boldsymbol{j}_1}^{n-k}. \tag{2.19}$$

Equating coefficients of $\exp(i\hat{\Sigma}d_q\eta_q)$ in (2.18) gives equations for the computation of the $B_{k,\boldsymbol{d}}$'s and this could be done on a "once and for all" basis. However the storage problem created makes this impracticable. In an attempt to overcome this generating functions were examined, these involving a summation of the product of $B_{k,\boldsymbol{d}}$, a time function and powers of the $x_p$'s. Only the one space variable case will be considered, the higher dimensional cases leading to a quadratic matrix equation which is a generalisation of (2.23). This is considered in the next section.

For the left quarter space problem a result similar to (2.19) is obtained with the 1, 0 in $\boldsymbol{j}_l$ replaced by $N-1$ and $N$ respectively. The $B_k$ sequence is generated from $P_-^{-1} \sim \xi\alpha_{-1}$ and is as in (2.18) with $\alpha_1$ and $\alpha_{-1}$ interchanged. The stability of these boundary approximations is easily established. In the given problem the method of constructing the boundary conditions is such that the solution at mesh points on the artificial boundaries and within the region enclosed by them is identical to the solution obtained by applying the difference scheme at the same points for the infinite problem. Since the finite difference representation of the infinite problem is stable then the finite system is necessarily also stable and results in no reflection at the boundaries.

### 2.3. One dimensional time dependent boundary conditions

In one space dimension (2.1) is replaced by (1.1) whose linearised form is

$$\frac{\partial \boldsymbol{u}}{\partial t} + E\frac{\partial \boldsymbol{u}}{\partial x} = 0. \tag{2.20}$$

https://doi.org/10.1017/S0013091500004053 Published online by Cambridge University Press

The generating function in this case is, with $c_k$ the one dimensional form of $B_{k,\boldsymbol{d}}$,

$$g(t)=\sum_{k=1}^{\infty} c_k t^{k-1}=c_1+c_2t+\sum_{k=3}^{\infty} c_k t^{k-1}. \tag{2.21}$$

Initially the right quarter plane problem is considered and for this (2.18) gives

$$c_1=\alpha_1, \quad c_2=\alpha_0 c_1 \quad \text{and} \quad c_k=\alpha_0 c_{k-1}+\alpha_{-1}\sum_{l'=1}^{k-2} c_{k-1-l'}c_{l'} \quad \text{for} \quad k \geqq 3.$$

$$\sum_{k=3}^{\infty} c_k t^{k-1}=\sum_{k=3}^{\infty} \alpha_0 c_{k-1} t^{k-1}+\sum_{k=3}^{\infty} \alpha_{-1}\sum_{l'=1}^{k-2} c_{k-1-l'}c_{l'}t^{k-1}. \tag{2.22}$$

Writing $k'+k''=k-1$ so that

$$\sum_{k=3}^{\infty}\sum_{l'=1}^{k-2} c_{k-1-l'}c_{l'}t^{k-1}=t^2\sum_{k'=1}^{\infty} c_{k'}t^{k'-1}\sum_{k''=1}^{\infty} c_{k''}t^{k''-1}$$

it follows that the last term in (2.22) reduces to $\alpha_{-1}t^2g^2(t)$ and the first term in (2.22) combines with $\alpha_0 c_1 t$ to give $\alpha_0 t g(t)$. Thus (2.21) becomes

$$\alpha_{-1}t^2g^2(t)-(I-\alpha_0 t)g(t)+\alpha_1=0 \tag{2.23}$$

Since only one space variable is involved the problem can be reduced to diagonal form, with say $b_1$ of the eigenvalues $\lambda$ of $E$ negative and $b-b_1$ of them positive. The unknown is now represented by $\boldsymbol{v}=X^{-1}\boldsymbol{u}$, $X$ the matrix of eigenvectors of $E$. Representation can thus be made as a series of scalar equations in the components $v$ of $\boldsymbol{v}$ so that (2.23) can be regarded as a scalar equation. Given $q=\dfrac{\lambda\,\Delta t}{\Delta x}$, where $\Delta t$ and $\Delta x$ are respectively time and space increments, the $\alpha$'s will be functions of $q$ and a study of the characteristics at an artificial boundary shows that a condition at such a boundary is 'essential' if $q<0$. 'Essential' conditions are considered first.

Since $g(t)$ must be non-singular at $t=0$ the negative sign applies to the discriminant of the quadratic. With $a=t(\alpha_0^2-4\alpha_1\alpha_{-1})^{1/2}$ and $\mu=\alpha_0(\alpha_0^2-4\alpha_1\alpha_{-1})^{-1/2}$ the solution of (2.23) is

$$g=\frac{1-\alpha_0 a(\alpha_0^2-4\alpha_1\alpha_{-1})^{-1/2}-(1-2\mu a+a^2)^{1/2}}{2\alpha_{-1}a^2(\alpha_0^2-4\alpha_1\alpha_{-1})^{-1}}. \tag{2.24}$$

Define $Q_\phi$ by

$$(1-2\mu a+a^2)^{1/2}=\sum_{\phi=0}^{\infty} a^\phi Q_\phi(\mu), \tag{2.25}$$

so that $Q_0=1$, $Q_1=-\mu$. Using the generating function for Legendre polynomials $P_\phi$ it follows that

$$Q_\phi(\mu)=P_\phi(\mu)-2\mu P_{\phi-1}(\mu)+P_{\phi-2}(\mu),$$

and by using the recurrence relation between $P_\phi$, $P_{\phi-1}$ and $P_{\phi-2}$ that

$$Q_\phi(\mu)=\frac{-1}{2\phi-1}\{P_\phi(\mu)-P_{\phi-2}(\mu)\}, \quad \phi \geqq 2. \tag{2.26}$$

Substituting (2.25) into (2.24), simplifying and using (2.21) gives, with $\phi = k$,

$$\sum_{k=1}^{\infty} c_k t^{k-1} = \frac{-1}{2\alpha_{-1}} \sum_{k=2}^{\infty} (\alpha_0^2 - 4\alpha_1\alpha_{-1})^{k/2} t^{k-2} Q_k(\mu), \quad c_k = c_k(\mu).$$

Equating the co-efficients of $t^{k+1}$ in this and using (2.26) produces

$$c_{k+1} = \frac{(\alpha_0^2 - 4\alpha_1\alpha_{-1})^{k/2+1}}{2\alpha_{-1}(2k+3)} \{P_{k+2}(\mu) - P_k(\mu)\}, \quad k \geqq 2. \tag{2.27}$$

Using the above mentioned recurrence relation for Legendre polynomials it follows that

$$P_{\phi+1}(\mu) - P_{\phi-1}(\mu) = \frac{2\phi+1}{\phi+1} \mu\{P_\phi(\mu) - P_{\phi-2}(\mu)\} - \frac{(\phi-2)(2\phi+1)}{(\phi+1)(2\phi-3)} \{P_{\phi-1}(\mu) - P_{\phi-3}(\mu)\} \tag{2.28}$$

Substituting (2.27) into (2.28), with $\phi = k$, leads to, for $k \geqq 3$,

$$c_k = \frac{2k-1}{k+1} (\alpha_0^2 - 4\alpha_1\alpha_{-1})^{1/2} \mu c_{k-1} - \frac{k-2}{k+1} (\alpha_0^2 - 4\alpha_1\alpha_{-1}) c_{k-2}. \tag{2.29}$$

This is a three point recurrence relation by means of which successive $c_k$'s can be evaluated since $c_1 = \alpha_1$, $c_2 = \alpha_0\alpha_1$. It can be simplified once a particular difference scheme has been selected. For instance, the Lax–Wendroff approximation of (2.20) gives $\alpha_1 = -\frac{1}{2}q(1-q)$, $\alpha_0 = 1 - q^2$ and $\alpha_{-1} = \frac{1}{2}q(1+q)$. It then follows that $c_k = c_k(\alpha_0^{1/2})$ and (2.29) becomes

$$c_k = \alpha_0 \left\{ \frac{2k-1}{k+1} c_{k-1} - \frac{k-2}{k+1} c_{k-2} \right\}, \quad k \geqq 3, \tag{2.30}$$

with $c_1$ and $c_2$ as above. Thus, for the right quarter plane problem, equation (2.19) gives

$$V_0^n = \sum_{k=1}^{n} c_k V_1^{n-k} \tag{2.31}$$

for each component $V$ corresponding to a negative eigenvalue $\lambda$. Note that a different $c_k$ sequence is required for each such component. Equation (2.30), or (2.29) in the general case, can be used to generate the $c_k$ coefficients on a "once and for all basis". The components of $\boldsymbol{v}$ would still be required to be stored back to the initial level. The stability requirements for the Lax–Wendroff scheme imply $|q| < 1$ so that $|\alpha_0^2| < 1$, $|\alpha_1| < 1$ resulting in $|c_1| < 1$, $|c_2| < 1$ and it easily follows from this and (2.30) that $|c_k| < 1$ for $k < 8$. By using the second theorem of Stieltjes to give a bound on $|P_k(\alpha_0^{1/2}) - P_{k+2}(\alpha_0^{1/2})|$ it follows that, for $k \geqq 8$, $|c_k| < 1$ provided $|\alpha_0^{1/2}| \leqq 0.9988$. The Stieltjes bound is too conservative for $|\alpha_0^{1/2}| > 0.9988$. However as $|\alpha_0^{1/2}| \to 1$ the right hand side of (2.27) tends to zero so that $|c_k| < 1$. In fact $c_k \to 0$ as $k \to \infty$ since $\alpha_0 < 1$ and $\lim_{k\to\infty} P_k(\alpha_0^{1/2}) = 0$ for $|\alpha_0^{1/2}| < 1$. Although $n$ in (2.31) can tend to infinity this last result implies that only a finite number of terms of the $c_k$ sequence are required, being as low as twelve for some values of $\alpha_0$.

For the left quarter plane problem a sequence of functions $\hat{c}_k$ satisfying (2.27), with $\alpha_1$ and $\alpha-1$ interchanged, and (2.30) is produced such that $\hat{c}_1=\alpha_{-1}$, $\hat{c}_2=\alpha_0\hat{c}_1$. A boundary condition is 'essential' for components with $\lambda>0$ and is given by

$$V_N^n=\sum_{k=1}^{n}\hat{c}_k V_{N-1}^{n-k}. \tag{2.32}$$

Given either $\lambda>0$ for the right or $\lambda<0$ for the left quarter plane problems a boundary condition is "non-essential". Consider $\boldsymbol{V}$ partitioned into $\boldsymbol{V}^{\mathrm{I}}$ for components with $\lambda<0$ and $\boldsymbol{V}^{\mathrm{II}}$ for components with $\lambda>0$. Suitable 'non-essential' boundary conditions for the right quarter plane problem are

$$\boldsymbol{V}_0^{\mathrm{II}}=S_0\boldsymbol{V}_0^{\mathrm{I}} \tag{2.33}$$

and for the left quarter plane problem are

$$\boldsymbol{V}_N^{\mathrm{I}}=S_1\boldsymbol{V}_N^{\mathrm{II}}, \tag{2.34}$$

where $S_0$, $S_1$ are constant rectangular matrices for which $\|S_0\|<1$, $\|S_1\|<1$.

The relationship between the left and right half-plane problems and between the $\hat{c}_k$ and $c_k$ sequences implies that for a particular component the condition at $x=N\,\Delta x$ with $\lambda>0$ for the Lax–Wendroff scheme is the same as would be obtained on $x=0$ if $\lambda$ had been negative. This suggests a procedure, which will not upset stability, for 'non-essential' boundary conditions at $x=0$ and $x=N\,\Delta x$, namely replace the eigenvalue by its negative and proceed as for an 'essential' condition at that boundary. The 'non-essential' boundary condition for each component can thus be given at $x=0$ by (2.31) with $\hat{c}_k$ replacing $c_k$ and at $x=N\,\Delta x$ by (2.32) with $\hat{c}_k$ replaced by $c_k$. The boundary conditions for the diagonalised system follow from this, finally giving for the one space variable problem

$$\boldsymbol{U}_0^n=\sum_{k=1}^{n}A_k\boldsymbol{U}_1^{n-k} \tag{2.35a}$$

$$\boldsymbol{U}_N^n=\sum_{k=1}^{n}\hat{A}_k\boldsymbol{U}_{N-1}^{n-k} \tag{2.35b}$$

where, for example, $A_k=XD_kX^{-1}$ with $D_k=\operatorname{diag}(c_k^{(1)},\ldots,c_k^{(b_1)},\hat{c}_k^{(b_1+1)},\ldots,\hat{c}_k^{(b)})$, the superscript indicating the component.

The stability of the strip problem follows from Section 2.3 and their manner of derivation ensures the boundary conditions place no further restriction on $q$. Normal mode analysis checks this and involves, for the right quarter plane problem, substituting $V_j^n=\hat{V}\kappa^jZ^n$ into both the Lax–Wendroff scheme and the boundary condition (2.31). Provided $q<0$ examination of the spectrum of $Z$ for $|\kappa|<1$ shows no solution exists for which $|Z|>1$ for $|\kappa|<1$. The corresponding result for the left quarter plane problem requires $q>0$. $q=0$ gives a stable situation since it leads to a constant solution problem. The results of Kreiss [**12**] therefore imply the stability of the strip problem. Thus the presence of the derived boundary conditions does not alter the stability requirements of the Lax–Wendroff scheme.

A typical strip problem is (1.1) with boundary conditions (2.35) and initial conditions

$$\left[1 \quad 1 \quad \frac{1}{\gamma(\gamma-1)M_\infty^2}\right],$$

https://doi.org/10.1017/S0013091500004053 Published online by Cambridge University Press

$M_\infty$ the freestream Mach number and $\gamma = 1.41$. The stability of the boundary conditions was tested by considering values 0.3(0.1)0.9 for $M_\infty$ with $\Delta x = 0.008$ and $\Delta t$ chosen to be approximately 90% of the maximum allowed by the Lax–Wendroff stability criterion, the two step Richtmyer version of the scheme being used at internal points. Each run was for 1000 time steps and only in the case of $M_\infty = 0.9$ was there any indication of instability. This was overcome by decreasing $\Delta t$ slightly and was probably caused by the fact that one of the $c_k$ sequences for $M_\infty = 0.9$ converged very slowly.

In the one dimensional problem the method of constructing the boundary approximation implies infinite accuracy. However all that is required is that they be second order accurate. This is so provided

$$\sum_{k=1}^{\infty} c_k = 1 \tag{2.36a}$$

$$q^{-1} + \sum_{k=1}^{\infty} k c_k = 0 \tag{2.36b}$$

$$1 + 2q \sum_{k=1}^{\infty} (k-1)c_k + q^2 \sum_{k=1}^{\infty} (k-1)^2 c_k = q^2. \tag{2.36c}$$

Using the generating function $g(t)$ these equations can be written in terms of $g(1)$, $g'(1)$ and $g''(1)$ and verified in an elementary manner using equation (2.23). Equations (2.36) prove useful in the extension of the one space variable results to higher dimensions.

A three point recurrence relation differing from (2.19) only in the coefficient of $c_{k-2}$ is obtained when a lower order term $K\boldsymbol{u}$ is added to (2.20). The results of this section also apply to the MacCormack schemes since they give the same $\alpha_1, \alpha_0, \alpha_{-1}$ as in the Lax–Wendroff scheme.

## 2.4. Modified boundary conditions

The $c_k$ sequence developed in Section 2.3 is such that $c_k \to 0$ as $k \to \infty$. However in some cases the convergence rate is very slow and it is desirable to speed this up using modifications which are still local in space but not in time.

Since

$$V_0^{n+1} = \sum_{k=1}^{n} c_k V_1^{n-k+1} \quad \text{and} \quad V_0^n = \sum_{k=1}^{n} c_k V_1^{n-k}$$

it follows that

$$V_0^{n+1} = \delta V_0^n + \sum_{k=1}^{n} (c_k - \delta c_{k-1}) V_1^{n-k+1} \quad \text{with} \quad c_k = 0 \quad \text{if} \quad k < 1.$$

There is no $\delta$ to ensure the rapid convergence of $c_k - \delta c_{k-1}$. A similar result occurs when two parameters are used except that when the $c_k$ sequence converges very slowly the coefficient of $V_1$ quickly becomes very small. The four parameter case, with $S_k = c_k - \delta_1 c_{k-1} - \delta_2 c_{k-2} - \delta_3 c_{k-3} - \delta_4 c_{k-4}$ is therefore considered and for a particular component of the right quarter plane problem gives

$$V_0^{n+1} = \delta_1 V_0^n + \delta_2 V_0^{n-1} + \delta_3 V_0^{n-2} + \delta_4 V_0^{n-3} + \sum_{k=1}^{n+1} S_k V_1^{n-k+1}. \tag{2.37}$$

https://doi.org/10.1017/S0013091500004053 Published online by Cambridge University Press

By using (2.30) $S_k$ can be expressed in terms of $c_k$, $c_{k-1}$ and $c_{k-2}$ and by letting $k \to \infty$ it follows that $\delta_1 = 4\alpha_0$, $\delta_2 = -2\alpha_0(1+2\alpha_0)$, $\delta_3 = 4\alpha_0^2$ and $\delta_4 = -\alpha_0^2$ with an error $E_k$ in $S_k$ given by

$$E_k = \frac{15\alpha_0}{(k+1)(k-3)(k-4)}\{-4c_{k-1} + [3+\alpha_0 - k(1-\alpha_0)]c_{k-2}\}.$$

This error tends to zero faster than the corresponding error for the single and two parameter cases in almost every situation. However for certain values of $\alpha_0$ more than fifty $c_k$ terms are required to ensure $|S_k| < 10^{-6}$. The situation is not improved by increasing the number of parameters to six since more $c_k$ terms are sometimes needed to ensure $|S_k| < 10^{-6}$.

A result similar to (2.37) can be obtained for the left quarter plane problem and the pair used as a basis for the boundary conditions for the strip problem of Section 2.3. These conditions were applied to that problem and similar results were obtained regarding stability.

The number of $c_k$ terms to be evaluated is still higher than was hoped for and, in an attempt to minimise this, conditions which are local in time but not in space are now considered.

## 3. Spacewise boundary conditions in one space variable

In the right quarter plane problem discussed in Section 2 the boundary value at $x = 0$ for each component at time level $(n+1)\,\Delta t$ can be obtained as a linear combination of the values at $x = 0$ for the previous time levels. A boundary condition is 'essential' for such a component when the associated eigenvalue is negative. By considering the characteristics from a point on $x = 0$ it is seen that the values at time level $n\,\Delta t$ for $x > 0$ are dependent on the values at $x = 0$ for $0 \leqq t < n\,\Delta t$. It would thus seem that the boundary condition at $x = 0$, $(n+1)\,\Delta t$ can be given as a linear combination of values at $n\,\Delta t$, $x \geqq 0$ without affecting the internal values too much provided sufficient terms are taken. This procedure cannot apply at $x = 0$ to components for which the eigenvalue is positive, the boundary value for such components having to be obtained independently.

### 3.1. The half plane problems

Equation (2.20) can be reduced to diagonal form in which the $k$th scalar equation of the system is

$$\frac{\partial v^{(k)}}{\partial t} + d^{(k)}\frac{\partial v^{(k)}}{\partial x} = 0, \tag{3.1}$$

$d^{(k)}$ being an eigenvalue of $E$. The problem is discretised with time and space increments $\Delta t$ and $\Delta x$ respectively and (3.1) is approximated by

$$V_j^{n+1} = \sum_{l=-1}^{1} q_l V_{j+l}^n. \tag{3.2}$$

Suitable initial conditions must be prescribed and the boundary condition at $x = 0$ is taken as, with $s \to \infty$,

$$V_0^{n+1} = \sum_{p=0}^{s} b_p V_p^n. \tag{3.3}$$

With $r = \Delta t/\Delta x$, $\alpha = \alpha_j \Delta x$ and $q = rd^{(k)}$ real, (3.1) has a solution proportional to

$$\exp(i\{-qn\alpha + j\alpha\}). \tag{3.4}$$

Substituting this into (3.2) and truncating for second order accuracy gives $q_l$, $l = -1, 0, 1$ appropriate to the one space variable Lax–Wendroff scheme. Stability of this implies $|q| \leqq 1$. Substituting (3.4) into (3.3) produces

$$e^{-iq\alpha} = \sum_{p=0}^{s} b_p e^{ip\alpha}. \tag{3.5}$$

This can be expanded in Taylor series, and powers of $\alpha$ equated to yield a system of linear equations for the $b$'s. The coefficient matrix of the system is a Vandermonde type and unless its dimension is small there will be instability. However writing the complex exponential in terms of a sine and a cosine and equating real and imaginary parts produces

$$\cos q\alpha = \sum_{p=0}^{\infty} b_p \cos p\alpha \quad \text{and} \quad \sin q\alpha = -\sum_{p=0}^{\infty} b_p \sin p\alpha.$$

Multiplying the first of these by $\cos \bar{\gamma}\alpha$, the second by $\sin \bar{\gamma}\alpha$, adding and integrating from 0 to $\pi$ gives

$$b_p = \begin{cases} \sin \dfrac{(p+q)\pi}{\pi(p+q)} = \dfrac{(-1)^p \sin \pi q}{\pi(p+q)} & \text{if} \quad p \geqq 0 \\ 0 \quad \text{if} \quad p < 0. \end{cases} \tag{3.6}$$

Thus $b_p \to 0$ as $p \to \infty$. A constant solution of (3.3) implies $\sum_{p=0}^{\infty} b_p = 1$ and from (3.6) $|b_p| < 1$. With the possible exception of the first, these terms alternate in sign but the rate at which $b_p \to 0$ is in general too slow for $|b_p|$ to be small enough to be ignored when $p$ is small.

The sign change between successive $b_p$ terms suggests considering a sequence of elements $\frac{1}{2}(b_p + b_{p+1})$, $p = 0, 1, \dots$ with $b_p = 0$ if $p < 0$. These elements tend to zero and, except for the first two, the terms alternate in sign. This suggests extending the averaging process and produces the boundary condition

$$V_0^{n+1} = \sum_{p=0}^{\infty} h_p V_p^n, \tag{3.7}$$

with, for $p = 0, 1, \dots,$

$$h_p = \tfrac{1}{32}(b_p + 5b_{p-1} + 10b_{p-2} + 10b_{p-3} + 5b_{p-4} + b_{p-5}) \tag{3.8}$$

and $b_p = 0$ for $p < 0$. (3.6) can be used to express $h_p$ in a more suitable form. Thus for $p \geqq 5$

$$h_p = \frac{-15(-1)^p \sin \pi q}{4\pi(p+q)(p+q-1)(p+q-2)(p+q-3)(p+q-4)(p+q-5)}, \tag{3.9}$$

https://doi.org/10.1017/S0013091500004053 Published online by Cambridge University Press

$h_p$, $0 \leqq p < 5$, being obtained from (3.8). It is clear that $|h_p| < 1$ and, after the first six positive terms, the signs of $h_p$ alternate. Moreover $h_p = O(p^{-6})$ so that they decay rapidly. The requisite number of terms for a given accuracy depends on the value of $q$, the choice $p = 14$ being adequate to satisfy $|h_p| < 10^{-6}$. Considering a constant solution of (3.7) gives

$$\sum_{p=0}^{\infty} h_p = 1. \tag{3.10a}$$

Applying the averaging process to the $h_p$'s gives terms which are $O(p^{-7})$. This results in greater accuracy for large $p$ but the multiplicative constant and the additional denominator term $(p + q - 6)$ ensure that the gain for small $p$ is minimal.

Equation (3.7) is therefore chosen as the basic boundary condition, stability being established for $q \leqq 0$ using normal mode analysis as in Section 2.3. The case $q > 0$ leads to instability, implying boundary conditions such as (2.33) must be used for such components. Equations (2.33) and (3.7), with at most fourteen values of $h_p$ stored, constitute suitable boundary conditions for the right quarter plane problem.

The left quarter plane problem is similar to the one described in Section 2.3 except that 'essential' boundary conditions for $q \geqq 0$ are given by

$$V_N^{n+1} = \sum_{p=0}^{-\infty} h_{p+N} V_{p+N}^n \tag{3.11}$$

with (2.34) providing suitable conditions for components with $q < 0$.

Second order accuracy of the boundary conditions (3.7) and (3.11) requires (3.10a) to be satisfied together with

$$\sum_{p=0}^{\infty} p h_p = -q \tag{3.10b}$$

$$\sum_{p=0}^{\infty} p^2 h_p = q^2 \tag{3.10c}$$

these being established using residue calculus.

An alternative for a 'non-essential' boundary condition is obtained by changing the sign of the eigenvalue and using the $h_p$ sequence generated from the corresponding $q$. Stability is not upset by this choice and a result similar to (2.35) is obtained for the stable quarter plane problems. The results of Kreiss **[12]** then imply the stability of the strip problem.

The stability of the boundary conditions was tested using the flow problem of Section 2.3 with boundary conditions (3.7), (3.11), (2.33) and (2.34) with $S_0 = [\frac{1}{2} \ \ \frac{1}{4}]^T$ and $S_1 = [\frac{1}{2} \ \ \frac{1}{4}]$. The same time and space increments were used as before and fourteen values of $h_p$ stored. There was no evidence of any instability after 1000 time steps.

## 4. Higher dimensional problems

### 4.1. Problem in two space variables

In two space variables (2.1) takes the form

$$\frac{\partial \boldsymbol{u}}{\partial t} + \frac{\partial F(\boldsymbol{u})}{\partial x} + \frac{\partial G(\boldsymbol{u})}{\partial y} = 0 \tag{4.1}$$

and has to be solved in the region $0 \leqq x \leqq a$, $0 \leqq y \leqq b$ with suitable initial and boundary conditions prescribed. The problem is discretised with $a = K_1 \Delta x$, $b = K_2 \Delta y$, $K_1$ and $K_2$ natural numbers, and (4.1) approximated by the Strang [**18**] scheme

$$\boldsymbol{U}^{n+1} = \tfrac{1}{2}(L_x L_y + L_y L_x)\boldsymbol{U}^n. \tag{4.2}$$

In this $L_x$ and $L_y$ are the operators associated with the one dimensional problems

$$\frac{\partial \boldsymbol{u}}{\partial t} + \frac{\partial F(\boldsymbol{u})}{\partial x} = 0 \quad \text{and} \quad \frac{\partial \boldsymbol{u}}{\partial t} + \frac{\partial G(\boldsymbol{u})}{\partial y} = 0$$

respectively. Equation (4.2) is implemented at all internal mesh points using the standard formulation of Morris and Gourlay [**4**]. This formulation uses the standard two step Richtmyer scheme for $L_y$ and $L_x$ with the same approximations used for them on boundaries on which $x$ and $y$ are respectively constant. Special schemes are required for the operators $L_x$ and $L_y$ on boundaries on which $x$ and $y$ are respectively constant. A 'time dependent' form for these is considered first.

The one dimensional nature of the basic operators in (4.2) and in the other Strang schemes allows the stability of the boundary problem to be treated in the same fashion as in Section 2.3. The approximations for $L_x\boldsymbol{U}$ at the boundaries $x = 0$, $x = K_1 \Delta x$ are of the form of (2.35) and are given by

$$\boldsymbol{U}^n_{0,l} = \sum_{k=1}^{n} A_k \boldsymbol{U}^{n-k}_{1,l} \tag{4.3a}$$

$$\boldsymbol{U}^n_{K_1,l} = \sum_{k=1}^{n} \hat{A}_k \boldsymbol{U}^{n-k}_{K_1-1,l}. \tag{4.3b}$$

The one dimensional 'y' problem has boundaries at $y = 0$, $y = K_2 \Delta y$ and the approximations used for $L_y\boldsymbol{U}$ at these boundaries are respectively

$$\boldsymbol{U}^n_{j,0} = \sum_{k=1}^{n} B_k \boldsymbol{U}^{n-k}_{j,1} \tag{4.4a}$$

$$\boldsymbol{U}^n_{j,K_2} = \sum_{k=1}^{n} \hat{B}_k \boldsymbol{U}^n_{j,K_2-1}. \tag{4.4b}$$

The results of Section 2 imply that the boundary approximations for $L_x$ and $L_y$ will be bounded and the stability of the initial boundary value problem in two space variables follows from the stability analysis of the basic Strang scheme.

Consistency and accuracy comparable with the internal scheme is also required for the boundary approximation. At $x = 0$ the approximations for $L_x\boldsymbol{U}^n$ and $L_y\boldsymbol{U}^n$ are respectively (4.3a) and the standard Lax–Wendroff approximation. By expanding these approximations in Taylor series and using the extension of (2.36) to a system it follows that

$$L_x L_y \boldsymbol{U}^n = O(\Delta t^2), \quad L_y L_x \boldsymbol{U}^n = O(\Delta t^2) \quad \text{with} \quad \tfrac{1}{2}(L_x L_y + L_y L_x)\boldsymbol{U}^n = O(\Delta t^3), \tag{4.5}$$

just as for the internal scheme. The same results hold for the boundary $x = K_1 \Delta x$ using (4.3b) and also for the boundaries $y = 0$, $y = K_2 \Delta y$ using the standard Lax–Wendroff approximation for $L_x\boldsymbol{U}^n$ and (4.4a), (4.4b) respectively for $L_y\boldsymbol{U}^n$. At the corner $x = 0$,

$y=0$ the boundary operators are (4.3a) and (4.4a). Expanding these in Taylor series and using (2.36) establishes (4.5) with similar results for the other corner points. The boundary approximations are therefore consistent and of the same order of accuracy as the internal scheme.

The boundary conditions given above extend easily to higher dimensions and can also be adapted for use with the Strang [**19**] scheme $\boldsymbol{U}^{n+1}=L_{x/2}L_yL_{x/2}$ in its compounded form. Stability follows readily and the accuracy of the boundary and internal approximations are comparable. The main difficulty would lie in feeding in the boundary information.

Equations (3.3) and (2.33) can be used to approximate $L_x\boldsymbol{U}^n$ at $x=0$, with similar approximations used at the other boundaries. Stability follows readily but the required accuracy, which need only be first order as is shown in Section 4.2, cannot be established. In fact consistency only holds for a special relation between the elements of $\boldsymbol{u}$. This is also true of the conditions given by Gottlieb and Turkel [**7**]. The 'space dependent' condition can be employed with an $h_p$ sequence defined on each boundary for each component. The stability of such a scheme is easily established and the required accuracy obtained from equations (3.10).

Analysis of the accuracy of the approximations follows the same pattern as given above.

### 4.2. Minimum accuracy for boundary approximations

Equation (4.1) is to be considered with initial conditions $\boldsymbol{u}_0(x, y)$ and homogeneous boundary conditions which are either specified or are derivable in some way. $L_x$ is defined so that

$$L_x\boldsymbol{u}_0(x, y)=\boldsymbol{V}(x, y, \Delta t)+O(\Delta t^3)$$

with $\boldsymbol{V}(x, y, 0)=\boldsymbol{u}_0(x, y)$ and where $\boldsymbol{V}(x, y, \Delta t)$ is the solution of $\boldsymbol{V}_t+E\boldsymbol{V}_x=0$ with homogeneous boundary conditions applying. $M_x$ is taken as the exact solution step operator for this, that is $M_x\boldsymbol{u}_0(x, y)=\boldsymbol{V}(x, y, \Delta t)$, $M_y$ and $M$ being similarly defined for the corresponding 'y problem' and the two dimensional problem.

The numerical scheme (4.2) will have the required accuracy provided

$$M\boldsymbol{u}_0(x, y)=\tfrac{1}{2}(M_xM_y+M_yM_x)\boldsymbol{u}_0(x, y)+O(\Delta t^3).$$

Since the properties of the one dimensional operators are well established it suffices to consider $M_x$, $M_y$ and $M$, rather than the numerical operators $L_x$ and $L_y$. The results of Strang [**18**] ensure only mesh points for which $M\boldsymbol{u}_0(x, y)$ is dependent on one or more boundary conditions need be investigated.

Assuming no explicit dependence on $t$ it follows from the semi group property $M(t_1+t_2)=M(t_1)M(t_2)$ that

$$M(t)=\left\{M\left(\frac{t}{n}\right)\right\}^n, \tag{4.6}$$

with similar results holding for $M_x$ and $M_y$. Let

$$M(t)=I+tm_1+t^2m_2+\dots, \tag{4.7}$$

similar definitions applying to $M_x$ and $M_y$ with $m$ replaced by $m^{(x)}$ and $m^{(y)}$.

$$m_2=\tfrac{1}{2}m_1'm_1, \tag{4.8}$$

where $m_1'$ is a Fréchet derivative, this being also satisfied by $m^{(x)}$ and $m^{(y)}$. (4.6) can thus be expanded and powers of $t$ equated to give the operators $m_2, m_3, \ldots$ in terms of $m_1$. The operators $M_xM_y$ and $M_yM_x$ are formed and (4.8), for $m_2^{(x)}$ and $m_2^{(y)}$, used to give

$$\tfrac{1}{2}(M_xM_y+M_yM_x)=I+t(m_1^{(x)}+m_1^{(y)})+\tfrac{1}{2}t^2(m_1'^{(x)}+m_1'^{(y)})(m_1^{(x)}+m_1^{(y)})+O(t^3) \tag{4.9}$$

so that if $m=m_1^{(x)}+m_1^{(y)}$ then

$$\tfrac{1}{2}(M_xM_y+M_yM_x)=M+O(t^3).$$

Thus it is only necessary to ensure that the first order term is correct. This result is also true of the Strang scheme $\boldsymbol{U}^{n+1}=L_{x/2}L_yL_{x/2}\boldsymbol{U}^n$ in which the exact operator $M_{x/2}$ corresponding to $L_{x/2}$ is defined in a manner similar to that for $M_x$. Equation (4.9) is true for the internal points and this indicates that the conditions used on the boundary need only be first order accurate, agreeing with the one dimensional result of Gustafsson [**8**].

## 5. Conclusions

In Section 2 the solution of the equation $\dfrac{\partial \boldsymbol{u}}{\partial t}+\dfrac{\partial \boldsymbol{F}(\boldsymbol{u})}{\partial x}=0$ was considered over an infinite domain. The problem was discretised and artificial boundaries introduced in order that a numerical solution could be effected. 'Time dependent' conditions which were local in space but non local in time were obtained for these boundaries. Although they were stable certain components required more storage of past information than was desirable. This was overcome by developing, for these artificial boundaries, conditions which were local in time but non local in space. The stable 'space dependent' conditions involved no storage of past information.

Both the 'time dependent' and 'space dependent' boundary approximations in one space variable were adapted for use with higher dimensional problems. In both cases stable conditions of the required accuracy were obtained for the two space variable problem.

Finally it was shown for the Strang schemes that it is only necessary to ensure that the boundary approximations are of first order accuracy. This applies to problems with the boundary conditions specified as well as to problems involving artificial boundaries.

**Acknowledgement.** Part of this work was carried out by the author at the University of Strathclyde under the direction of Professor D. S. Butler. The author would like to thank Professor Butler for his advice and encouragement during the course of that research and also the referee of this paper for the useful suggestions he made.

REFERENCES

**1.** A. M. Burns, *Math. Comp.* **32** (1978), 707–724.

**2.** B. Engquist and A. Majda, *Math. Comp.* **31** (1977), 629–651.

**3.** B. Engquist and A. Majda, *Comm. Pure App.* **32** (1979), 313–357.

**4.** A. R. Gourlay, and J. Ll. Morris, *Math. Comp.* **22** (1968), 549–555.

**5.** A. R. Gourlay and J. Ll. Morris, *Math. Comp.* **22** (1968), 715–720.

**6.** A. R. Gourlay and J. Ll. Morris, *J. Comp. Phys.* **5** (1970), 229–243.

**7.** D. Gottlieb and E. Turkel, *J. Comp. Phys.* **26** (1978), 181–196.

**8.** B. Gustafsson, *Math. Comp.* **29** (1975), 396–406.

**9.** B. Gustafsson, H. O. Kreiss and A. Sundström, *Math. Comp.* **26** (1972), 649–686.

**10.** B. Gustafsson and H. O. Kreiss, *J. Comp. Phys.* **30** (1979), 333–351.

**11.** H. O. Kreiss, On Difference Approximations, Sym. Uni. Maryland (1965).

**12.** H. O. Kreiss, Difference Approximations for initial boundary value problems for hyperbolic differential equations, Sym. Mad. Wis. (1966).

**13.** H. O. Kreiss, *Math. Comp.* **22** (1968), 703–714.

**14.** J. Ll. Morris and G. McGuire, J.I.M.A. **10** (1972), 150–165.

**15.** J. Ll. Morris and G. McGuire, J.I.M.A. **67** (1976), 53–67.

**16.** S. Osher, *Trans. Amer. Math. Soc.* **137** (1969), 177–201.

**17.** S. Osher, *Math. Comp.* **32** (1969), 335–340.

**18.** G. Strang, *Arch. ration. Mech. Anal.* **12** (1963), 392–402.

**19.** G. Strang, *SIAM J. Numer. Anal.* **5** (1968), 506–517.

Department of Mathematics and Computing
Paisley College of Technology