Stability and convergence of linear parabolic problems on nonuniform grids
Mattheij, R.M.M.; Smooke, M.D.

Published: 01/01/1989

**Document Version**
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the author’s version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

**Citation for published version (APA):**
STABILITY AND CONVERGENCE
OF LINEAR PARABOLIC PROBLEMS
ON NONUNIFORM GRIDS
by
R.M.M. Mattheij
M.D. Smooke
Many problems in the physical sciences can be reduced to the solution of a system of time-dependent partial differential equations. Of particular interest to us are problems in combustion and heat and mass transfer, i.e., hydrocarbon ignition and catalytic combustion. The governing equations in these applications can be formulated as a system of parabolic mixed initial-boundary value problems. The numerical stiffness that results from solving these problems on a discrete mesh combined with the inherent stiffness of the disparate decay rates of the various chemical species necessitate the use of implicit time differencing methods. The problems also produce solutions that contain regions of high spatial activity, i.e., sharp peaks and steep fronts. Although an equispaced mesh could be used in the calculations, it is often more efficient to employ a nonuniform adaptive grid in the anticipation that the high activity regions will be better resolved. Important questions in such studies are the effects that adaptive time and space steps (both fixed and variable numbers of points) have on the stability and convergence of the parabolic solver. We investigate these issues in this paper for a class of linear mixed initial-boundary value problems.
1. Introduction

A special class of difference methods for solving partial differential equations occurs when the derivatives with respect to only one of the independent variables are replaced by difference quotients. This is what has traditionally been called the method of lines. The method replaces a given partial differential equation by a system of ordinary or partial differential equations with a smaller number of independent variables. The method can be viewed as lying somewhere between analytic and fully discrete methods. The idea of replacing one of the independent variables in a partial differential equation by an appropriate difference expression such that the original problem is reduced to a new problem—still differential but of lower dimension—appears to have been applied first by Rothe [7] to parabolic partial differential equations.

Liskovets makes the important observation that in applied problems (either hyperbolic or parabolic in nature) most of the effort put into the method of lines has been in what is termed the “longitudinal” method [3]. The method used by Rothe is called the “transverse” method. In applying the longitudinal method the spatial derivatives are discretized and the partial differential equation is reduced to a system of initial value problems. With the advent of high speed digital computers and the development of variable step algorithms for the integration of systems of ordinary differential equations, the longitudinal method of lines has been used with increasing regularity for the solution of complicated partial differential equations (see e.g. [4,5]).

The longitudinal method is particularly well suited to problems in which the spatial variation of the solution does not contain any sharp peaks or steep fronts as a function of time. As a result, an equispaced or mildly nonuniform grid can be chosen a-priori and then used for all subsequent time levels. However, for problems in which the spatial variation of the dependent solution components changes rapidly, such an approach may not be efficient. Instead, we want to be able to concentrate grid points in regions where the dependent solution components vary most rapidly. By discretizing the time derivatives of the original mixed initial-boundary value problem, a system of two-point boundary value problems is obtained. These boundary value problems can then be solved with adaptive two-point boundary value methods (see e.g. [1,8]). It is interesting to note that, although

* Work supported in part by the Air Force Office of Scientific Research and the NATO Scientific Affairs Division
the transverse method of lines was the original formulation of the method, there have been relatively few applications of it in practice. The most obvious reason for this is the fact that the development of adaptive two-point boundary value software has generally lagged the development of variable step initial value problem software by several years. We believe, however, that there are classes of problems for which the transverse method of lines is particularly well suited.

We point out that many problems in the physical sciences can often be reduced to the solution of a system of time-dependent partial differential equations. Of particular interest to us are problems in combustion and heat and mass transfer, i.e., hydrocarbon ignition and catalytic combustion. The governing equations in these applications can be formulated as a system of coupled nonlinear parabolic mixed initial-boundary value problems [12]. The numerical stiffness that results from solving these problems on a discrete mesh combined with the inherent stiffness of the disparate decay rates of the various chemical species necessitate the use of implicit time differencing methods. The problems also produce solutions that contain regions of high spatial activity, i.e., sharp peaks and steep fronts. Although an equispaced mesh could be used in the calculations, it is often more efficient to employ a nonuniform adaptive grid in the anticipation that the high activity regions will be better resolved. It is the high activity regions, however, that force a decrease in the size of the time step to satisfy accuracy considerations. For problems with very sharp solution profiles on highly refined grids, variable order initial value problem codes often reduce the order of the algorithm to at most one or two. Accurate and efficient solution of this class of problems necessitates combining the best features of both the longitudinal and the transverse method of lines—adaptive time and spatial steps, respectively.

There exists an extensive literature devoted to convergence and stability of difference methods for parabolic problems (see, e.g., [10,11,14]). A number of more recent papers have considered the solution of time-dependent partial differential equations using adaptive spatial grids [2,6,13]. In most of these papers a fixed number of adaptive points are used at each time level to satisfy an equidistribution criterion. In the paper by Smooke and Koszykowski [9], however, the number as well as the location of the grid points could change from one time level to another. Important questions in such studies are the effects that adaptive time and space steps (both fixed and variable numbers of points) have on the stability and convergence of the parabolic solver. The investigation of the convergence properties of finite difference solutions to such problems on a uniform grid can be obtained by standard matrix stability techniques. We will review these methods and extend them to more general cases paying special attention to infinite time intervals.

In this paper we investigate the stability and convergence of variable step (space and time) algorithms in the solution of one-dimensional, linear, parabolic, mixed initial-boundary value problems. Although nonlinear problems are more important in practice, they are more demanding and so we have chosen to treat them separately in a subsequent
paper. The manuscript is organized as follows. In the next section we present some preliminary results and notation. In Section 3 we investigate the stability and convergence of autonomous problems on a fixed nonuniform grid. In Section 4 we develop the same analysis for nonautonomous problems. The case where both the time steps and the spatial grids can change is examined in Section 5. Finally, in Section 6 we present some numerical results.

2. Preliminaries

In this paper we consider the one-dimensional parabolic problem

\[ u_t = a(x,t)u_{xx} + b(x,t)u_x + c(x,t)u + f(x,t), \quad [x,t] \in (0,1) \times (0,T), \]

\[ u(0,t) = u_0, \]

\[ u(1,t) = u_1, \]

\[ u(x,0) = g(x), \]

when nonuniform spatial grids are employed. The quantities \( u_0 \) and \( u_1 \) are constants and \( (\cdot)_x \) and \( (\cdot)_t \) denote differentiation with respect to \( x \) and \( t \), respectively. The quantities \( a, b, c, f \), and \( g \) are assumed to be sufficiently smooth functions of the independent variables \( x \) and \( t \). We point out that combustion problems of physical interest are almost always nonlinear. Problems similar in form to (2.1) can result, however, from a linearization of the original equations.

In the numerical solution of (2.1) we start from the initial point \( t^0 = 0 \) and obtain a solution of (2.1) at time levels \( 0 = t^0 < t^1 < t^2 < \ldots < t^J \), where the time step \( k_j = t^{j+1} - t^j \). The spatial discretization is performed on the mesh \( \mathcal{M}_j \),

\[ \mathcal{M}_j = \{0 = x_0^j < x_1^j < \ldots < x_{M_j}^j = 1\}, \]

where \( h_i^j = x_i^j - x_{i-1}^j, \quad i = 1,2,\ldots,M_j \), and \( h^j = \max_{1 \leq i \leq M_j} h_i^j \). The time level superscript on the mesh is used to account for the fact that, in general, the number and/or the location of the grid points can differ from one time level to another. We approximate the spatial derivatives in (2.1) using finite difference expressions. In particular, omitting the discretization error, we write

\[ u_{xx}(x_i^j, t_j) \approx \left( \frac{2}{h_i^j + h_{i+1}^j} \right) \left( \frac{u_{i+1,j} - u_{i,j}}{h_i^j} - \frac{u_{i,j} - u_{i-1,j}}{h_{i+1}^j} \right), \]

and

\[ u_x(x_i^j, t_j) \approx \frac{u_{i+1,j} - u_{i-1,j}}{h_i^j + h_{i+1}^j}, \]

where we define \( u_{i,j} = u(x_i^j, t^j), i = 0,1,\ldots,M_j \), and \( j = 0,1,2,\ldots,J \). By replacing the continuous differential operators in (2.1) by expressions similar to those in (2.3) and
(2.4), we convert the problem of finding an analytic solution of (2.1) to one of finding an approximation to this solution at each point \((x'_i, t'_i)\).

If we set \(a(x, t) = 1, b(x, t) = c(x, t) = f(x, t) = 0\) and employ the spatial differences in (2.3) and (2.4) along with a single-step time discretization procedure, we obtain the linear system

\[
[I + \theta k_j A_{j+1}] U_{j+1} = [I - (1 - \theta) k_j A_j] U_j,
\]  

at each time level where \(0 \leq \theta \leq 1\). The matrix \(A_j\) can be written in the form

\[
A_j = \begin{bmatrix}
\gamma^j_1 & \beta^j_1 \\
\alpha^j_2 & \gamma^j_2 & \beta^j_2 \\
\alpha^j_3 & \gamma^j_3 & \beta^j_3 & \quad \\
\alpha^j_{M-2} & \beta^j_{M-2} & \gamma^j_{M-1} & \beta^j_{M-1} \\
\end{bmatrix},
\]

where

\[
\beta^j_i = -\frac{2}{(h^i_j + h^i_{j+1})h^i_{j+1}},
\]

\[
\alpha^j_i = \frac{h^i_{j+1}}{h^i_j} \beta^j_i,
\]

\[
\gamma^j_i = \frac{2}{h^i_j h^i_{j+1}}.
\]

The vector \(U_j\) is given by

\[
U_j = \begin{bmatrix}
u_{1,j} \\
u_{2,j} \\
\vdots \\
u_{M-2,j} \\
u_{M-1,j}
\end{bmatrix}.
\]

Of particular interest to us is the case where \(\theta = 1\) (i.e., backward Euler), but in our analysis we will pursue the more general case \(\theta \leq 1\).

If we add the lower order terms \((bu_x + cu)\) to the (2.5), we have

\[
[I + \theta k_j B_{j+1}] U_{j+1} = [I - (1 - \theta) k_j B_j] U_j,
\]  

where \(B_j = A_j + C_j\) with

\[
C_j = \begin{bmatrix}
c^j_1 & \delta^j_1 \\
\epsilon^j_2 & c^j_2 & \delta^j_2 \\
\vdots & \vdots & \vdots \\
\epsilon^j_{M-2} & c^j_{M-2} & \delta^j_{M-2} & \quad \\
\epsilon^j_{M-1} & c^j_{M-1} & \delta^j_{M-1}
\end{bmatrix},
\]
where

\[ c'_i = c(x'_i, t'), \quad b'_i = b(x'_i, t'), \]  
\[ \delta^i_j = \frac{b'_i}{(h_i' + h_{i+1}')}, \]  
\[ \varepsilon^i_j = -\frac{b'_i}{(h_i' + h_{i+1}')}, \]

(2.13) \hspace{1cm} (2.14) \hspace{1cm} (2.15)

Since the inhomogeneous function \( f \) does not affect the stability of the finite difference schemes under consideration, we find (2.11) to be our most general form. We will investigate the size of the errors due to discretization in both space and time. This subject is not new and has been studied before in various forms. Our approach differs from these previous investigations in that we ultimately combine variable spatial grids and variable time steps, thus setting the framework for difficult nonlinear problems. In the discussion that follows we denote \( \| \cdot \| = \| \cdot \|_\infty \). In addition, \( \prod_{j=1}^m M_j = M_m \ldots M_1 \), if \( m \geq l \) and \( \prod_{j=1}^m M_j = I \), if \( m < l \) and similarly \( \sum_{j=1}^m M_j = M_m + \ldots + M_1 \) if \( m \geq l \) and \( \sum_{j=1}^m M_j = 0 \) if \( m < l \).

The convergence analysis contained in the remainder of this paper will make use of the following definitions.

**Definition 2.1** We call the numerical method stable if for some constants \( \delta_1, \delta_2, C, T \) and \( h_i' \leq \delta_1, k_j \leq \delta_2 \), we have

\[ \| \prod_{j=1}^m (I + \theta k_j B_{j+1})^{-1}(I - (1 - \theta)k_j B_j) \| \leq C, \]  
\[ \text{for all } l, m \text{ and with } \sum k_j \leq T. \] The method is called uniformly stable if (2.16) holds for arbitrarily large \( T \) and \( \delta_2 \).

Stability relates to convergence results on finite intervals. An important problem type occurs when, for example, one tries to approximate a steady-state, i.e., \( \partial u / \partial t = 0 \), in as few time steps as possible. For this class of problems, effectively on infinite intervals, we need the concept of uniform stability.

**Definition 2.2** We let \( U_j^\Delta \) be the exact solution defined on the grid \( M^j \) at time level \( t^j \). We call the method convergent if

\[ \lim_{h_j \to 0} \| U_j^\Delta - U_j^i \| = 0, \text{ for all } j. \]

(2.17)

**Definition 2.3** We let \( U_m^\Delta \) be the exact solution defined on the grid \( M^j \) at time level \( t^j \). We call the method asymptotically convergent if

\[ \lim_{h_j \to 0} \| U_m^\Delta - U_m^i \| = 0, \]

(2.18)
where $T = \sum_{j=1}^{m} k_j$ and $k_j$ is not $o(1)$ as $j \to \infty$.

We point out that asymptotic convergence is meaningful only if $\lim_{t \to \infty} u(x, t)$ exists, i.e., we have a steady-state solution.

3. Convergence Analysis for Autonomous Problems with Fixed Nonuniform Grids

In this section we consider the case where the number and location of the grid points do not change from one time level to another, i.e., $h_i = h_{i+1}$ and $M_i = M_{i+1}$ and, to start with, the time step is also kept constant. We realize, of course, that this is an unrealistic assumption for problems in which the regions of high activity move as a function of time and for problems in which dissipation enables large time steps to be taken. Therefore, we study separately the case where $T$ can be arbitrarily large. Clearly this type of analysis is meaningful only in cases where the coefficients of (2.1) do not depend on $t$ explicitly (which, in turn, has interesting ramifications for nonlinear autonomous problems). Nevertheless, it illustrates the fundamental concepts of our analysis. To simplify our notation in the discussion that follows, we omit the time level superscript. We begin with the following result.

Lemma 3.1 For the matrix $B_j = B$ (cf. (2.11) with $k_j = k$) the eigenvalues are contained in the interval $[\min, kc_i, k\lambda]$ where

$$\lambda = \max_i, c_i + \frac{4}{h_i h_{i+1}} + \frac{2b_i}{h_i + h_{i+1}}.$$ (3.1)

The proof follows from a Gershgorin argument.

For stability we must have

$$\| [(I + \theta k B)^{-1} (I - (1 - \theta) k B)]^m \| \leq C,$$ (3.2)

for all $m$ and some constant $C$. This is a classical problem the result of which is summarized in the following theorem.

Theorem 3.1 Let $k_j = k$. If $\frac{1}{2} \leq \theta \leq 1$, then (2.11) is stable and, if $0 \leq \theta < \frac{1}{2}$, then (2.11) is stable only if $\theta < \frac{1}{2} - \frac{1}{k\lambda}$, where $\lambda = \max_i, c_i + \frac{1}{h_i h_{i+1}} + \frac{2b_i}{h_i h_{i+1}}$.

Proof: There exists a matrix $T$ that diagonalizes $B$, i.e., $B = T^{-1} D T$ for some diagonal matrix $D$. This follows from the fact that $B$ is similar to a symmetric matrix which can be seen by defining a similarity transformation $S^{-1} BS$ where $S$ is a diagonal matrix. Since the matrix $[I + \theta k B]$ has the same eigenvectors as $[I - (1 - \theta) k B]$, it is sufficient to show that $|1 + (1 - \theta) k \lambda |(1 + \theta k \lambda)| \leq 1$ for all eigenvalues of $B$. To this end we define the
function \( g(\mu) \) where

\[
g(\mu) = \frac{1 - (1 - \theta)\mu}{1 + \theta \mu}.
\] (3.3)

If \( g \) is plotted versus \( \mu \), one can show that \( g(\mu) \leq 1 \) for \( \mu \leq 0 \) and that \( g(\mu) \geq -1 \) for \( \theta \geq \frac{1}{2} \) and \( g(\mu) < -1 \) for \( \theta < \frac{1}{2} \). Since for \( \theta < \frac{1}{2} \), \( g(\mu) = -1 \) when \( \mu = \frac{1-\theta}{\theta} \), the stability requirement for \( 0 \leq \theta < \frac{1}{2} \) follows from Lemma 3.1. 

If we assume sufficient smoothness properties of the solution, then the local error due to the spatial discretization at time level \( t^i \), i.e., \( T_j = (T_{1j}, \ldots, T_{Mj})^t \) has a local estimate \( T'_j = O(h_j^t(h_{i+1}^j - h_i^j)) \). Hence, we can write

\[
dU^\Delta_j \quad = \quad BU_j + T_j,
\] (3.4)

which becomes after time discretization

\[
|I + \theta kB|U_{j+1} = |I - (1 - \theta)kB|U_j + k(T_j + \tilde{T}_j),
\] (3.5)

where \( \tilde{T}_j = (\tilde{T}_1^j, \ldots, \tilde{T}_M^j)^t \) is the local error due to the time discretization. We have \( \tilde{T}_j = O((\frac{1}{2} - \theta)k) \), \( \theta \neq \frac{1}{2} \) and \( \tilde{T}_j = O(\frac{k^2}{2}) \), \( \theta = \frac{1}{2} \). Using Theorem 3.1 we find

**Theorem 3.2** Let (2.1) be autonomous. The method in (2.11) is convergent for \( h^j \) and \( k \) small enough, i.e.,

\[
\| U_j^\Delta - U_j \| = O(\max, h_j(h_{i+1} - h_i)) + O(k),
\] if \( \theta \neq 1/2 \) and

\[
\| U_j^\Delta - U_j \| = O(\max, h_j(h_{i+1} - h_i)) + O(k^2),
\] if \( \theta = 1/2 \).

**Proof:** Upon denoting \( D = |I + \theta kB|^{-1}[I - (1 - \theta)kB] \), we can write

\[
\| U_j^\Delta - U_j \| = \| \sum_{i=1}^j D^{i-1} |I + \theta kB|^{-1} k(T_i + \tilde{T}_i) \|. \] (3.8)

For \( h^j \) and \( k \) small enough, we have stability and hence

\[
\| U_j^\Delta - U_j \| \leq jk \max, (T_i + \tilde{T}_i). \] (3.9)

Of course, posing the problem in this way is not very practical since we are not interested usually in taking exceedingly small steps as \( t \) increases; the actual stepsize restrictions on \( h^j \) for \( \theta < \frac{1}{2} \) are given in Theorem 3.1. It is well known that (2.11) enables one to increase the steplength as the time coordinate increases. Although Theorem 3.1 was proven for constant step size, it is simple to see that a similar result also holds when we have a
variable time step. In this case the matrices $B_j$ are still uniformly diagonalizable. Hence, we conclude that we really need $\theta > 1/2$. Since dissipativity is a crucial notion, we have

**Theorem 3.3** Let (2.1) be autonomous and $\theta > 1/2$. The method in (2.11) is uniformly stable, more precisely, upon denoting the product

$$D(l,j) = \prod_{k=1}^{j} (I + \theta k, B)^{-1} (I - (1 - \theta) k, B),$$

(3.10)

there exists a constant $C$ independent of $l, j$ such that

$$\| D(l,j) \| \leq C \prod_{k=1}^{j} \max \left( \frac{1 - \theta}{\theta}, \frac{1}{1 + \theta k, \lambda} \right),$$

(3.11)

where $\lambda$ is defined by equation (3.1).

**Proof:** Consider the quantity $g$ defined in (3.3). We can derive the fact that $g(\mu) \geq \frac{\mu - 1}{\mu}$. Moreover, for $\mu < \frac{1}{1 + \rho}$, $g(\mu) \leq \frac{1}{1 + \rho \mu}$. If we substitute the bound for the smallest eigenvalue of $B$ and use the monotonicity of $g$ for $\mu > 0$, we obtain the desired result.

One should note that the bound for $\| D(l,j) \|$ in Theorem 3.3 is sharp as $k_* \to \infty$ for $\theta < 1$ and $\theta = 1$ separately. In particular, for $\theta < 1$ we obtain a bound equal to $((1 - \theta)/\theta)^{j-1}$ and for $\theta = 1$ we obtain $\prod_{k=1}^{j} (1/(1 + k, \lambda))$. We use these results to derive results for infinite interval problems. We have

**Theorem 3.4** Let (2.1) be autonomous and assume $\theta > \frac{1}{2}$. If $\theta k, \lambda$ is not $o(1)$ as $s \to \infty$, then there exists a constant $C(\theta)$ such that

$$\| U_j^\Delta - U_j \| \leq C(\theta) \| T_j \| + \| \tilde{T}_j \|,$$

(3.12)

**Proof:** If $j$ is small, say $j \leq \hat{j}$, then the result follows from the proof of Theorem 3.2. Let $\hat{j}$ be such that $\| D(l,j) \| \leq C |\rho(\theta)|^{j-1}$, for $l \geq \hat{j}$ where $\rho(\theta)$ is sufficiently smaller than one. If we denote the global error at $t^\star$ by $S_i$, then for some constant $\hat{C}$

$$\| U_j^\Delta - U_j \| = \| \sum_{l=1}^{j} D(l,j) (I + \theta k_l, B)^{-1} k_l (T_l + \tilde{T}_l) + D(\hat{j},j) S_i \| \leq$$

$$\hat{C} \sum_{l=1}^{j} \rho^{j-l} \frac{k_l}{1 + \theta k_l \lambda} (\| T_l \| + \| \tilde{T}_l \|) + \rho^{j-\hat{j}} \| S_i \|,$$

(3.13)

Because of the exponential factor $\rho^{j-\hat{j}}$ the global error is essentially of the order of the local error.

**Remark 3.1** It follows from the proof of Theorem 3.4 that optimal damping effects are achieved for $\theta = 1$. 9
In addition, we have

**Corollary 3.1** Let (2.1) be autonomous and \( \theta > \frac{1}{2} \). If we assume that \( \lim_{t \to \infty} T_t = 0 \), then \( k \) may grow unbounded and there exists a constant \( C \) independent of \( j \) such that

\[
\| U_j^\alpha - U_j \| \leq C \| \tilde{T}_j \|, \quad (3.14)
\]

i.e., the method is asymptotically convergent.

4. Convergence Analysis for Nonautonomous Problems on Nonuniform Grids

The results derived in Theorems 3.2 and 3.4 can be generalized to problems with separable coefficients, i.e., the coefficients \( a, b \) and \( c \) in (2.1) have a common factor, say \( d(t) \), such that

\[
\begin{align*}
a(x, t) &= d(t)a(x), \\
b(x, t) &= d(t)b(x), \\
c(x, t) &= d(t)c(x).
\end{align*}
\]

(4.1a) (4.1b) (4.1c)

Our simplifying assumption \( a = 1 \) (which can always be achieved by properly rescaling the problem) must be replaced by \( \bar{a} = 1 \). A further rescaling of the time variable reduces the problem to the one considered before. A more elegant approach is to replace the time step \( k_j \) by \( \bar{k}_j = k_j d(t_j) \) and to then apply directly Theorems 3.2 and 3.4 to obtain the appropriate convergence result. We point out that this result depends upon having \( d(t) \geq \rho > 0 \).

If (2.1) is not separable, the arguments on which the stability and the convergence analysis in Section 3 were based fail to work. In particular, we no longer have similar matrices \( B_j \) and so an eigenvalue analysis is not meaningful in general. Therefore we proceed in another way obtaining bounds that are generally less sharp. We estimate both factors arising in (2.5) separately. We have

**Property 4.1** Assume \( h_i^t, h_{i+1}^t < b_i^t/2 \) and \( |(1 - \theta)k_j(2h_i^t + h_{i+1}^t) + c_i^t| < 1 \), then

\[
\| I - (1 - \theta)k_j B_j \| \leq 1 + (1 - \theta)k_j \max_i c_i^t, \quad (4.2)
\]

**Proof:** In the \( i \)th row, the codiagonal elements of \( B_j \) are given by

\[
k_j(1 - \theta) \left[ \frac{2}{(h_i^t + h_{i+1}^t)h_i^t} - \frac{b_i^t}{(h_i^t + h_{i+1}^t)} \right]
\]

and

\[
k_j(1 - \theta) \left[ \frac{2}{(h_i^t + h_{i+1}^t)h_{i+1}^t} + \frac{b_i^t}{(h_i^t + h_{i+1}^t)} \right]
\]

10
which are positive if \(|b'_i| > 2/h'_i, 2/h'_{i+1}\). The corresponding diagonal element is given by
\[
1 - \frac{2(1 - \theta)k_j}{h'_i h'_{i+1}} - (1 - \theta)k_j c'_i
\]
which is positive if
\[
(1 - \theta)k_j \left[ \frac{2}{h'_i + h'_{i+1}} + c'_i \right] < 1. \tag{4.3}
\]
The largest absolute row sum is then equal to \(1 + (1 - \theta)\max c'_i\).

In the discussion that follows we denote \(glb(A) = \min_{\|x\|=1} \|Ax\|\). We have

**Property 4.2** Under the assumptions of Property 4.1, we have
\[
\text{glb}(I + \theta k_j B_{j+1}) \geq 1 + \theta k_j \max \left[ \min c'_i, c'_i, c'_i + \frac{2}{(h'_i + h'_{i+1})h''_i h''_{i+1} + \frac{2}{(h''_{i-1} + h''_{i-2})h''_{i-1}}} \right], \tag{4.4}
\]

**Proof:** There exists some vector \(v\) of which at least one maximal component equals one such that \(\| (I + \theta k_j B_{j+1})v \| = \text{glb} (I + \theta k_j B_{j+1})\). Let this component be denoted by \(v_i\) and consider the ith element of \((I + \theta k_j B_{j+1})v\). If \(v_{i-1}\) and \(v_{i+1}\) were not both equal to one, then this element is always larger than \(c'_i + 1\) and so \(\| (I + \theta k_j B_{j+1})v \|\). Hence the minimum value of \(\| (I + \theta k_j B_{j+1})v \|\) is achieved for all coordinates equal to one.

If we combine the results from Properties 4.1 and 4.2, we find

**Theorem 4.1**
\[
\| (I + \theta k_j B_{j+1})^{-1}(I - (1 - \theta)k_j B_j) \| \leq \frac{1 + (1 - \theta)k_j c'_i}{1 + \theta k_j c'_{i\max}} \tag{4.5}
\]
where \(c'_{i\max} = \max c'_i\), \(c'_{i\min} = \min c'_i\). If \(c'_{i\min} = 0\), then this bound can be sharpened to
\[
\| (I + \theta k_j B_{j+1})^{-1}(I - (1 - \theta)k_j B_j) \| \leq \frac{1 + (1 - \theta)k_j c'_{i\max}}{1 + 2\theta k_j c'_{i\max}}, \tag{4.6}
\]
where
\[
\rho_{j+1} = \max [h''_i(h''_i + h''_{i+1}), h''_{i+1}(h''_{i+1} + h''_{i+2})]. \tag{4.7}
\]

**Theorem 4.2** Let (2.1) be nonautonomous and \(\max h_i\) be small enough. The method in (2.11) is then convergent. More precisely, we have
\[
\| U^\Delta_j - U_j \| = O(\max h_i(h_{i+1} - h_i)) + O(\max k_j), \text{ if } \theta \neq \frac{1}{2}, \tag{4.8}
\]
Proof: The proof is essentially the same as that for Theorem 3.2.

Obviously the bounds in Theorem 4.1 become sharper as the method becomes more implicit, i.e., $\theta \to 1$. As a result, it may be better to employ backward Euler if one wants to make use of the stability properties of this method. We have the following stability estimate.

**Theorem 4.3** For the product $D(l,j) = \prod_{m=l}^{j} (I + k_i B_l)^{-1}$ we have

$$
\| D(l,j) \| \leq \prod_{m=l}^{j} \frac{1}{1 + k_m c_{m+1}}. 
$$

(4.10)

In an analogy with (3.2) and (3.3), we define the local truncation error due to spatial discretization by $T_j$ and the corresponding time discretization by $\tilde{T}_j$. Hence we can write

$$
dU_j^\Delta = BjU_j^\Delta + T_j, 
$$

(4.11)

(where $B$ has to be understood as $B(t)$) and

$$
[I + k_j B_j]U_{j+1} = U_j + k_j (T_j + \tilde{T}_j). 
$$

(4.12)

We then obtain for infinite intervals

**Theorem 4.4** Let (2.1) be nonautonomous and let $\theta = 1$. If $\theta k, \lambda$ is not $o(1)$ as $s \to \infty$, then there exists a constant $C$ such that

$$
\| U_j^\Delta - U_j \| \leq C \| T_j \|. 
$$

(4.13)

**Corollary 4.1** Let (2.1) be nonautonomous and $\theta = 1$. If we assume that $\lim_{s \to 0} T_i = 0$, then $k, \lambda$ can grow unbounded and we have an asymptotically convergent method.

6. Convergence Analysis for Autonomous Problems on Variable Nonuniform Grids

If the problem in (2.1) admits solutions that exhibit regions of high spatial activity, it is crucial that the grid points be placed adaptively in these regions to obtain an efficient solution algorithm. The use of an equispaced or mildly nonuniform grid can require a large number of points to obtain the solution to the same accuracy. Many of the methods that have been used to obtain adaptive grid spacings at each time level can be interpreted as
equidistributing a positive weight function on a given interval. Essentially one attempts
to determine a mesh $M^j$ such that the weight function achieves the same variation over
each subinterval.

In this section we consider a more realistic situation—the number and the location of
the grid points can change from one time level to another, i.e., $h_i^j \neq h_i^{j+1}$ and $M^j \neq M^{j+1}$. Much of the work that implements adaptive spatial grids in the solution of time-dependent
partial differential equations employ a fixed number of grid points from one time level to
another. Another approach can be found in [9] in which the number of points is allowed
to change (either increase or decrease) as the calculation progresses from one time level to
another.

Given a solution $U_j$ at time level $j$ defined on a grid $M^j$, a strategy for determining
a new grid is stated as follows. We may change the location defining a new point as a
convex combination of two successive old points and we may add at most one point on a
newly formed interval $[x_i^{j+1}, x_i^{j+1}]$. We can also remove points with precisely the opposite
strategy. To use the actual discrete time integration we first have to interpolate values of
$U_j$ on the new grid thus forming $\tilde{U}_j$. Suppose we would like to know the ith coordinate of
$\tilde{U}_j$, say $\tilde{u}_i^j$ where $x_i^{j+1}$ is located in the interval $[x_i^m, x_i^{m+1}]$. We obtain $\tilde{u}_i^j$ by forming

$$\tilde{u}_i^j = u_i^m + \frac{u_i^{m+1} - u_i^m} {x_i^{m+1} - x_i^m} (x_i^{j+1} - x_i^m). \quad (5.1)$$

This regridding can be written in operator notation as

$$\tilde{U}_j = P_j U_j, \quad (5.2)$$

where $P_j$ is an $M^{j+1} \times M^j$ matrix. We have immediately

Property 5.1 $\| P_j \|_\infty = 1$.

This regridding procedure introduces another error source. For simplicity, we will
restrict ourselves to backward Euler, i.e.,

$$(I + k_j B_{j+1}) U_{j+1} = P_j U_j. \quad (5.3)$$

and upon combining Theorem 3.3 and Property 5.1, we have

Theorem 5.1 The method in (5.3) is uniformly stable.

To investigate convergence we have to realize that we also introduce a local interpola-
tion error $\tilde{T}_j$ so that the method in (5.3) can be rewritten in the form

$$(I + k_j B_{j+1}) U_{j+1}^\Delta = P_j U_j^\Delta + k_j [T_j + \tilde{T}_j] + \tilde{T}_j, \quad (5.4)$$

for the ith component of $\tilde{T}_j$ we have

$$\tilde{T}_i^j = O((x_i^{j+1} - x_i^m)(x_i^{j+1} - x_i^{m+1})). \quad (5.5)$$
Theoretically, this interpolation error might jeopardize convergence. However, if we decrease the time step, the interpolation error will decrease in a similar way. In particular, we see in (5.5) that $T_i' = O(h_i, k_j)$ as $x_i^{l+1}$ will shift $O(k_j)$ at most; this is evident in the example contained in Section 6.

**Theorem 5.2** For a grid where $h = \max h'$ and $k = \max k_j$ and with $\sum k_j \leq T < \infty$, the method in (5.3) converges, i.e., $\| U_j^\Lambda - U_j \| \to 0$ if $\max (h, k) \to 0$.

**Proof:** We have

$$\| U_j^\Lambda - U_j \| \leq \| \sum_{i=1}^{j} \left( \prod_{s=1}^{i}(I + k_s B_{s+1})^{-1} P_s \right) [k_i(T_i + \bar{T}_i) + \bar{T}_i] \| . \quad (5.6)$$

If we let $c = \min c(x,t)$, $x,t \in [0,1] \times [0,T]$, then for $k_j$ small enough we have

$$\| \prod_{s=1}^{j}(I + k_s B_{s+1})^{-1} \| < C \exp (-c(t_j - t_l)), \quad (5.7)$$

where $C$ is independent of $j,l$. Hence we have

$$\| U_j^\Lambda - U_j \| \leq C \sum_{i=1}^{j} \exp (-c(t_j - t_i)) [k_i(T_i + \bar{T}_i) + \bar{T}_i] \leq$$

$$C \exp (-ct_j) \sum_{i=1}^{j} \exp (ct_i) \| \bar{T}_i \| + O(k + h). \quad (5.8)$$

The first term on the right goes to zero as $\max (h, k) \to 0$ since $\| \bar{T}_i \| = O(hk_j)$ due to (5.5).

If we have an infinite time interval it would seem that the interpolation errors could interfere with convergence—in particular when $c \to 0$ (cf. (5.8)). However, on such infinite intervals it is most likely that one is approaching a steady-state. This implies that the time steps are growing and the grid is not changing asymptotically, i.e., $\lim_{t \to \infty} \bar{T}_i = 0$. This gives us

**Theorem 5.3** Let $\lim_{t_i \to \infty} \bar{T}_i = 0$. Then there exists a constant $C$ such that

$$\| U_j^\Lambda - U_j \| < C[\| T_j \| + \| \bar{T}_j \|]. \quad (5.9)$$

**Proof:** From Theorem 3.3 we see that for $j$ sufficiently large we have the damping property of the backward Euler method combined with the asymptotically zero interpolation error which essentially reduces this case to the one we dealt with in Theorem 4.4.

We can finally write
Corollary 6.1 If \( \lim_{t \to -\infty} T_t = \lim_{t \to -\infty} T_t = 0 \), then \( k \) can grow unbounded and the method is asymptotically convergent.

We note that long time integration problems, i.e., problems in which one desires a steady-state, bring up the question of how to find the large steps so that sufficient damping is achieved. From an efficiency point of view, it is important to have large steps. On the other hand, since the differential equation in (3.4) is stiff in the sense that it may have time scales of various orders (embodied by the eigenvalues of \( B \)), we have to resolve layers initially. Otherwise a change of the grid might introduce large interpolation errors, i.e., the derivative terms hidden in (5.5) can be prohibitively large to make the method successful.

6. Numerical Results

Test Problem

In this section we illustrate the ideas of the previous sections with an example that is constructed in such a way as to have the regions of high spatial activity move periodically as a function of time. Specifically, we are interested in solving

\[
\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + f,
\]

\[
u\left(\frac{1}{2}, t\right) = \exp \left(\frac{1 - \sin t}{2}\right),
\]

\[
u\left(-\frac{1}{2}, t\right) = \exp \left(\frac{1 + \sin t}{2}\right),
\]

(6.1)

\[
u(x, 0) = \exp \left(-x^2\right),
\]

where \([x, t] \in (-\frac{1}{2}, \frac{1}{2}) \times (0, T)\). The function \( f \) is given by

\[
f = \left[\frac{x - \frac{1}{2} \sin t}{\epsilon} \cos t + \frac{2}{\epsilon} - 4 \frac{x - \frac{1}{2} \sin t}{\epsilon^2}\right] u,
\]

(6.2)

and so the system in (6.1) admits the analytic solution

\[
u = \exp \left(-\frac{(x - \frac{1}{2} \sin t)^2}{\epsilon}\right).
\]

(6.3)

We note that the degree of "steepness" of the solution profile is controlled by the size of \( \epsilon \). In addition, as the solution is advanced in time, the profile propagates first to the right, then to the left and finally back again to \( x = 0 \). A full cycle takes \( 2\pi \) dimensionless time units.

Solution Algorithm
To solve this problem we utilize a variable space-variable time solution algorithm in which the boundary value problems at each time level are solved adaptively. This is particularly useful for problems similar to the one considered here. In this way the location of the grid points can change from one time level to another with the hope that the spatial portion of the discretization error will be uniformly bounded over all time steps. Specifically, we employ a backward Euler time differencing scheme with the spatial derivative approximated by second-order centered finite differences.

The adaptive selection of the grid points at each time level is determined by equidistributing a positive weight function. Omitting the time level superscript, we say that the mesh $M$ is equidistributed with respect to the constant $C$ and the weight function $w$ if

$$\int_{x_i}^{x_{i+1}} w \, dx = C,$$

More precisely, we attempt to construct a mesh $M$ on $[-1/2, 1/2]$ by equidistributing the difference in the components of the solution and its gradient between adjacent mesh points, i.e.,

$$\int_{x_i}^{x_{i+1}} \frac{du}{dx} \, dx \leq \delta \left( \max_{-1/2 \leq x \leq 1/2} u - \min_{-1/2 \leq x \leq 1/2} u \right) \quad i = 0, 1, \ldots, M - 1,$$

and

$$\int_{x_i}^{x_{i+1}} \frac{d^2u}{dx^2} \, dx \leq \gamma \left( \max_{-1/2 \leq x \leq 1/2} \frac{du}{dx} - \min_{-1/2 \leq x \leq 1/2} \frac{du}{dx} \right) \quad i = 1, 2, \ldots, M - 1,$$

where $\delta$ and $\gamma$ are small numbers less than one and the values of $\max u, \min u, \max \frac{du}{dx}, \text{ and } \min \frac{du}{dx}$ are estimated from a numerical solution on a previously determined grid.

We implement this procedure in such a way that the number of grid points remains fixed but the location of the points can change from one time level to another. We first compute the integrals in (6.5) and (6.6) over the entire domain $[-1/2, 1/2]$. These quantities are then divided by the specified number of subintervals desired so that we have an estimate of the equidistribution constant in (6.4). This represents the desired change in the solution and its derivative from one point to another. We next evaluate the integrals one subinterval at a time and when we determine an interval in which the cumulated value of the integrals becomes greater than $C$, the location of the first grid point can be determined by applying an inverse interpolation algorithm. This process can then be restarted and the integrals evaluated until they again become larger than $C$. In this way we can determine the location of each of the remaining interior grid points. This algorithm can, however, produce grid point locations that differ from one time level to another. This requires an interpolation strategy.

We observe that solution of the boundary value problem in (5.3) at time level $j$ requires a knowledge of the solution at time level $j - 1$ at the points of the mesh $M^j$. If the same spatial grid were used for the entire calculation at time levels $t^0, t^1, \ldots, t^j$ we would
have solution values available at the proper mesh points. The physical locations of the grid points at one time level would be identical to the locations of the points at any other. However, since we employ an inverse interpolation procedure, it is likely that the location of the grid points will change from one time step to another. In such cases we interpolate solution values from the previous time level to obtain solution information corresponding to the location of the grid points at the current time. An additional spatial discretization error is introduced into the finite difference algorithm as a result of the interpolation procedure (see also (5.4)).

Finally, the time steps \( k_j, j = 1, 2, \ldots, J \), are chosen such that an approximation to the local error per unit step associated with the time differencing scheme is below some given value. This procedure is carried out for each of the \( M^j - 2 \) interior nodes. We choose \( k_j \) as the smallest of all the calculated values. If after a successful calculation at the \( j^{th} \) level, we find that the time step determined from the error estimate should be smaller than the value we have used, we redo the calculation at the \( j^{th} \) level with the new value of \( k_j \). In addition, to prevent the time steps from increasing too rapidly, we impose the criteria \( k_{j+1}/k_j \leq 2 \).

Discussion

In Figure 1 we illustrate the location of the maximum value of the solution as a function of time for one complete cycle. The corresponding velocity of propagation for this maximum is given in Figure 2. Due to the lower velocity of the front as it approaches the boundaries of the domain, we anticipate smaller interpolation errors as compared to when the front is moving through the center of the domain. In Figure 3 we illustrate the maximum grid movement \( \max_i |x_{i}^{t+1} - x_{i}^{t}| \) for several fixed values of the time step \( k_j \) and 100 adaptively determine grid points. We note immediately the rapid reduction in the movement of the mesh as the front approaches the boundaries of the computational domain. To further investigate the influence of the time step on the grid movement (and, as a result, on the interpolation error) we illustrate in Figure 4 the maximal grid movement as a function of time for a set of calculations with \( k_j = 0.1 \) and \( M^j = 25, 50, 100 \). There is a similarity between the results in Figure 3 and the results in Figure 2. In addition, we see that the grid movement is almost independent of the number of nodes.

Finally, in Figure 5 we illustrate the coupled effects of the maximal grid movement and the size of the time steps. For each time step considered, three calculations were computed \( (M^j = 25, 50, 100) \). The results contained in the figure indicate that the maximal grid movement is approximately proportional to \( k_j \) which gives evidence to the result contained in Section 5, i.e., \( T_i = O(h_{m}^{j}k_j) \).
REFERENCES


Figure 1. Illustration of the solution profile at $t = 0$ (solid line), $t = \pi/2$ (dashed line) and $t = 3\pi/2$ (dotted line). The degree of steepness of the profile is controlled by the size of $\varepsilon$. 
Figure 2. Illustration of the velocity profile for the solution maximum.
Figure 3. Illustration of the maximal grid movement $|x_{i}^{t+1} - x_{m}^{t}|$ for 100 adaptively chosen grid points and $k_j = 10^{-1}, 10^{-2}, 10^{-3}$. 
Figure 4. Illustration of the maximal grid movement as a function of time for $k_j = 10^{-1}$ and $M_j = 25, 50, 100$. The grid movement is almost independent of the number of points.
Figure 5. Illustration of the size of the time steps versus the maximal grid movement. For each time step considered, three calculations were computed ($M_j = 25, 50, 100$). We note that the maximal grid movement is proportional to $k_j$. 