An Analysis of the Box and Trapezoidal Schemes for Linear Singularly Perturbed Boundary Value Problems

By Richard Weiss*

Abstract. Stability and convergence results are derived for the box and trapezoidal schemes applied to boundary value problems for linear singularly perturbed first order systems of o.d.e.'s without turning points.

1. Introduction. Most numerical methods for singularly perturbed boundary value problems in ordinary differential equations analyzed to date utilize some upwinding procedure. So they are applicable only if the underlying problem exhibits some special form or can be transformed to such a form by analytic techniques; see Abrahamsson, Keller and Kreiss [2], Kreiss and Nichols [6], Ringhofer [9].

Frequently, however, such explicit transformations are not available. Then one has to resort to some standard difference scheme combined with an adaptive mesh selection procedure. Successful computations of this kind have been reported with the trapezoidal scheme by Abrahamsson [1] and by Ascher, Christiansen and Russell [3] and Maier [7], who have used collocation methods.

Recently, Ascher and Weiss [4] have set out to investigate the applicability of a particular class of difference schemes, i.e. collocation with piecewise polynomials, to singular perturbation problems. They gave a detailed analysis of these schemes when applied to singularly perturbed first order systems with constant coefficients.

The present paper provides the basis for the extension of these results to more general problems. We provide an analysis of the box and trapezoidal schemes applied to boundary value problems for linear first order systems with variable coefficients (without turning points). The box and trapezoidal schemes are the simplest members of the families of collocation methods based on Gauss and Lobatto points, respectively.

We consider the system of $n + m$ equations, with n equations singularly perturbed,

$$(1.1) \qquad \varepsilon y' = A_{11}(t, \varepsilon) y + A_{12}(t, \varepsilon) z + f_1(t, \varepsilon),$$

$$(1.2) \qquad z' = A_{21}(t, \varepsilon) y + A_{22}(t, \varepsilon) z + f_2(t, \varepsilon), \qquad 0 \leqslant t \leqslant 1, \varepsilon > 0,$$

plus the boundary conditions

$$(1.3) \qquad B_0 \binom{y}{z}(0) + B_1 \binom{y}{z}(1) = \beta.$$

The matrices $A_{ij}(t, \varepsilon)$ and the right-hand sides $f_i(t, \varepsilon)$ are assumed to be smooth on $[0, 1] \times [0, \varepsilon_0]$, for some $\varepsilon_0 > 0$.

A key assumption for our analysis is the existence of a smooth matrix-valued function $E(t)$ such that

$$(1.4) \qquad E^{-1}(t) A_{11}(t, 0) E(t) = \mathrm{diag}\{\lambda_1(t), \ldots, \lambda_n(t)\}, \qquad 0 \leqslant t \leqslant 1,$$

i.e. that $A_{11}(t, 0)$ can be diagonalized by a smooth transformation, and that

$$(1.5) \qquad \begin{aligned} \mathrm{Re}\, \lambda_i(t) &< 0, \quad i = 1, \ldots, n_- \leqslant n, \\ \mathrm{Re}\, \lambda_i(t) &> 0, \quad i = n_- + 1, \ldots, n, n_- + n_+ = n. \end{aligned}$$

Our analysis shows that the box and trapezoidal schemes perform well provided that

(i) the conditions (1.4) and (1.5) are satisfied,

(ii) (1.1), (1.2) and (1.3) represent a well-posed boundary value problem, uniformly in ε,

(iii) the $n \times n$ matrix

$$\begin{pmatrix} E_-^{-1}(0) \\ E_+^{-1}(1) \end{pmatrix}$$

is nonsingular, where $E_-^{-1}(0)$ stands for the first n_- rows of $E^{-1}(0)$ and $E_+^{-1}(1)$ denotes the last n_+ rows of $E^{-1}(1)$, and

(iv) a sufficiently fine grid with gridspacings of size comparable to ε is used in the layer regions (and a "coarse" grid, just fine enough to resolve the reduced solution, is used on the remaining part of the interval).

While condition (1.4) can be relaxed, all other assumptions are essential. In particular, the methods do not work without the fine grids in the layer regions since the errors in the layers otherwise pollute the solution on the whole interval. The structure of grids suitable for the layer regions depends on the desired accuracy in the layers, on the eigenvalues of $A_{11}(0, 0)$ with negative real parts and on the eigenvalues of $A_{11}(1, 0)$ with positive real parts, respectively. For instance, equidistribution of the local truncation error leads to the following grid generation rule at $t = 0$: The local truncation errors all have approximate size δ provided that

$$h_1 = (\delta/c)^{1/2} \varepsilon, \qquad h_j = h_{j-1} \exp\left\{ -\frac{\mu}{2} h_{j-1}/\varepsilon \right\},$$

where c is a constant, $\mu = \max_{i=1,\ldots,n_-}\{\mathrm{Re}\, \lambda_i(0)\}$ and h_j denotes the jth gridspacing, counting from $t = 0$. This strategy is employed until the contribution of layer has decayed to δ, i.e. until $t = -\mu\varepsilon \ln \delta$. The number of gridpoints on the interval $[0, -\mu\varepsilon \ln \delta]$ generated by the above procedure can be shown to be proportional to $\delta^{-1/2}$. Note that it is independent of ε.

When these grids for the layer regions are combined with an appropriate coarse grid in the interior of the interval, the local truncation error of a suitable general solution of (1.1), (1.2) is kept below some threshold for all meshpoints on $[0, 1]$, uniformly in ε. Still the schemes do not perform satisfactorily unless condition (iii) holds. This is in contrast to the common opinion that meshes based on the equidistribution of the local truncation error are always safe to use.

We conclude this section with an outline of the paper. In Section 2 we collect a number of basic analytic results on linear singularly perturbed boundary value

problems which will be required in the analysis of the difference schemes. The schemes are defined in Section 3. Section 4 contains a stability result for scalar variable coefficient problems. In Sections 5 and 6 we present the analysis for the box and trapezoidal schemes, respectively, and some numerical results illustrating our theory are given in Section 7.

2. Analytic Preliminaries. Here we collect a number of well-known results for linear singularly perturbed first order systems; for details see O'Malley [8], Kreiss and Nichols [6].

We consider the first order system

(2.1a)
$$\varepsilon y' = A_{11} y + A_{12} z + f_1,$$
(2.1b)
$$z' = A_{21} y + A_{22} z + f_2, \qquad 0 \leqslant t \leqslant 1,$$

where y, z are vectors with n and m components, respectively, and where $A_{ij} = (a_{kl}^{ij}(t, \varepsilon))$, $f_i = f_i(t, \varepsilon)$, $i, j = 1, 2$. For simplicity of presentation we assume that $A_{ij}, f_i \in C^\infty([0, 1] \times [0, \varepsilon_0])$ for some positive ε_0.

We assume that

(2.2)
$$A_{11}(t, 0) = E(t)\Lambda(t)E^{-1}(t)$$

with $E \in C^\infty[0, 1]$ and

(2.3)
$$\Lambda(t) = \mathrm{diag}\{\lambda_1(t), \ldots, \lambda_n(t)\}$$

where

(2.4)
$$\mathrm{Re}\,\lambda_i(t) \begin{cases} < 0, & i = 1, \ldots, n_- \\ > 0, & i = n_- + 1, \ldots, n, \, n_- + n_+ = n, \, t \in [0, 1]. \end{cases}$$

Given an $m \times n$ matrix-valued function $L \in C^1[0, 1]$, the linear transformation

(2.5)
$$r = y, \qquad s = z - \varepsilon L y$$

applied to (2.1) yields

(2.6a)
$$\varepsilon r' = (A_{11} + \varepsilon A_{12} L) r + A_{12} s + f_1,$$
(2.6b)
$$s' = (-\varepsilon L' - L A_{11} + \varepsilon(A_{22} L - L A_{12} L) + A_{21}) r$$
$$+ (A_{22} - L A_{12}) s - L f_1 + f_2.$$

If L is chosen so that

(2.7)
$$\varepsilon L' = -L A_{11} + \varepsilon(A_{22} L - L A_{12} L) + A_{21},$$

then (2.6b) is uncoupled from (2.6a). It is a consequence of (2.2), (2.3), (2.4) that for any $k \geqslant 0$ and ε sufficiently small, say $\varepsilon \leqslant \varepsilon_1$, there is a solution $L \in C^k[0, 1]$ of (2.7) which satisfies

$$\sum_{j=0}^{k} \left\| \frac{d^j L}{dt^j} \right\| \leqslant \mathrm{const}, \qquad 0 < \varepsilon \leqslant \varepsilon_1.$$

$(\|\cdot\|$ is the maximum norm.) This L has an asymptotic expansion in powers of ε,

$$L = L(t, \varepsilon) = \sum_{j=0}^{k} L_j(t)\varepsilon^j + O(\varepsilon^{k+1}),$$

which can be determined by equating powers of ε in (2.7).

With L chosen in this way, (2.6a, b) become

$$\varepsilon r' = (A_{11} + \varepsilon A_{12}L)r + A_{12}s + f_1,$$
$$s' = (A_{22} - LA_{12})s - Lf_1 + f_2,$$

This system can be simplified further by the linear transformation

(2.8) $u = E^{-1}r, \qquad v = s$

which yields

(2.9a) $\varepsilon u' = (\Lambda + \varepsilon B_{11})u + B_{12}v + g_1,$

(2.9b) $v' = B_{22}v + g_2,$

where

$$B_{11}(t, \varepsilon) = E^{-1}(t)\big[(A_{11}(t, \varepsilon) - A_{11}(t,0))E(t)\varepsilon^{-1}$$
$$+ A_{12}(t, \varepsilon)L(t, \varepsilon)E(t) - E'(t)\big],$$
$$B_{12}(t, \varepsilon) = E^{-1}(t)A_{12}(t, \varepsilon),$$
$$B_{22}(t, \varepsilon) = A_{22}(t, \varepsilon) - L(t, \varepsilon)A_{12}(t, \varepsilon),$$
$$g_1(t, \varepsilon) = E^{-1}(t)f_1(t, \varepsilon),$$
$$g_2(t, \varepsilon) = f_2(t, \varepsilon) - L(t, \varepsilon)f_1(t, \varepsilon).$$

The system (2.9) is a very convenient basis for further analysis since the equations for u in (2.9a) are a diagonal system, up to a matrix of size ε. In particular, using a contraction mapping argument as in Kreiss and Nichols [6], it is easy to obtain a representation of the general solution of (2.9) which completely reveals the structure inherent in such systems.

In order to state these results we introduce some notation. Let $P_- \in R^{n_- \times n}$ and $P_+ \in R^{n_+ \times n}$ be the linear maps defined by

$$P_-x = \begin{pmatrix} x_1 \\ \vdots \\ x_{n_-} \end{pmatrix}, \quad P_+x = \begin{pmatrix} x_{n_-+1} \\ \vdots \\ x_n \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Also, let

(2.10) $Hw = g$

be a shorthand notation for (2.9), with

$$w(t) = \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}, \qquad g(t) = \begin{pmatrix} g_1(t) \\ g_2(t) \end{pmatrix}.$$

Then we have the key result

THEOREM 2.1. *The system* (2.9) *subject to the boundary conditions*

$$P_-u(0) = \eta_- \in R^{n_-}, \quad P_+u(1) = \eta_+ \in R^{n_+}, \quad v(0) = \eta_0 \in R^m$$

has a unique solution provided ε is sufficiently small, say $\varepsilon \leqslant \varepsilon_2$. This solution satisfies

(2.12) $\|w\| \leqslant \mathrm{const}(\|g\| + \|\eta_-\| + \|\eta_+\| + \|\eta_0\|), \qquad 0 < \varepsilon \leqslant \varepsilon_2.$

An asymptotic expansion argument combined with this theorem yields

THEOREM 2.2. *For any* $k \geq 0$ *there is a* (*particular*) *solution* $w_p(t)$ *of* (2.9) (*or* (2.10)) *which satisfies*

$$(2.13) \qquad \sum_{j=0}^{k} \left\| \frac{d^j w_p}{dt^j} \right\| \leq \text{const}, \qquad 0 < \varepsilon \leq \varepsilon_2.$$

Using Theorem 2.1 we can define matrix solutions W_-, W_+ and W_0 of (2.10) with $g = 0$ in the following way:

(i)

$$W_- = \begin{pmatrix} U_- \\ 0 \end{pmatrix}, \qquad U_- \in R^{n \times n_-},$$

$$\varepsilon U_-' = (\Lambda + \varepsilon B_{11}) U_-; \quad P_- U_-(0) = I, \quad P_+ U_-(1) = 0,$$

(ii)

$$W_+ = \begin{pmatrix} U_+ \\ 0 \end{pmatrix}, \qquad U_+ \in R^{n \times n_+},$$

$$\varepsilon U_+' = (\Lambda + \varepsilon B_{11}) U_+; \quad P_- U_+(0) = 0, \quad P_+ U_+(1) = I,$$

(iii)

$$W_0 = \begin{pmatrix} U_0 \\ V_0 \end{pmatrix} \in R^{(n+m) \times m};$$

$$HW_0 = 0; \quad V_0(0) = I, \quad P_- U_0(0) = S_-(\varepsilon), \quad P_+ U_0(1) = S_+(\varepsilon),$$

where (according to Theorem 2.2) the matrices $S_-, S_+ \in R^{(n \mp) \times m}$ can be chosen such that

$$(2.14) \qquad \sum_{j=0}^{k} \left\| \frac{d^j W_0}{dt^j} \right\| \leq \text{const}, \qquad 0 < \varepsilon \leq \varepsilon_2.$$

With the aid of these matrix solutions and the particular solution defined in Theorem 2.2 we obtain the desired representation of the general solution of (2.9).

THEOREM 2.3. *Any solution* w *of* (2.9) *can be written as*

$$(2.15) \qquad w = W_- \gamma_- + W_+ \gamma_+ + W_0 \gamma_0 + w_p,$$

with $\gamma_- \in R^{n_-}$, $\gamma_+ \in R^{n_+}$ *and* $\gamma_0 \in R^m$.

The standard method yields the existence of the asymptotic expansions

$$(2.16) \qquad \begin{cases} U_-(t) = \displaystyle\sum_{j=0}^{k} U_{-j}(t/\varepsilon) \varepsilon^j + O(\varepsilon^{k+1}), \\[2mm] U_+(t) = \displaystyle\sum_{j=0}^{k} U_{+j}((t-1)/\varepsilon) \varepsilon^j + O(\varepsilon^{k+1}), \\[2mm] W_0(t) = \displaystyle\sum_{j=0}^{k} W_{0j}(t) \varepsilon^j + O(\varepsilon^{k+1}). \end{cases}$$

It is clear that

$$(2.17) \qquad U_{-0} \begin{pmatrix} \exp(\Lambda_-(0)t/\varepsilon) \\ 0 \end{pmatrix}, \qquad U_{+0} = \begin{pmatrix} 0 \\ \exp(\Lambda_+(1)(t-1)/\varepsilon) \end{pmatrix},$$

where

$$(2.18) \quad \Lambda_-(t) = \mathrm{diag}\{\lambda_1(t),\ldots,\lambda_{n_-}(t)\}, \qquad \Lambda_+ = \mathrm{diag}\{\lambda_{n_-+1}(t),\ldots,\lambda_n(t)\}.$$

Now we return to (2.1) and assume that boundary conditions

$$(2.19) \qquad B_0(\varepsilon)\binom{y}{z}(0) + B_1(\varepsilon)\binom{y}{z}(1) = \beta(\varepsilon)$$

are given, where B_0, B_1 and β depend smoothly on ε. By the transformations (2.5) and (2.8) these boundary conditions are changed to

$$(2.20) \qquad C_0(\varepsilon)w(0) + C_1(\varepsilon)w(1) = \zeta(\varepsilon).$$

Substituting (2.15) into (2.20), we obtain a linear system

$$(2.21) \qquad M(\varepsilon)\gamma = \xi(\varepsilon)$$

for $\gamma = (\gamma_-, \gamma_+, \gamma_0)$. The matrix M has an expansion

$$M(\varepsilon) = \sum_{j=0}^{k} M_j \varepsilon^j + O(\varepsilon^k).$$

For the sequel we assume that M_0 is nonsingular. This is equivalent to assuming that the boundary value problem (2.1), (2.19) be well posed, uniformly in ε, i.e. that

$$(2.22) \qquad \left\| \begin{matrix} y \\ z \end{matrix} \right\| \leqslant \mathrm{const}\left(\left\| \begin{matrix} f_1 \\ f_2 \end{matrix} \right\| + \|\beta\| \right), \qquad 0 < \varepsilon \leqslant \varepsilon_3,$$

with a constant independent of ε.

In the analysis of the numerical methods we shall require certain representations of the general solution of (2.9) not only on the interval $[0,1]$ but on any interval $[\underline{t}, \bar{t}]$ with $0 \leqslant \underline{t} < \bar{t} \leqslant 1$. This is achieved by defining W_-, W_+, W_0 as previously, but with $t = 0$ and $t = 1$ replaced by $t = \underline{t}$ and $t = \bar{t}$. W_-, W_+ and W_0 so defined have asymptotic expansions analogous to (2.16). Of course, $\underline{t}$ and $\bar{t}$ now enter into the coefficients of these expansions. For instance, corresponding to (2.17) we have

$$(2.23) \quad U_{-0} = \begin{pmatrix} \exp(\Lambda(\underline{t})(t-\underline{t})/\varepsilon) \\ 0 \end{pmatrix}, \qquad U_{+0} = \begin{pmatrix} 0 \\ \exp(\Lambda_+(\bar{t})(t-\bar{t})/\varepsilon) \end{pmatrix}.$$

Denoting by $(U_-)_l$ and $(U_+)_l$ the lth column of U_- and U_+, respectively, we obtain from (2.23) and (2.9) the following estimates:

$$(2.24) \qquad \left\| \frac{d^i(U_-)_l}{dt^i} \right\| \leqslant \mathrm{const}\, \varepsilon^{-i}\big(\exp\{\mathrm{Re}\,\lambda_l(\underline{t})(t-\underline{t})/\varepsilon\} + O(\varepsilon)\big),$$

$$\underline{t} \leqslant t \leqslant \bar{t}, l = 1,\ldots, n_-, i = 0, 1,\ldots, k,$$

$$\left\| \frac{d^i(U_+)_l}{dt^i} \right\| \leqslant \mathrm{const}\, \varepsilon^{-i}\big(\exp\{\mathrm{Re}\,\lambda_{n_-+l}(\bar{t})(t-\bar{t})/\varepsilon\} + + O(\varepsilon)\big),$$

$$\underline{t} \leqslant t \leqslant \bar{t}, l = 1,\ldots, n_+, i = 0, 1,\ldots, k.$$

3. The Difference Schemes. On $[0, 1]$ we introduce the partition

(3.1) $$\Delta = \{0 = t_1 < t_2 < \cdots < t_N < t_{N+1} = 1\}$$

and define $h_i = t_{i+1} - t_i$, $t_{i+1/2} = (t_i + t_{i+1})/2$, $i = 0, \dots, N$.
The box scheme for (2.1), (2.19) is

(3.2) $$\varepsilon \frac{y_{i+1} - y_i}{h_i} = A_{11}(t_{i+1/2})\frac{y_i + y_{i+1}}{2} + A_{12}(t_{i+1/2})\frac{z_i + z_{i+1}}{2} + f_1(t_{i+1/2}),$$

$$\frac{z_{i+1} - z_i}{h_i} = A_{21}(t_{i+1/2})\frac{y_i + y_{i+1}}{2} + A_{22}(t_{i+1/2})\frac{z_i + z_{i+1}}{2} + f_2(t_{i+1/2}),$$

$$i = 1, \dots, N,$$

(3.3) $$B_0\begin{pmatrix} y_1 \\ z_1 \end{pmatrix} + B_1\begin{pmatrix} y_{N+1} \\ z_{N+1} \end{pmatrix} = \beta.$$

(For reasons of brevity we do not indicate the potential dependence of A_{ij}, f_i, B_i and β on ε.)

In the trapezoidal scheme, (3.2) is replaced by

(3.4) $$\varepsilon \frac{y_{i+1} - y_i}{h_i} = \frac{A_{11}(t_i)y_i + A_{11}(t_{i+1})y_{i+1}}{2} + \frac{A_{12}(t_i)z_i + A_{12}(t_{i+1})z_{i+1}}{2}$$

$$+ \frac{f_1(t_i) + f_1(t_{i+1})}{2},$$

$$\frac{z_{i+1} - z_i}{h_i} = \frac{A_{21}(t_i)y_i + A_{21}(t_{i+1})y_{i+1}}{2} + \frac{A_{22}(t_i)z_i + A_{22}(t_{i+1})z_{i+1}}{2}$$

$$+ \frac{f_2(t_i) + f_2(t_{i+1})}{2}, \qquad i = 1, \dots, N.$$

An important step in the analysis of the difference schemes is a transformation of the discrete variables (x_i, y_i) analogous to the transformation of (2.1) to the form (2.9). Introducing the new variables

(3.5) $$\begin{pmatrix} u_i \\ v_i \end{pmatrix} = \begin{pmatrix} E^{-1}(t_i) & 0 \\ -\varepsilon L(t_i) & I \end{pmatrix}\begin{pmatrix} y_i \\ z_i \end{pmatrix}, \qquad i = 1, \dots, N+1,$$

we obtain, after some straightforward algebra for the box scheme (3.2),

(3.6) $$\begin{pmatrix} \varepsilon \dfrac{u_{i+1} - u_i}{h_i} \\ \dfrac{v_{i+1} - v_i}{h_i} \end{pmatrix} = B(t_{i+1/2})\begin{pmatrix} \dfrac{u_i + u_{i+1}}{2} \\ \dfrac{v_i + v_{i+1}}{2} \end{pmatrix} + R_i[u_i, u_{i+1}, v_i, v_{i+1}]$$

$$+ \begin{pmatrix} E^{-1}(t_{i+1/2})f_1(t_{i+1/2}) \\ -L(t_{i+1/2})f_1(t_{i+1/2}) + f_2(t_{i+1/2}) \end{pmatrix}, \quad i = 1, \dots, N,$$

where

(3.7) $$B(t) = \begin{pmatrix} \Lambda(t) + \varepsilon B_{11}(t) & B_{12}(t) \\ 0 & B_{22}(t) \end{pmatrix}$$

and where the R_i are linear maps from $R^{2(n+m)}$ in R^{n+m} with

(3.8) $$\|R_i\| \leqslant Kh_i, \qquad i = 1, \dots, N; K = \text{const.}$$

The corresponding equations for the trapezoidal scheme (3.4) are

$$(3.9) \quad \begin{pmatrix} \varepsilon \dfrac{u_{i+1} - u_i}{2} \\ \dfrac{v_{i+1} - v_i}{2} \end{pmatrix} = \frac{1}{2} \left\{ B(t_i) \begin{pmatrix} u_i \\ v_i \end{pmatrix} + B(t_{i+1}) \begin{pmatrix} u_{i+1} \\ v_{i+1} \end{pmatrix} \right\} + \hat{R}_i [u_i, u_{i+1}, v_i, v_{i+1}]$$

$$+ \frac{1}{2} \begin{bmatrix} E^{-1}(t_{i+1/2}) & 0 \\ -L(t_{i+1/2}) & I \end{bmatrix} \begin{bmatrix} f_1(t_i) + f_1(t_{i+1}) \\ f_2(t_i) + f_2(t_{i+1}) \end{bmatrix}, \quad i = 1, \dots, N,$$

where the $\hat{R}_i$ are maps analogous to the R_i and they satisfy an analogous estimate.

Part of the analysis of the difference schemes will be based on (3.6) and (3.9) rather than on (3.2) and (3.4). Note that due to the occurrence of the R_i, $\hat{R}_i$, (3.6) and (3.9) are not precisely the box and trapezoidal schemes for (2.9). This would be the case only when E and L are constant matrices.

For the analysis of these schemes we assume that the partition Δ has the following structure: The meshspacings h_i are comparable to ε for $i = 1, \dots, N^{(0)}$ and $i = N - N^{(1)} + 1, \dots, N$, where $N^{(0)}$, $N^{(1)}$ are given. In between, i.e. for $i = N^{(0)} + 1, \dots, N - N^{(1)}$, no such restriction is posed on the size of the gridspacings and we will show that the choice $h_i \gg \varepsilon$ is feasible.
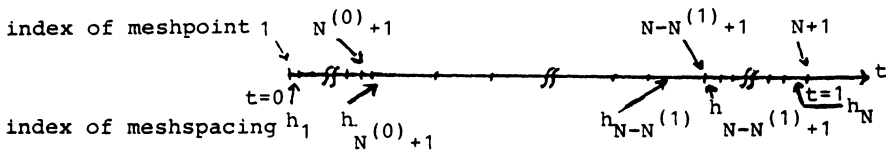

FIGURE 3.1. *The grid*

4. A Stability Result and Notation. Here we establish an estimate for the solution of the difference equation arising from the discretization of linear scalar singularly perturbed differential equations. In the subsequent sections this result will be used for the treatment of vector systems. Also, we introduce some shorthand notation.

To state our estimate we shall employ a grid

$$(4.1) \qquad \{ 0 \leqslant \tau_1 < \tau_2 < \cdots < \tau_I < \tau_{I+1} \leqslant 1 \},$$

which in later sections will be identified with different portions of the grid (3.1). As before, let

$$h_i = \tau_{i+1} - \tau_i, \qquad \tau_{i+1/2} = (\tau_i + \tau_{i+1})/2, \qquad i = 1, \dots, I,$$

and in addition, let

$$(4.2) \quad \bar{h} = \max\{ h_i \mid i = 1, \dots, I \}, \qquad \lambda(t) = \alpha(t) + i\beta(t), \qquad \tau_1 \leqslant t \leqslant \tau_{I+1},$$

with $\alpha, \beta \in C[\tau_1, \tau_{I+1}], \alpha(t) < 0, \tau_1 \leqslant t \leqslant \tau_{I+1}$, and define

$$\underline{\alpha} = \min\{ \alpha(t) \mid \tau_1 \leqslant t \leqslant \tau_{I+1} \}, \qquad \bar{\alpha} = \max\{ \alpha(t) \mid \tau_1 \leqslant t \leqslant \tau_{I+1} \},$$

$$\bar{\gamma}^2 = \max\{ (\beta(t)/\alpha(t))^2 \mid \tau_1 \leqslant t \leqslant \tau_{I+1} \},$$

$$\sigma = -\bar{h}\underline{\alpha}/2\varepsilon, \qquad \rho = (1 + \sigma)^2 + \sigma^2 \bar{\gamma}^2.$$

Then we have

LEMMA 4.1. *The solution* $\langle y_1, \ldots, y_{I+1} \rangle$ *of the recursion*

$$(4.3) \qquad \varepsilon \frac{y_{i+1} - y_i}{h_i} - \lambda(\tau_{i+1/2}) \frac{y_{i+1} + y_i}{2} = f_i, \qquad i = 1, \ldots, I,$$

satisfies

$$|y_i| \leqslant |y_1| \exp\{\bar{\alpha}(\tau_i - \tau_1)/\varepsilon\rho\} + \frac{\rho}{|\bar{\alpha}|} \max\{|f_j|, j = 1, \ldots, i - 1\}, \qquad i = 1, \ldots, I + 1.$$

Proof. Let ω be a complex number with $\xi = \operatorname{Re} \omega < 0$ and $\eta = \operatorname{Im} \omega / \xi$. Then

$$\left| \frac{1 + \omega}{1 - \omega} \right|^2 = \frac{(1 + \xi)^2 + \xi^2 \eta^2}{(1 - \xi)^2 + \xi^2 \eta^2} = 1 + \frac{4\xi}{(1 - \xi)^2 + \xi^2 \eta^2}$$

$$\leqslant \exp\{4\xi/((1 - \xi)^2 + \xi^2 \eta^2)\},$$

whence

$$(4.4) \qquad \left| \frac{1 + \omega}{1 - \omega} \right| \leqslant \exp\{2\xi/((1 - \xi)^2 + \xi^2 \eta^2)\}.$$

Rewriting (4.3) as

$$y_{i+1} = y_i \left(\frac{1 + \omega_i}{1 - \omega_i} \right) + \frac{h_i}{\varepsilon} (1 - \omega_i)^{-1} f_i,$$

where

$$(4.5) \qquad \omega_i = \lambda(\tau_{i+1/2}) h_i / 2\varepsilon,$$

we obtain the solution

$$(4.6) \qquad y_i = \sum_{j=1}^{i-1} \left[\prod_{l=j+1}^{i-1} \left(\frac{1 + \omega_l}{1 - \omega_l} \right) \right] \frac{h_j}{\varepsilon} (1 - \omega_j)^{-1} f_j + \prod_{l=1}^{i-1} \left(\frac{1 + \omega_l}{1 - \omega_l} \right) y_1.$$

By (4.4) and (4.5)

$$(4.7) \qquad \left| \frac{1 + \omega_i}{1 - \omega_i} \right| \leqslant \exp\{\bar{\alpha} h_i / \rho \varepsilon\}.$$

Hence, by (4.6)

$$|y_i| \leqslant \varepsilon^{-1} \max\{|f_j| \, | \, j = 1, \ldots, i - 1\} \sum_{j=1}^{i-1} h_j \exp\{\bar{\alpha}(\tau_i - \tau_{j+1})/\rho\varepsilon\}$$

$$+ |y_1| \exp\{\bar{\alpha}(\tau_i - \tau_1)/\rho\varepsilon\}, \qquad i = 1, \ldots, I + 1.$$

The lemma follows since

$$\varepsilon^{-1} \sum_{j=1}^{i-1} h_j \exp\{\bar{\alpha}(\tau_i - \tau_{j+1})/\rho\varepsilon\} \leqslant \varepsilon^{-1} \int_{\tau_1}^{\tau_i} \exp\{\bar{\alpha}(\tau_i - s)/\rho\varepsilon\} \, ds \leqslant \rho/|\bar{\alpha}|. \qquad \square$$

Note that Lemma 4.1 is mainly useful when $\bar{h} \leqslant K_0 \varepsilon$, where K_0 is a constant, and when the ratio of imaginary to real parts of $\lambda(t)$ is of reasonable size (i.e. the problem is not highly oscillatory).

Now we introduce some notation. Given a grid (4.1) and vectors or matrices $s_i \in R^l$, for $i = 1, \ldots, I + 1$ (or $i = 1, \ldots, I$) and some integer l, we set $s_\Delta = \{s_1, \ldots, s_{I+1}\}$ (or $s_\Delta = \{s_1, \ldots, s_I\}$). We define

$$\|s_\Delta\| = \max\{\|s_i\|, \ i = 1, \ldots, I + 1\}$$

and analogously for the other case, where $\|s_i\|$ is the maximum norm of s_i. Given a function $s \in C[\tau_1, \tau_{I+1}]$, we define the restriction of s to the grid,

$$\Delta s = \{s(\tau_1), \ldots, s(\tau_{I+1})\}.$$

By $c_1, c_2, \ldots$ we shall denote positive constants.

5. The Analysis of the Box Scheme. The analysis proceeds in two main steps: First representations of the general solution of the difference equations (3.2) viz (3.6) are derived separately for each of the intervals $[t_1, t_{N^{(0)}+1}]$, $[t_{N^{(0)}+1}, t_{N-N^{(1)}+1}]$, $[t_{N-N^{(1)}+1}, 1]$, and these representations are related to the general solution of the differential system (2.1) viz. (2.9). Then the three representations are combined to yield the general solution of the difference equations on the whole interval $[0, 1]$, and the remaining free constants in this general solution are determined through the boundary conditions (3.3).

5.1. *The Interval* $[t_1, t_{N^{(0)}+1}]$. Here it is convenient to use the transformed difference equations (3.6). We write these equations for $i = 1, \ldots, N^{(0)} + 1$ in compact form as

$$(5.1) \qquad L_\Delta \begin{bmatrix} u_\Delta \\ v_\Delta \end{bmatrix} = \begin{bmatrix} \tilde{L}_\Delta^{11} & \tilde{L}_\Delta^{12} \\ \tilde{L}_\Delta^{21} & \tilde{L}_\Delta^{22} \end{bmatrix} \begin{bmatrix} u_\Delta \\ v_\Delta \end{bmatrix} + \begin{bmatrix} g_\Delta^1 \\ g_\Delta^2 \end{bmatrix},$$

where

$$u_\Delta = \{u_1, \ldots, u_{N^{(0)}+1}\}, \qquad v_\Delta = \{v_1, \ldots, v_{N^{(0)}+1}\},$$

$$L_\Delta \begin{bmatrix} u_\Delta \\ v_\Delta \end{bmatrix} = \begin{bmatrix} L_\Delta^1 & u_\Delta \\ L_\Delta^2 & v_\Delta \end{bmatrix} = \left\{ \begin{array}{l} \varepsilon \dfrac{u_{i+1} - u_i}{h_i} - \Lambda(t_{i+1/2}) \dfrac{u_i + u_{i+1}}{2} \\[2mm] \dfrac{v_{i+1} - v_i}{h_i} - B_{22}(t_{i+1/2}) \dfrac{v_i + v_{i+1}}{2} \end{array} \right\}, \qquad i = 1, \ldots, N^{(0)}.$$

The $\tilde{L}_\Delta^{ij}$ stand for the remaining parts of the difference operator, and g_Δ^1, g_Δ^2 contain the inhomogeneous terms. From (3.7) it is clear that

$$(5.2) \quad \|\tilde{L}_\Delta^{11}\|, \|\tilde{L}_\Delta^{21}\|, \|\tilde{L}_\Delta^{22}\| \leqslant c_1(\bar{h}_- + \varepsilon), \qquad \bar{h}_- = \max\{h_i \mid i = 1, \ldots, N^{(0)}\},$$

and $\|\tilde{L}_\Delta^{12}\| \leqslant c_2$.

Now we impose boundary conditions

$$(5.3) \qquad P_- u_1 = \eta_- \in R^{n_-}, \quad P_+ u_{N^{(0)}+1} = \eta_+ \in R^{n_+}, \quad v_1 = \eta_0 \in R^m$$

and proceed to show that (5.1) subject to these boundary conditions has a unique solution provided that

$$(5.4) \qquad h_i \leqslant K_0 \varepsilon, \qquad i = 1, \ldots, N^{(0)},$$

with a suitable constant K_0.

We begin with a discussion of the structure of the difference operator L_Δ. The first n components, $L_\Delta^1 u_\Delta$, are scalar recursions which can be analyzed with the aid of Lemma 4.1 in the following way: When treating the first n_- of these recursions,

which belong to the eigenvalues of Λ with negative real parts (see (2.4)), the grid (4.1) is identified with the grid $\{0 = t_1 < t_2 < \cdots < t_{N^{(0)}+1}\}$, i.e. with the fine left end portion of the grid (3.1). And the function $\lambda(t)$ of (4.2) is replaced by $\lambda_i(t)$ of (2.3) when considering the ith recursion; $i = 1, \ldots, n_-$. When dealing with the remaining n_+ recursions, the grid (4.1) is again identified with the fine left end portion of the grid (3.1), but now according to the labeling $\tau_j = t_{N^{(0)}+1} - t_{N^{(0)}+2-j}$, $j = 1, \ldots, N^{(0)} + 1$. And for $\lambda(t)$ of (4.2) we take $\lambda_i(t_{N^{(0)}+1} - t)$, $i = n_- + 1, \ldots, n$. Then it follows from Lemma 4.1 that the problem

$$L_\Delta^1 u_\Delta = g_\Delta, \quad P_- u_1 = \eta_-, \quad P_+ u_{N^{(0)}+1} = \eta_+$$

has a unique solution for all g_Δ, η_-, η_+, and this solution satisfies

(5.5) $$\|u_\Delta\| \leq \|\eta_-\| + \|\eta_+\| + d\|g_\Delta\|,$$

where the quantity d is the largest of the constants $\rho/|\bar{\alpha}|$ obtained from Lemma 4.1. When $\bar{h}_- = K_0 \varepsilon$, then $d = d(K_0)$ in (5.5) is

(5.6) $$d = \left((1 - K_0 \underline{\alpha}/2)^2 + K_0^2 \underline{\alpha}^2 \bar{\gamma}^2/4\right)/|\bar{\alpha}|.$$

The last m components of L_Δ are easily analyzed as well. It is clear that the problem $L_\Delta^2 v_\Delta = g_\Delta$, $v_1 = \eta_0$, has a unique solution for all g_Δ, η_0 provided that

(5.7) $$\bar{h}_- \leq \|B_{22}\|/4,$$

and this solution satisfies

(5.8) $$\|v_\Delta\| \leq e(\|\eta_0\| + \|g_\Delta\|), \qquad e = \text{const.}$$

Applying the estimates (5.5) and (5.8) to (5.1), (5.3) and using (5.2), we obtain

$$\|u_\Delta\| \leq \|\eta_-\| + \|\eta_+\| + d\{c_1(\bar{h}_- + \varepsilon)\|u_\Delta\| + c_2\|v_\Delta\| + \|g_\Delta^1\|\}$$
$$\|v_\Delta\| \leq e\{\|\eta_0\| + c_1(\bar{h}_- + \varepsilon)(\|u_\Delta\| + \|v_\Delta\|) + \|g_\Delta^2\|\}.$$

When

(5.9) $$c_1(\bar{h}_- + \varepsilon)(d + e + c_2 de) \leq (2 - \sqrt{2})/2,$$

this yields the final stability result for (5.1):

(5.10) $$\|u_\Delta\| \leq 2\{\|\eta_-\| + \|\eta_+\| + d\|g_\Delta^1\| + dc_2 e(\|\eta_0\| + \|g_\Delta^2\|)\},$$
$$\|v_\Delta\| \leq 2e\{\|\eta_0\| + c_1(\bar{h}_- + \varepsilon)(\|\eta_-\| + \|\eta_+\| + \|g_\Delta^1\|) + \|g_\Delta^2\|\}.$$

Next we state a representation of the general solution of (5.1) which is the discrete counterpart of the representation (2.15) and is a simple consequence of (5.10). For reasons of brevity we write (5.1) as

$$H_\Delta w_\Delta = g_\Delta.$$

THEOREM 5.1. *The general solution of* (5.1) *on* $[t_0, t_{N^{(0)}+1}]$ *can be written as*

(5.11) $$w_\Delta = {}_1W_\Delta^- \eta_- + {}_1W_\Delta^+ \eta_+ + {}_1W_\Delta^0 \eta_0 + {}_1w_{p,\Delta},$$

where

$${}_1W_\Delta^{-,+,0} = \left\{{}_1W_j^{-,+,0}\right\}_{j=1}^{N^{(0)}+1},$$

with

$$_1W_j^- = \begin{pmatrix} _1U_j^- \\ _1V_j^- \end{pmatrix} \in R^{(n+m)\times n_-}, \qquad _1W_j^+ = \begin{pmatrix} _1U_j^+ \\ _1V_j^+ \end{pmatrix} \in R^{(n+m)\times n_+},$$

$$_1W_j^0 = \begin{pmatrix} _1U_j^0 \\ _1V_j^0 \end{pmatrix} \in R^{(n+m)\times m},$$

are matrix solutions defined by

$$H_\Delta {}_1W_\Delta^- = 0, \qquad P_- {}_1U_1^- = I, \qquad P_+ {}_1U_{N^{(0)}+1}^- = 0, \qquad {}_1V_1^- = 0,$$

$$H_\Delta {}_1W_\Delta^+ = 0, \qquad P_- {}_1U_1^+ = 0, \qquad P_+ {}_1U_{N^{(0)}+1}^+ = I, \qquad {}_1V_1^+ = 0,$$

$$H_\Delta {}_1W_\Delta^0 = 0, \qquad P_- {}_1U_1^0 = {}_1S_-, \qquad P_+ {}_1U_{N^{(0)}+1}^0 = {}_1S_+, \qquad {}_1V_1^0 = I,$$

and

$$_1w_{p,\Delta} = \begin{Bmatrix} _1u_{p,\Delta} \\ _1v_{p,\Delta} \end{Bmatrix}$$

is a particular solution defined by

$$H_\Delta {}_1w_{p,\Delta} = g_\Delta, \quad P_- {}_1u_{p,1} = P_- u_p(0), \quad P_+ {}_1u_{p,N^{(0)}+1} = P_+ u_p(t_{N^{(0)}+1}),$$

$$_1v_{p,1} = v_p(0).$$

Here the matrices $_1S_-$ and $_1S_+$ are defined by

$$_1S_- = P_- U_0(0), \qquad _1S_+ = P_+ U_0(t_{N^{(0)}+1}),$$

where $U_0(t)$ is the "upper part" of $W_0(t)$ (see Theorem 2.3), and $w_p(t) = (u_p(t), v_p(t))$ is defined in Theorem 2.2. (The subscript "1" in $_1W^{-,+,0}$ and so forth indicates that the representation (5.11) is valid for the first portion of the grid (3.1).)

It is clear that $_1W_\Delta^0$ and $_1w_{p,\Delta}$ are approximations to $W_0(t)$ and $w_p(t)$. Because of the estimates (2.13), (2.14) and of (5.10), the standard consistency plus stability argument yields

(5.12) $$\|_1W_\Delta^0 - \Delta W_0\| \leqslant c_3 \bar{h}_-^2,$$

(5.13) $$\|_1w_{p,\Delta} - \Delta w_p\| \leqslant c_3 \bar{h}_-^2.$$

Also, $_1W_\Delta^-$ and $_1W_\Delta^+$ are approximations to $W_-(t)$, $W_+(t)$ on $[0, t_{N^{(0)}+1}]$. However, since negative powers of ε occur in the estimates (2.24), the derivation of bounds analogous to (5.12), (5.13) is more delicate. First we consider the local truncation error $H_\Delta[\Delta W_-]$: From (5.2) and (2.24) it follows that the local truncation error of the lth column of W_- is

$$H_\Delta \begin{bmatrix} \Delta(U_-)_l \\ 0 \end{bmatrix} = \left\{ \begin{pmatrix} r_i \\ s_i \end{pmatrix} \right\}_{i=1}^{N^{(0)}}, \qquad l = 1, \dots, n_-,$$

with $r_i = (r_{i1}, \dots, r_{in_-})^T$,

$$|r_{ik}| \leqslant c_4 \left(\left(\frac{h_i}{\varepsilon} \right)^2 + h_i \right) (\exp\{\operatorname{Re} \lambda_l(0) t_i/\varepsilon\} \delta_{kl} + \varepsilon), \qquad k = 1, \dots, n_-,$$

and

$$\|s_i\| \leqslant c_4 h_i (\exp\{\operatorname{Re} \lambda_l(0) t_i/\varepsilon\} + \varepsilon),$$

for $i = 1, \ldots, N^{(0)}$. Here δ_{kl} is the Kronecker delta. If the grid is such that

$$(5.14) \qquad \left\| \begin{pmatrix} r_i \\ s_i \end{pmatrix} \right\| \leqslant \delta, \qquad i = 1, \ldots, N^{(0)},$$

for a given "tolerance" δ, then, using (5.10),

$$(5.15) \qquad \|_1 W_\Delta^- - \Delta W_- \| \leqslant c_5 \delta.$$

An efficient way to achieve (5.14) is to equidistribute the principal part of the local error, i.e., to determine the grid such that

$$(5.16) \qquad c_4 \left(\frac{h_i}{\varepsilon} \right)^2 \exp\{\operatorname{Re} \lambda_l(0) t_i / \varepsilon\} = c_4 \left(\frac{h_1}{\varepsilon} \right)^2 = \delta.$$

Taking the (an) l for which $\operatorname{Re} \lambda_l(0)$ is smallest in absolute value, this leads to the strategy already developed in Ascher and Weiss [4],

$$(5.17) \qquad h_i = h_1 \exp\{-\operatorname{Re} \lambda_l(0) t_i / 2\varepsilon\}, \qquad i = 2, \ldots, N^{(0)},$$

$$h_1 = (\delta/c_4)^{1/2} \varepsilon.$$

This generates an increasing sequence $\{h_i \mid i = 1, \ldots, N^{(0)}\}$. Of course, now the question arises whether $\bar{h}_- = h_{N^{(0)}}$ obtained in this way is not too large for the previous existence theory to be meaningful. An estimate of the magnitude of $h_{N^{(0)}}$ can be obtained in the following way: It is natural to terminate the strategy (5.17) once a value $t_{N^{(0)}}$ is obtained such that

$$\exp\{\operatorname{Re} \lambda_l(0) t_{N^{(0)}} / \varepsilon\} \approx \delta,$$

i.e. when the contribution of the layer has decreased to the magnitude of δ. Then, using (5.16),

$$c_4 \left(\frac{h_{N^{(0)}}}{\varepsilon} \right)^2 \delta \approx \delta; \qquad h_{N^{(0)}} \approx \varepsilon c_4^{1/2}.$$

So the constant K_0 in (5.4) has the magnitude of $c_4^{1/2}$, essentially independent of ε.

Equation (5.15) expresses the fact that $_1 W_\Delta^-$ is an approximation to $W_-(t)$ on the fine grid, provided the grid is selected according to (5.14). Given such a grid, we now analyze $_1 W_\Delta^+$. The reason for determining the grid on the basis of $_1 W_\Delta^-$ and not on the basis of $_1 W_\Delta^+$ is that $_1 W_\Delta^-$ will contribute significantly to the general solution of the difference equation (3.2) on the whole interval [0, 1]. The contribution of $_1 W_\Delta^+$ on the other hand will turn out to be insignificant, once the representations of the solution on the three subintervals are combined to one representation valid on [0, 1]. Note that there is no analytic counterpart to $_1 W_\Delta^+$ in (2.15).

We write

$$(5.18) \qquad _1 W_\Delta^+ = \begin{pmatrix} _1 U_\Delta^+ \\ _1 V_\Delta^+ \end{pmatrix}$$

and consider $P_{+\,1} U_\Delta^+$. This is a set of $N^{(0)} + 1$ matrices of size $n_+ \times n_+$. For $i = 1, \ldots, N^{(0)} + 1$ we denote the lth column of the corresponding matrix by u_i^l, so that

$$P_{+\,1} U_\Delta^+ = \left\{ \left(u_i^l \right)_{l=1}^{n_+} \right\}_{i=1}^{N^{(0)}+1}.$$

Then we decompose u_i^l in the form

(5.19) $$u_i^l = \xi_i e_l + \eta_i^l, \qquad i = 1, \ldots, N^{(0)} + 1, \eta_i^l \in R^{n_+},$$

where $e_l = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ is the lth unit vector, the scalars ξ_i are defined by

(5.20) $$\varepsilon \frac{\xi_{i+1} - \xi_i}{h_i} - \lambda_k(t_{i+1/2}) \frac{\xi_i + \xi_{i+1}}{2} = 0, \qquad i = 1, \ldots, N^{(0)}; \; k = n_- + l,$$

$$\xi_{N^{(0)}+1} = 1,$$

and where the vectors η_i^l will turn out to be small. From (5.20) it follows that

$$\xi_i = \prod_{j=i}^{N^{(0)}} \frac{1 - \omega_j}{1 + \omega_j},$$

where $\omega_j = \lambda_k(t_{j+1/2}) h_j / 2\varepsilon$. By (4.4)

$$\left| \frac{1 - \omega_j}{1 + \omega_j} \right| \leq \exp\{-\underline{\alpha} h_j / \rho \varepsilon\},$$

where $\underline{\alpha} = \min_{0 \leq t \leq t_{N^{(0)}+1}} \langle \operatorname{Re} \lambda_k(t) \rangle$, and $\rho = (1 + \sigma)^2 + \sigma^2 \bar{\gamma}^2$ with $\sigma = \bar{h}_- \bar{\alpha} / 2\varepsilon$,

$$\bar{\alpha} = \max_{0 \leq t \leq t_{N^{(0)}+1}} \langle \operatorname{Re} \lambda_k(t) \rangle, \qquad \bar{\gamma}^2 = \max_{0 \leq t \leq t_{N^{(0)}+1}} \left\langle \left(\frac{\operatorname{Im} \lambda_k(t)}{\operatorname{Re} \lambda_k(t)} \right)^2 \right\rangle.$$

Thus

(5.21) $$\xi_i \leq \exp\{-\underline{\alpha}(t_{N^{(0)}+1} - t_i)/\rho \varepsilon\}, \qquad i = 1, \ldots, N^{(0)} + 1.$$

On substitution of (5.19) into the equations defining $_1 W_\Delta^+$, we obtain a system of difference equations for the η_i^l and the lth columns of $P_- \, _1 U_\Delta^+$ and $_1 V_\Delta^+$. This system has homogeneous boundary conditions and an inhomogeneity of size $\varepsilon + \bar{h}_-$. Hence, the stability result (5.10) applied to this system for each l, $l = 1, \ldots, n_+$, yields

(5.22) $$\|\eta_\Delta\|, \|P_{-1} U_\Delta^+\|, \|_1 V_\Delta^+\| \leq c_6 (\bar{h}_- + \varepsilon).$$

This completes the analysis of $_1 W_\Delta$.

   5.2. *The Interval* $[t_{N-N^{(1)}+1}, t_{N+1}]$. The analysis proceeds in the same way as on the first interval. So we only state the results.

   THEOREM 5.2. *If* $h_i \leq \bar{h}_+ \leq K_1 \varepsilon$, $i = N - N^{(1)} + 1, \ldots, N$, *with a suitable constant* K_1, *then the general solution of* (5.1) *on* $[t_{N-N^{(1)}+1}, t_{N+1}]$ *can be written as*

(5.23) $$w_\Delta = {}_3 W_\Delta^- \eta_- + {}_3 W_\Delta^+ \eta_+ + {}_3 W_\Delta^0 \eta_0 + {}_3 w_{p,\Delta},$$

where the ${}_3 W_\Delta^{-,+,0}$ *are structured like the* ${}_1 W_\Delta^{-,+,0}$, *and*

$$H_\Delta \, {}_3 W_\Delta^- = 0, \quad P_- \, {}_3 U_{N-N^{(1)}+1}^- = I, \quad P_+ \, {}_3 U_{N+1}^- = 0, \quad {}_3 V_{N-N^{(1)}+1}^- = 0,$$

$$H_\Delta \, {}_3 W_\Delta^+ = 0, \quad P_- \, {}_3 U_{N-N^{(1)}+1}^+ = 0, \quad P_+ \, {}_3 U_{N+1}^+ = I, \quad {}_3 V_{N-N^{(1)}+1}^+ = 0,$$

$$H_\Delta \, {}_3 W_\Delta^0 = 0, \quad P_- \, {}_3 U_{N-N^{(1)}+1}^0 = {}_3 S_-, \quad P_+ \, {}_3 U_{N+1}^0 = {}_3 S_+, \quad {}_3 V_{N-N^{(1)}+1}^0 = {}_3 S_0,$$

and

$$H_\Delta \, {}_3 w_{p,\Delta} = g_\Delta, \qquad P_- \, {}_3 u_{p,N-N^{(1)}+1} = P_- u_p(t_{N-N^{(1)}+1}),$$

$$P_+ \, {}_3 u_{p,N+1} = P_+ u_p(t_{N+1}), \qquad {}_3 v_{p,N-N^{(1)}+1} = v_p(t_{N-N^{(1)}+1}).$$

Here the matrices $_3S_{-,+,0}$ are defined by

$$_3S_- = P_- U_0(t_{N-N^{(1)}+1}), \quad _3S_+ = P_+ U_0(t_{N+1}), \quad _3S_0 = V_0(t_{N-N^{(1)}+1}),$$

where $U_0(t)$, $V_0(t)$ are the lower and upper parts of $W_0(t)$ defined in Section 2.

For $_3W_\Delta^{-,+,0}$ we have the following estimates corresponding to (5.12), (5.13), (5.15), (5.21) and (5.22):

$$(5.24) \qquad \|_3W_\Delta^0 - \Delta W_0\| \leqslant c_7 \bar{h}_+^2,$$

$$(5.25) \qquad \|_3w_{p,\Delta} - \Delta w_p\| \leqslant c_8 \bar{h}_+^2,$$

$$(5.26) \qquad \|_3W_\Delta^+ - \Delta W_+\| \leqslant c_9 \delta;$$

$$_3W_\Delta^- = \begin{pmatrix} _3U_\Delta^- \\ _3V_\Delta^- \end{pmatrix}, \qquad P_{-\,3}U_\Delta^- = \left\{ \left(u_i^l \right)_{l=1}^{n_-} \right\}_{i=N-N^{(1)}+1}^{N+1}$$

with $u_i^l = \xi_i e_l + \eta_i^l$, $i = N - N^{(1)} + 1, \dots, N + 1$, $\eta_i^l \in R^{n_-}$,

$$(5.27) \qquad |\xi_i| \leqslant \exp\{\underline{\alpha}(t_i - t_{N-N^{(1)}+1})/\rho\varepsilon\}, \qquad i = N - N^{(1)} + 1, \dots, N + 1,$$

$$(5.28) \qquad \|\eta_\Delta\|, \|P_{+\,3}U_\Delta^-\|, \|_3V_\Delta^-\| \leqslant c_{10}(\varepsilon + \bar{h}_+).$$

5.3. *The Interval* $[t_{N^{(0)}+1}, t_{N-N^{(1)}+1}]$. For brevity of notation we set $\underline{i} = N^{(0)} + 1$, $\bar{i} = N - N^{(1)} + 1$. Here it is convenient to work with the difference scheme in its original form (3.2). First we consider the "reduced" problem

$$(5.29a) \qquad 0 = A_{11}(t_{i+1/2}) \frac{y_i + y_{i+1}}{2} + A_{12}(t_{i+1/2}) \frac{z_i + z_{i+1}}{2} + f_i^1,$$

$$(5.29b) \qquad \frac{z_{i+1} - z_i}{h_i} = A_{21}(t_{i+1/2}) \frac{y_i + y_{i+1}}{2} + A_{22}(t_{i+1/2}) \frac{z_i + z_{i+1}}{2} + f_i^2,$$

$$i = \underline{i}, \dots, \bar{i} - 1,$$

with $f_i^1 \in R^n$, $f_i^2 \in R^m$. Substitution of

$$(5.30) \qquad \frac{y_i + y_{i+1}}{2} = -A_{11}^{-1}(t_{i+1/2})\left(A_{12}(t_{i+1/2}) \frac{z_i + z_{i+1}}{2} + f_i^1 \right)$$

into (5.29b) yields

$$(5.31) \qquad \frac{z_{i+1} - z_i}{h_i} = \left(A_{22} - A_{21}A_{11}^{-1}A_{12} \right)(t_{i+1/2}) \frac{z_{i+1} + z_i}{2}$$

$$+ f_i^2 - \left(A_{21}A_{11}^{-1} \right)(t_{i+1/2}) f_i^1, \qquad i = \underline{i}, \dots, \bar{i} - 1.$$

Let $Z_i^j \in R^{m \times m}$ satisfy

$$\frac{Z_{i+1}^j - Z_i^j}{h_i} = \left(A_{22} - A_{21}A_{11}^{-1}A_{12} \right)(t_{i+1/2}) \frac{Z_i^j + Z_{i+1}^j}{2}, \qquad i = j, \dots, \bar{i} - 1,$$

$$Z_j^j = I; \quad j \geqslant \underline{i}.$$

Then the general solution of (5.31) can be written as

$$(5.32) \qquad z_i = Z_i^{\underline{i}} z_{\underline{i}} + \sum_{j=\underline{i}}^{i-1} h_j Z_i^{j+1} \left(I - h_j \left(A_{22} - A_{21}A_{11}^{-1}A_{12} \right)(t_{j+1/2})/2 \right)^{-1}$$

$$\times \left(f_j^2 - \left(A_{21}A_{11}^{-1} \right)(t_{j+1/2}) f_j^1 \right), \qquad i = \underline{i}, \dots, \bar{i}.$$

The general solution of (5.30) is

$$(5.33) \quad y_i = (-1)^{i-\underline{i}} y_{\underline{i}} - 2 \sum_{j=\underline{i}}^{i-1} (-1)^{i-1-j} \left(C(t_{j+1/2}) \frac{z_j + z_{j+1}}{2} + A_{11}^{-1}(t_{j+1/2}) f_j^1 \right)$$

$$= (-1)^{i-\underline{i}} y_{\underline{i}} - C(t_{\underline{i}-1/2}) z_{\underline{i}} + (-1)^{i-\underline{i}} C(t_{i+1/2}) z_i$$

$$+ \sum_{j=\underline{i}+1}^{i-1} (-1)^{i-1-j} \left(C(t_{j-1/2}) - C(t_{j+1/2}) \right) z_j$$

$$- 2 \sum_{j=\underline{i}}^{i-1} (-1)^{i-1-j} A_{11}^{-1}(t_{j+1/2}) f_j^1,$$

where $C = A_{11}^{-1} A_{12}$. The last sum in this expression can be written as

$$(5.34) \quad \sum_{j=\underline{i}}^{i-1} (-1)^{i-1-j} A_{11}^{-1}(t_{j+1/2}) f_j^1$$

$$= \sum_{j=0}^{(i-\underline{i})/2-1} A_{11}^{-1}(t_{\underline{i}+2j+3/2}) \left(f_{\underline{i}+2j+1}^1 - f_{\underline{i}+2j}^1 \right)$$

$$+ \sum_{j=0}^{(i-\underline{i})/2-1} \left(A_{11}^{-1}(t_{\underline{i}+2j+3/2}) - A_{11}^{-1}(t_{\underline{i}+2j+1/2}) \right) f_{\underline{i}+2j}^1$$

in the case when $i - \bar{i}$ is even, and in a similar way when $i - \underline{i}$ is odd.

From (5.32), (5.33) and (5.34) we obtain

LEMMA 5.1. *Provided* $\bar{h} = \max\{ h_i \mid i = \underline{i}, \ldots, \bar{i} - 1 \}$ *is sufficiently small, say* $\bar{h} \leqslant h_0$, *then (5.29) subject to initial conditions specifying* $y_{\underline{i}}, z_{\underline{i}}$ *has a unique solution, and this solution satisfies*

$$(5.35) \quad \|y_\Delta\| \leqslant \|y_{\underline{i}}\| + c_{11} \left(\|z_{\underline{i}}\| + \|f_\Delta^1\| + \|f_\Delta^2\| + \sum_{j=\underline{i}}^{\bar{i}-1} \|f_{j+1}^1 - f_j^1\| \right),$$

$$\|z_\Delta\| \leqslant c_{11} \left(\|z_{\underline{i}}\| + \|f_\Delta^1\| + \|f_\Delta^2\| \right).$$

Now we turn to the "unreduced" problem, with $\varepsilon(y_{i+1} - y_i)/h_i$ replacing 0 on the left-hand side of (5.29a). We decompose the solution of the problem in the form

$$(5.36) \quad \left. \begin{array}{l} y_i = y_i^r + \eta_i \\ z_i = z_i^r + \zeta_i \end{array} \right\}, \qquad i = \underline{i}, \ldots, \bar{i},$$

where y_i^r, z_i^r stand for the solution of the reduced problem using the starting values $y_{\underline{i}}, z_{\underline{i}}$ such that $\eta_{\underline{i}} = 0$, $\xi_{\underline{i}} = 0$. Substitution of (5.36) into the unreduced equations yields

$$(5.37) \quad A_{11}(t_{i+1/2}) \frac{\eta_i + \eta_{i+1}}{2} A_{12}(t_{i+1/2}) \frac{\zeta_i + \zeta_{i+1}}{2}$$

$$= \frac{\varepsilon}{h_i} (\eta_{i+1} - \eta_i) - \frac{\varepsilon}{h_i} (y_{i+1}^r - y_i^r),$$

$$\frac{\zeta_{i+1} - \zeta_i}{h_i} = A_{21}(t_{i+1/2}) \frac{\eta_i + \eta_{i+1}}{2} + A_{22}(t_{i+1/2}) \frac{\zeta_i + \zeta_{i+1}}{2}, \qquad i = \underline{i}, \ldots, \bar{i} - 1,$$

$$\eta_{\underline{i}} = 0, \quad \zeta_{\underline{i}} = 0.$$

Applying Lemma 5.1 to (5.37), we obtain

$$\|\eta_\Delta\| \leqslant c_{11}\left(\frac{2\varepsilon}{h_{\min}} + 4\varepsilon \sum_{j=\underline{i}}^{\overline{i}-1} h_j^{-1}\right)(\|\eta_\Delta\| + \|y_\Delta^r\|),$$

$$\|\zeta_\Delta\| \leqslant c_{11}\left(\frac{2\varepsilon}{h_{\min}}\right)(\|\eta_\Delta\| + \|y_\Delta^r\|),$$

where $h_{\min} = \min\{h_i \,|\, i = \underline{i}, \ldots, \overline{i} - 1\}$. Hence if

$$(5.38) \qquad \varepsilon c_{11}\left(\frac{2}{h_{\min}} + 4\sum_{j=\underline{i}}^{\overline{i}-1} h_j^{-1}\right) \leqslant K < 1,$$

which we will assume from now on, then, by the contraction principle, (5.37) has a unique solution which satisfies

$$(5.39) \qquad \|\eta_\Delta\| \leqslant \frac{K}{1-K}\|y_\Delta^r\|, \qquad \|\zeta_\Delta\| \leqslant \frac{K}{1-K}\|y_\Delta^r\|.$$

We now derive a representation of the general solution of (3.2) for $i = \underline{i}, \ldots, \overline{i} - 1$. For ease of presentation we abbreviate (3.2) as

$$H_\Delta x_\Delta = f_\Delta.$$

A particular solution $_2 x_{p,\Delta}$ is defined by

$$(5.40) \qquad H_\Delta \,_2 x_{p,\Delta} = f_\Delta, \qquad _2 x_{p,\underline{i}} = x_p(t_{\overline{i}}^-),$$

where $x_p(t)$ is the particular solution obtained from $w_p(t)$ by inverting the transformations (2.8) and (2.5). Let $e_\Delta = \,_2 x_{p,\Delta} - \Delta x_p$. Then e_Δ satisfies $H_\Delta e_\Delta = l_\Delta$, $e_{\underline{i}} = 0$, where $l_\Delta = \{l_{\underline{i}}, \ldots, l_{\overline{i}-1}\}$ represents the local truncation error of $x_p(t)$. It is well known that

$$l_i = \varphi(t_{i+1/2})h_i^2 + O(h_i^4)$$

for some smooth function $\varphi(t)$. Hence it follows from Lemma 5.1 and (5.39) that

$$(5.41) \qquad \|e_\Delta\| \leqslant c_{12}\overline{h}.$$

If the grid is locally almost uniform, i.e.

$$(5.42) \qquad h_{i+1} = h_i(1 + O(h_i)),$$

then $l_{i+1} - l_i = O(h_i^3)$, and from Lemma 5.1 plus (5.39) we obtain

$$(5.43) \qquad \|e_\Delta\| \leqslant c_{13}\overline{h}^2.$$

Both estimates, (5.41) and (5.43), are sharp, as will be shown by a numerical example in Section 6.

The general solution of the homogeneous discrete problem can be written as

$$(5.44) \qquad x_\Delta = X_\Delta^0 \gamma_0 + X_\Delta \gamma, \qquad \gamma_0 \in R^m, \gamma \in R^n,$$

where

$$X_\Delta^0 = \begin{pmatrix} Y_\Delta^0 \\ Z_\Delta^0 \end{pmatrix}, \qquad X_\Delta = \begin{pmatrix} Y_\Delta \\ Z_\Delta \end{pmatrix}$$

and

$$H_\Delta X_\Delta^0 = 0, \quad Y_i^0 = Y_0(t_i), \quad Z_i^0 = Z_0(t_i),$$
$$H_\Delta X_\Delta = 0, \quad Y_i = I, \quad Z_i = 0.$$

Here $Y_0(t)$, $Z_0(t)$ are obtained from $U_0(t)$, $V_0(t)$ by inverting (2.8) and (2.5). The same argument as used for the particular solution yields

(5.45)
$$\|X_\Delta^0 - \Delta X_0\| \leqslant c_{14}\bar{h}$$

and

(5.46)
$$\|X_\Delta^0 - \Delta X_0\| \leqslant c_{15}\bar{h}^2$$

if (5.42) is satisfied.

Applying the same arguments as in the derivation of (5.39) we obtain

(5.47)
$$Y_i = (-1)^{i-\bar{i}}I + \Psi_i, \quad i = \underline{i},\ldots,\bar{i},$$

with

(5.48a)
$$\|\Psi_\Delta\| \leqslant K/(1 - K)$$

and

(5.48b)
$$\|Z_\Delta\| \leqslant K/(1 - K).$$

So up to a perturbation of size $K/(1 - K)$ we have

(5.49)
$$X_i = \begin{pmatrix} (-1)^{i-\bar{i}}I \\ 0 \end{pmatrix}, \quad i = \underline{i},\ldots,\bar{i}.$$

REMARK. A key assumption for the analysis of this subsection was (5.38). If a quasiuniform family of grids is used on $[t_i, t_{\bar{i}}]$, then (5.38) is equivalent to requiring that $\varepsilon \leqslant \text{const } h_{\min}^2$ with an appropriate constant of size one. For such a quasiuniform family the analysis of this subsection can be extended to the case $\varepsilon \geqslant \text{const } h_{\min}^2$.

5.4. *Combining the Representations of the General Solution.* Here we consider the discrete boundary value problem consisting of Eqs. (3.9) for $i = 1,\ldots, N$ and the boundary conditions

(5.50)
$$P_-u_1 = \eta_-, \quad P_+u_N = \eta_+, \quad v_1 = \eta_0.$$

This is the counterpart of problem (2.9) subject to the boundary conditions (2.11), which is well posed, according to Theorem 2.1. We shall show that the discrete problem has a unique solution and that this solution approximates the solution of (2.9), (2.11) provided δ, $\bar{h}$ and K are sufficiently small.

On each of the three subintervals we have a different representation of the general solution of the difference scheme:

(5.51)
$$w_\Delta^1 = {}_1W_\Delta^-\eta_-^1 + {}_1W_\Delta^+\eta_+^1 + {}_1W_\Delta^0\eta_0^1 + {}_1w_{p,\Delta}$$

on $[t_1, t_{\underline{i}}]$, with $\eta_-^1 \in R^{n_-}$, $\eta_+^1 \in R^{n_+}$, $\eta_0^1 \in R^m$,

(5.52)
$$w_\Delta^2 = {}_2W_\Delta\eta^2 + {}_2W_\Delta^0\eta_0^2 + {}_2w_{p,\Delta}$$

on $[t_{\underline{i}}, t_{\bar{i}}]$, with $\eta^2 \in R^n, \eta_0^2 \in R^m$, and

(5.53)
$$w_\Delta^3 = {}_3W_\Delta^-\eta_-^3 + {}_3W_\Delta^+\eta_+^3 + {}_3W_\Delta^0\eta_0^3 + {}_3w_{p,\Delta}$$

on $[t_i^-, t_{N+1}]$, with $\eta_-^3 \in R^{n_-}$, $\eta_+^3 \in R^{n_+}$ and $\eta_0^3 \in R^m$. Here $_1W_\Delta^{-,+,0}$, $_3W_\Delta^{-,+,0}$ are as in Theorems 5.1, 5.2 and $_2W_\Delta$, $_2W_\Delta^0$ are obtained from X_Δ, X_Δ^0 via the transformation (3.5).

We now determine the parameter vectors $\eta_{-,+,0}^1$, η^2, η_0^2, $\eta_{-,+,0}^3$ such that

(5.54a)
$$P_- u_1^1 = \eta_-,$$

(5.54b)
$$v_1^1 = \eta_0,$$

(5.54c)
$$w_i^1 = w_i^2,$$

(5.54d)
$$w_i^2 = w_i^3,$$

(5.54e)
$$P_+ u_{N+1}^3 = \eta_+.$$

Equations (5.54) are a system of $3 \times (n + m)$ linear equations for the $3 \times (n + m)$ components of the unknown vectors. We shall show that this system has a uniformly bounded inverse if δ, $\bar{h}$ and K are sufficiently small. Written out explicitly, Eqs. (5.54) become

(5.55a)
$$P_{-\,1}U_1^- \eta_-^1 + P_{-\,1}U_1^+ \eta_+^1 + P_{-\,1}U_1^0 \eta_0^1 = \eta_- - P_{-\,1}u_{p,1},$$

(5.55b)
$$_1V_1^- \eta_-^1 + _1V_1^+ \eta_+^1 + _1V_1^0 \eta_0^1 = \eta_0 - _1v_{p,1},$$

(5.55c)
$$_1W_i^- \eta_-^1 + _1W_i^+ \eta_+^1 + _1W_i^0 \eta_0^1 - _2W_i \eta^2 - _2W_i^0 \eta_0^2 = _2w_{p,i} - _1w_{p,i},$$

(5.55d)
$$_2W_i \eta^2 + _2W_i^0 \eta_0^2 - _3W_i^- \eta_-^3 - _3W_i^+ \eta_+^3 - _3W_i^0 \eta_0^3 = _3w_{p,i} - _2w_{p,i},$$

(5.55e)
$$P_{+\,3}U_{N+1}^- \eta_-^3 + P_{+\,3}U_{N+1}^+ \eta_+^3 + P_{+\,3}U_{N+1}^0 \eta_0^3 = \eta_+ - P_{+\,3}u_{p,N+1}.$$

Most of the matrices in (5.55) have some special structure which is now discussed for each equation.

Structure in (5.55a). By definition,
$$P_{-\,1}U_1^- = I, \quad P_{-\,1}U_1^+ = 0, \quad P_{-\,1}U_1^0 = _1S_-.$$

Structure in (5.55b). By definition,
$$_1V_1^- = 0, \quad _1V_1^+ = 0, \quad _1V_1^0 = I.$$

Structure in (5.55c). By (2.17) and (5.15),
$$_1W_i^- = \begin{pmatrix} \exp(\Lambda_-(0)t_i/\varepsilon) \\ 0 \end{pmatrix} + O(\delta) = O(\delta).$$

By (5.22),
$$_1W_i^+ = \begin{pmatrix} O_{n_- \times n_+} \\ I_{n_+ \times n_+} \\ O_{m \times n_+} \end{pmatrix} + O(\varepsilon + \bar{h}_-),$$

and by (5.12)
$$_1W_i^0 = W^0(t_i) + O(\bar{h}_-^2).$$

By (5.47), (5.48), on using (3.5),
$$_2W_i = \begin{pmatrix} E^{-1}(t_i) \\ 0 \end{pmatrix} + O(K).$$

By (5.45), (5.46), on using (3.5),

$$_2W_l^0 = W_0(t_l) + O(\bar{h}^l), \qquad l = 1 \text{ or } 2.$$

Structure in (5.55d). By (5.47), (5.48), on using (3.5),

$$_2W_i^- = \begin{pmatrix} (-1)^{i-l}E^{-1}(t_i^-) \\ 0 \end{pmatrix} + O(K).$$

By (5.45), (5.46), on using (3.5),

$$_2W_i^0 = W_0(t_i^-) + O(\bar{h}^l), \qquad l = 1 \text{ or } 2.$$

By (5.28),

$$_3W_i^- = \begin{pmatrix} I_{n_-\times n_-} \\ O_{n_+\times n_-} \\ O_{m\times n_-} \end{pmatrix} + O(\varepsilon + \bar{h}_+).$$

By (2.23) and (5.26),

$$_3W_i^+ = \begin{pmatrix} 0 \\ \exp(\Lambda_+(1)(t_i^- - 1)/\varepsilon) \\ 0 \end{pmatrix} + O(\delta) = O(\delta).$$

By (5.24),

$$_3W_i^0 = W_0(t_i^-) + O(\bar{h}_+^2).$$

Structure in (5.55e). By definition,

$$P_+\,_3U_{N+1}^- = 0, \quad P_+\,_3U_{N+1}^+ = I, \quad P_+\,_3U_{N+1}^0 =\,_3S_+.$$

Hence, up to a small perturbation, the matrix of the linear system (5.55) has the form indicated in Figure 5.1, where R stands for different nonsingular matrices, rectangular matrices are indicated by S, and where the vector of unknowns is $(\eta_-^1, \eta_+^1, \eta_0^1, \eta^2, \eta_0^2, \eta_-^3, \eta_+^3, \eta_0^3)$.
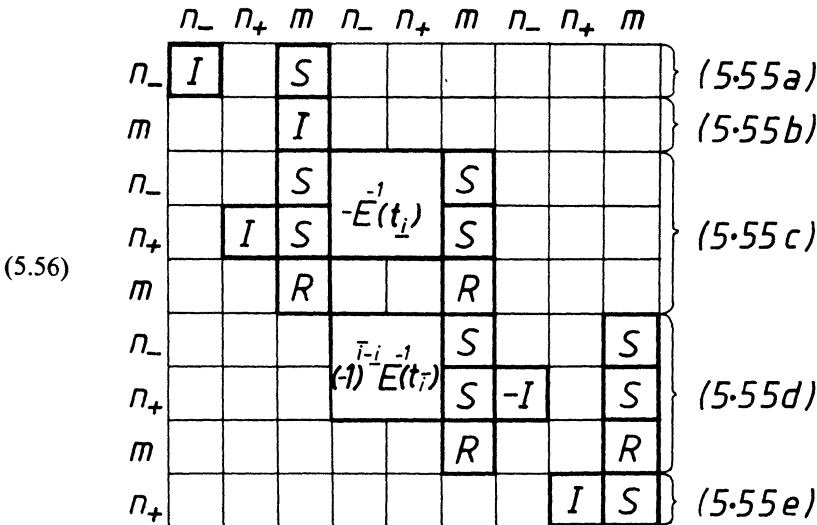

FIGURE 5.1. *The essential part of the matrix in* (5.55)

In the matrix (5.56) blocks which are not indicated explicitly are zero. The size of each block in (5.56) is determined by the corresponding entry in the first row and column of Figure 5.1. The correspondence between (5.56) and the splitting (5.55a−e) is indicated in the last column of Figure 5.1.

Using elementary row and column operations it is easy to see that the matrix (5.56) is nonsingular if and only if the $n \times n$ matrix

$$\begin{pmatrix} P_- E^{-1}(t_i) \\ P_+ E^{-1}(t_i^-) \end{pmatrix}$$

is nonsingular. As $t_i = O(\varepsilon \ln \delta)$, $1 - t_i^- = O(\varepsilon \ln \delta)$, this matrix has a uniformly bounded inverse for all relevant values of t_i, t_i^- if and only if

(5.57)
$$\det \begin{pmatrix} P_- E^{-1}(0) \\ P_+ E^{-1}(1) \end{pmatrix} \neq 0.$$

In this case, the full problem (5.55) also has a uniformly bounded inverse when δ, $\bar{h}$ and K are sufficiently small.

We shall now show that the solution of (5.55) generates an $O(\delta + \bar{h})$ approximation to the solution of the corresponding continuous problem, uniformly on the interval $[0, 1]$. According to Theorem 2.3 the solution of the continuous problem is

(5.58)
$$w = W_- \gamma_- + W_+ \gamma_+ + W_0 \gamma_0 + w_p,$$

where γ_-, γ_+ and γ_0 are uniquely determined through

(5.59)
$$\begin{cases} P_- U_-(0)\gamma_- + P_- U_+(0)\gamma_+ + P_- U_0(0)\gamma_0 = \eta_- - P_- u_p(0), \\ V_-(0)\gamma_- + V_+(0)\gamma_+ + V_0(0)\gamma_0 = \eta_0 - v_p(0), \\ P_+ U_-(1)\gamma_- + P_+ U_+(1)\gamma_+ + P_+ U_0(1)\gamma_0 = \eta_+ - P_+ u_p(1). \end{cases}$$

With the aid of γ_-, γ_+ and γ_0 determined in this way we define a vector

(5.60)
$$\rho = (\gamma_-, 0, \gamma_0, 0, \gamma_0, 0, \gamma_+, \gamma_0) \in R^{3(n+m)}.$$

It is easily verified with the aid of (5.59) and the approximation results for the $_k W_\Delta^{\cdot,0}$ and the particular solution that the vector ρ satisfies (5.55) up to a residual vector of size $O(\delta) + O(\bar{h}^l)$, where $l = 1$ or 2 depending on the coarse grid. Hence, denoting the solution vector of (5.55) by η, it follows that

(5.61)
$$\|\rho - \eta\| = O(\delta) + O(\bar{h}^l).$$

Combining the estimate (5.61) with the various approximation results for the $_k W_\Delta^{+,-,0}$ and for the particular solution, we obtain on comparing (5.58) with (5.51), (5.52) and (5.53):

LEMMA 5.2. *The problem* (3.9) *subject to the boundary conditions* (5.50) *has a unique solution* w_Δ *provided* δ, $\bar{h}$ *and* K *are sufficiently small and* (5.57) *holds. This solution satisfies*

$$\|w_\Delta - \Delta w\| \leq c_{16}(\delta + \bar{h}^l),$$

where $l = 1$ *or* 2, *depending on the coarse grid.*

Using the inverse transformation of (3.5) and the standard procedure of extending convergence results valid for a special set of linear boundary conditions to the case of general linear boundary conditions, as outlined for the continuous problem in Section 2, we obtain the principal result of this section.

THEOREM 5.3. *Assume that the boundary value problem* (2.1), (2.19) *is well posed uniformly in* ε, *for* $0 < \varepsilon \leqslant \varepsilon_0$, *and let* x *be its solution. Choose a grid as described in subsections* 5.1, 5.2 *and* 5.3, *and let the quantities* δ, $\bar{h}$ *and* K *characterizing this grid be small enough, i.e.*

$$0 < \delta \leqslant \delta_0, \quad 0 < \bar{h} \leqslant \bar{h}_0, \quad 0 < K \leqslant K_0,$$

for suitable constants δ_0, $\bar{h}_0$, K_0. *Also assume that* (5.57) *holds. Then the box scheme* (3.2), (3.3) *has a unique solution* $x_\Delta = (y_\Delta, z_\Delta)$ *which satisfies*

$$(5.62) \qquad\qquad \|x_\Delta - \Delta x\| \leqslant c_{17}(\delta + \bar{h}^l),$$

where $l = 2$ *if the coarse grid belongs to a family of locally almost uniform grids, and* $l = 1$ *otherwise.*

6. The Trapezoidal Scheme. The trapezoidal scheme (3.4) (or (3.9)) is analyzed in very much the same way as the box scheme: separate treatment of the fine grids and the coarse grid, and patching of the different representations of the general solution of the difference equations. There are only minor technical differences on the fine grids which we shall not dwell upon. Essentially all results of subsections 5.1 and 5.2 carry over to the trapezoidal scheme.

On the coarse grid however the global truncation error of the trapezoidal scheme differs substantially from that of the box scheme, as will be borne out by the subsequent analysis. We consider the problem

$$(6.1) \qquad \varepsilon \frac{y_{i+1} - y_i}{h_i} = \frac{A_{11}(t_i)y_i + A_{11}(t_{i+1})y_{i+1}}{2}$$

$$+ \frac{A_{12}(t_i)z_i + A_{12}(t_{i+1})z_{i+1}}{2} + \frac{f_1(t_i) + f_1(t_{i+1})}{2},$$

$$\frac{z_{i+1} - z_i}{h_i} = \frac{A_{21}(t_i)y_i + A_{21}(t_{i+1})y_{i+1}}{2}$$

$$+ \frac{A_{22}(t_i)z_i + A_{22}(t_{i+1})z_{i+1}}{2} + \frac{f_2(t_i) + f_2(t_{i+1})}{2}, \qquad i = \underline{i}, \ldots, \bar{i} - 1,$$

with $y_{\underline{i}}$, $z_{\underline{i}}$ given. By a slight variation of the analysis of the box scheme we obtain analogs to all results from Lemma 5.1 up to (5.39). We write (6.1) as

$$H_\Delta x_\Delta = f_\Delta$$

and define a particular solution $_2x_{p,\Delta}$ of this problem by

$$H_{\Delta \, 2}x_{p,\Delta} = f_\Delta, \qquad _2x_{p,\underline{i}} = x_p(t_{\underline{i}}),$$

as we did for the box scheme. The global truncation error $e_\Delta = {}_2x_{p,\Delta} - \Delta x_p$ satisfies

$$H_\Delta e_\Delta = l_\Delta, \qquad e_{\underline{i}} = 0,$$

where $l_\Delta = \{l_i, \ldots, l_{\bar{i}-1}\}$ is the local truncation error of $x_p(t)$. Now this local truncation error differs markedly from that of the box scheme. It has the structure

$$l_i = h_i^2 \begin{pmatrix} \dfrac{\varepsilon y'''}{12}(t_{i+1/2}) \\[2mm] \dfrac{z'''}{12}(t_{i+1/2}) \end{pmatrix} + \begin{pmatrix} \varepsilon O(h_i^4) \\[2mm] O(h_i^4) \end{pmatrix},$$

Hence, as $\varepsilon \ll \bar{h}$, it follows that

(6.2) $\|e_\Delta\| \leqslant c_{19}\bar{h}^2.$

This estimate is to be compared with the estimates (5.41), (5.43) for the box scheme. Note that when there is no unperturbed component z in (2.1), then (6.2) can be replaced by

(6.3) $\|e_\Delta\| \leqslant c_{20}\varepsilon\bar{h}^l,$

where $l = 2$ if the grid is locally almost uniform, and $l = 1$ otherwise.

There are corresponding results regarding the approximation of $X_0(t)$.

The remaining analysis of the trapezoidal scheme proceeds as for the box scheme leading to the counterpart of Theorem 5.3, with the estimate (5.62) replaced by

(6.4) $\|x_\Delta - \Delta x\| = c_{22}(\delta + \bar{h}^2),$

or

(6.5) $\|x_\Delta - \Delta x\| \leqslant c_{22}(\delta + \varepsilon\bar{h}^l),$ $l = 1$ or $2,$

in case (6.3) applies.

7. Numerical Results. To provide numerical evidence for the theory of Sections 5 and 6 we present some results for the problem.

(7.1) $\varepsilon y' = A(t; \lambda)y + f(t, \varepsilon; \lambda),$ $0 \leqslant t \leqslant 1,$

(7.2) $B_0 y(0) + B_1 y(1) = \beta,$

where $y(t) = (y_1(t), y_2(t))^T$, λ is a real parameter,

(7.3) $A(t; \lambda) = E(t; \lambda)\begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} E^{-1}(t, \lambda),$

with

$$E(t, \lambda) = E^{-1}(t, \lambda) = \begin{bmatrix} \sin \lambda t & \cos \lambda t \\ \cos \lambda t & -\sin \lambda t \end{bmatrix},$$

and

$$B_0 = \begin{pmatrix} P_- E^{-1}(0; \lambda) \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad B_1 = \begin{pmatrix} 0 \\ P_+ E^{-1}(1; \lambda) \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \cos \lambda & -\sin \lambda \end{pmatrix}.$$

Equation (7.1) can be solved explicitly. Introducing the new variable $u(t) = E^{-1}(t)y$, the homogeneous problem $\varepsilon y' - A(t, \lambda)y = 0$ becomes

$$\varepsilon u' = \begin{pmatrix} -1 & \lambda\varepsilon \\ -\lambda\varepsilon & 2 \end{pmatrix} u.$$

A fundamental solution matrix of this system is

$$\Phi(t) = \begin{pmatrix} \varepsilon \exp\{(t-1)(2+\varepsilon\alpha)/\varepsilon\} & -\dfrac{\lambda\varepsilon}{\alpha}\exp\{-t(1+\varepsilon\alpha)/\varepsilon\} \\[2ex] -\dfrac{\lambda\varepsilon}{\alpha}\exp\{(t-1)(2+\varepsilon\alpha)/\varepsilon\} & \varepsilon \exp\{-t(1+\varepsilon\alpha)/\varepsilon\} \end{pmatrix},$$

where

$$\alpha = \frac{1}{2\varepsilon}\left[-3 + \sqrt{9 - 4\varepsilon^2\lambda^2}\,\right] = -\frac{1}{3}\varepsilon\lambda^2 - \frac{1}{27}\varepsilon^3\lambda^4 + O(\varepsilon^5),$$

so the general solution of (7.1) is

$$(7.4) \qquad\qquad y(t) = E(t)\Phi(t)s + y_p(t),$$

where $y_p(t)$ is a particular solution of (7.1), and $s = (s_1, s_2)^T$. Substitution of (7.4) into (7.2) leads to a linear system for s, which is well conditioned provided that $|\lambda| \gg \varepsilon$.

Note that the matrix

$$\begin{pmatrix} P_- E^{-1}(0; \lambda) \\ P_+ E^{-1}(1; \lambda) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \cos\lambda & -\sin\lambda \end{pmatrix}$$

is singular when $\lambda = \pi(k + 1/2)$, $k = 0, \pm 1, \dots$, so that condition (5.57) is not satisfied for these values of λ, and hence Theorem 5.3 and the estimates (6.4), (6.5) for the trapezoidal scheme do not apply.

We now report computational results for the values $\lambda = \pi/4$ and $\lambda = \pi/2$.

7.1. *The Case* $\lambda = \pi/4$. The fine grids were constructed according to (5.17) with c_4 set equal to 1. The fine grid at the left endpoint ends once $\exp(-t_{N^{(0)}+1}/\varepsilon) \leqslant \delta$, and that at the right endpoint ends when $\exp(2(t_{N-N^{(1)}+1} - 1)/\varepsilon) \leqslant \delta$. The coarse grids were either chosen to be uniform or of the form $h_i = h$, $i = N^{(0)} + 1, N^{(0)} + 3, \dots$, $h_i = h/2$, $i = N^{(0)} + 2, N^{(0)} + 4, \dots$, which is *not* locally almost uniform.

We take $\delta = 10^{-6}$, which, for all values of ε considered, leads to fine grids with a total number of 1164 points. The coarse grids are then obtained by inserting $I = 9, 19, 39, \dots$ points between the two endpoints of the fine grids. δ was taken so small in order to be able to verify the rates of convergence of the schemes when I is increased.

The forcing term $f(t, \varepsilon; \lambda)$ in (7.1) is chosen such that $y_p(t) = (e^t, e^{-t})$ is a particular solution of (7.1), i.e. $f(t, \varepsilon; \lambda) = \varepsilon y_p(t) - A(t, \lambda)y_p$. The vector β was taken to be

$$\beta_1 = 3, \qquad \beta_2 = \cos(\lambda)\exp(1) - \sin(\lambda)\exp(-1) + 1,$$

so that y has layers at both ends.

In the following tables the values of

$$e_l = \max_{1 \leqslant j \leqslant N+1} |y_l(t_j) - y_{l,j}|, \qquad l = 1, 2,$$

i.e. the absolute values of the maximal error in the two components, are given for specific values of ε and I. The maximum is always obtained on the coarse grid. Away from the coarse grid towards the left and the right endpoints, the errors decrease rapidly and after a fairly small number of gridpoints they become of size δ.

Table 7.1 contains the errors obtained by the box scheme for $\varepsilon = 10^{-6}$ using uniform coarse grids. The results for the same ε and nonuniform coarse grids are given in Table 7.2.

TABLE 7.1

I	9	19	39
e_1	7.25 E–3	1.85 E–3	4.67 E–4
e_2	2.38 E–3	6.09 E–4	1.54 E–4

TABLE 7.2

I	9	19	39
e_1	1.55 E–2	8.39 E–3	4.36 E–3
e_2	1.16 E–2	5.54 E–3	2.70 E–3

We observe convergence of order 2 for the uniform grids and convergence of order 1 for the others. This agrees with Theorem 5.3.

In Table 7.3 we list the values of e_2 obtained by the trapezoidal scheme for different ε using uniform coarse grids. The analogous entries for the nonuniform grids are given in Table 7.4.

TABLE 7.3

ε \ I	9	19	39
1. E–2	9.22 E–6	2.11 E–6	1.96 E–6
5. E–3	7.45 E–6	1.96 E–6	1.96 E–6
2.5 E–3	4.64 E–6	1.96 E–6	1.96 E–6
1.25 E–3	2.60 E–6	1.96 E–6	1.96 E–6
6.25 E–4	1.96 E–6	1.96 E–6	1.96 E–6

TABLE 7.4

ε \ I	9	19	39
1. E–2	1.18 E–5	2.80 E–6	1.96 E–6
5. E–3	1.72 E–5	2.85 E–6	1.96 E–6
2.5 E–3	1.51 E–5	3.12 E–6	1.96 E–6
1.25 E–3	1.11 E–5	2.93 E–6	1.96 E–6
6.25 E–4	6.81 E–6	2.26 E–6	1.96 E–6

The entry 1.96 E–6 in Table 7.3 is the maximal error in the right layer, and it pollutes the whole interval [0, 1]. Whenever this entry occurs in Table 7.3, it means that the error due to the discretization on the coarse grid lies below 1.96 E–6. The entries in the first row and column support the estimate (6.4). The values of e_1 behave like those of e_2, and are therefore not given.

RICHARD WEISS

The entry 1.96 in Table 7.4 occurs for the same reason as above. In the first column of the table we observe an error proportional to ε, while not much can be inferred about the dependence on I. In our computer environment we cannot decrease δ further.

7.2. *The Case* $\lambda = \pi/2$. We have analyzed the performance of the two schemes in the situation when the matrix

$$\begin{bmatrix} P_- E^{-1}(0) \\ P_+ E^{-1}(1) \end{bmatrix}$$

is singular, for systems having no z-component and for meshes having a uniformly spaced coarse part. Assuming that all conditions of Theorem 5.3, except for (5.57), are satisfied, we obtained the bounds for the global truncation error

$$(7.5) \qquad \|y_\Delta - \Delta y\| \leq c_{23}(\delta + h^2)\left(1 + \frac{h^2}{\varepsilon}\right)$$

for the box scheme, and

$$(7.6) \qquad \|y_\Delta - \Delta y\| \leq c_{24}(\delta + \varepsilon h^2)\left(1 + \frac{h^2}{\varepsilon}\right)$$

for the trapezoidal scheme, respectively.

Table 7.5 contains the results for the box scheme applied to the previous problem with $\delta = 10^{-6}$ and $I = 9$ for different values of ε.

TABLE 7.5

ε	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
e_1	1.73 E–3	5.54 E–3	3.68 E–2	3.49 E–1	3.47 E 0	3.47 E + 1
e_2	8.39 E–2	2.12 E–3	2.26 E–3	2.28 E–3	2.28 E–3	2.27 E–3

The behavior of e_1 supports the validity of (7.5). That of e_2 can be explained by looking at the particular structure of the problem (7.1), (7.2).

Table 7.6 contains the analogous entries for the trapezoidal scheme.

TABLE 7.6

ε	1. E–2	5. E–3	2.5 E–3	1.12 E–3	6.25 E–4	3.125 E–4
e_1	1.04 E–5	7.76 E–6	5.26 E–6	3.85 E–6	3.14 E–6	2.79 E–6
e_2	6.53 E–6	4.49 E–6	2.87 E–6	2.32 E–6	2.03 E–6	1.96 E–6

In this table the errors do not behave linearly in ε, as they did for $\lambda = \pi/4$. Further agreement with (7.6) is not apparent, and there is a need for more analysis.

The computations were done on the CDC Cyber 174 of the Technical University of Vienna, using single precision (14 digits). The code SOLVEBLOK of de Boor and Weiss [5] was used in the implementation of the schemes. Due to the implicit scaling feature of SOLVEBLOK, no conditioning problems were encountered even when ε was very small.

Acknowledgement. I wish to thank Uri Ascher for many valuable discussions.

Institut für Angewandte und Numerische Mathematik
Technische Universität Wien
Gusshausstrasse 27-29
A-1040, Austria

1. L. R. ABRAHAMSSON, *Numerical Solution of a Boundary Value Problem in Semiconductor Physics*, Report #81, Department of Computer Science, Uppsala University, 1979.

2. L. R. ABRAHAMSSON, H. B. KELLER & H. O. KREISS, "Difference approximations for singular perturbations of systems of ordinary differential equations," *Numer. Math.*, v. 22, 1974, pp. 367 − 391.

3. U. ASCHER, I. CHRISTIANSEN & R. D. RUSSELL, "A collocation solver for mixed order systems of boundary value problems," *Math. Comp.*, v. 33, 1979, pp. 659 − 679.

4. U. ASCHER & R. WEISS, *Collocation for Singular Perturbation Problems* I: *First Order Systems With Constant Coefficients*, Report 81-2, Department of Computer Science, The University of British Columbia, 1981.

5. C. DE BOOR & R. WEISS, "SOLVEBLOK: A package for solving almost block diagonal linear systems, with application to spline approximation and the numerical solution of ordinary differential equations," *ACM Trans. Math. Software*, v. 6, 1980, pp. 80 − 91.

6. H. O. KREISS & N. NICHOLS, *Numerical Methods for Singular Perturbation Problems*, Report #57, Department of Computer Science, Uppsala University, 1975.

7. M. MAIER, *Die numerische Lösung von Halbleitermodellen mit Hilfe des Kollokationsverfahrens PITHOP unter Verwendung einer automatischen Schrittweitenkontrolle*, Master's thesis, Department of Mathematics, Technical University München, 1979.

8. R. E. O'MALLEY, JR., *Introduction to Singular Perturbations*, Academic Press, New York, 1974.

9. CH. RINGHOFER, *On Collocation Schemes for Quasilinear Singularly Perturbed Boundary Value Problems*, Doctoral thesis, Department of Applied Mathematics and Numerical Analysis, Technical University of Vienna, 1981.