THE NUMERICAL SOLUTION OF BOUNDARY VALUE PROBLEMS FOR SECOND ORDER FUNCTIONAL DIFFERENTIAL EQUATIONS BY FINITE DIFFERENCES

by

Colin W. Cryer

Computer Sciences Technical Report #127

June 1971
<table>
<thead>
<tr>
<th>CONTENTS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2. Preliminaries</td>
<td>4</td>
</tr>
<tr>
<td>3. LU factorization of Jacobi matrices</td>
<td>9</td>
</tr>
<tr>
<td>4. Properties of $I_\alpha$</td>
<td>12</td>
</tr>
<tr>
<td>5. Perturbations of monotone matrices</td>
<td>18</td>
</tr>
<tr>
<td>6. The numerical method</td>
<td>23</td>
</tr>
<tr>
<td>7. A numerical example</td>
<td>31</td>
</tr>
<tr>
<td>Appendix A. Further properties of $I_\alpha$</td>
<td>40</td>
</tr>
<tr>
<td>Appendix B. The program NEVERS</td>
<td>51</td>
</tr>
<tr>
<td>References</td>
<td>56</td>
</tr>
</tbody>
</table>
THE NUMERICAL SOLUTION OF BOUNDARY VALUE PROBLEMS FOR SECOND ORDER FUNCTIONAL DIFFERENTIAL EQUATIONS BY FINITE DIFFERENCES

by

Colin W. Cryer*

1. INTRODUCTION

In the present paper we consider numerical methods for computing the solution of the boundary value problem

\[
\begin{align*}
\ddot{x}(t) &= g(t,x(t)) + (\gamma x)(t), \quad 0 < t < 1, \\
x(0) &= x(1) = 0,
\end{align*}
\]

(1.1)

where, \( \gamma : \mathbb{R}^1 \to \mathbb{R}^1 \), and \( g : \mathbb{R}^2 \to \mathbb{R}^1 \).

It will be assumed that (1.1) has a unique twice continuously differentiable solution which will be denoted by \( x \) throughout the paper. It will also be assumed that \( g \) is continuously differentiable and that

\[
\frac{\partial g(t,y)}{\partial y} \geq \beta > -\pi^2, \tag{1.2}
\]

for \( t \in [0,1] \) and all \( y \in \mathbb{R}^1 \). No explicit assumptions about \( \gamma \) will be made. However, it will be assumed that \( \gamma \) can be approximated (in a sense explained later) by a Lipschitz continuous mapping \( \gamma_h \).

Examples of boundary value problems which can easily be cast into the form (1.1) are:

---

* Sponsored by the Mathematics Research Center, United States Army, Madison, Wisconsin, under Contract No.: DA-31-124-ARO-D-462, and the Office of Naval Research under Contract No.: N00014-67-A-0128-0004. The computations were supported by the University of Wisconsin Grants Committee.
1. The two-point boundary value problem

\[
\begin{align*}
\ddot{y}(s) &= g(s, y(s)), \ a < s < b, \\
y(a) &= y(b) = 0;
\end{align*}
\]  

(1.3)

2. The boundary value problem

\[
\begin{align*}
\ddot{y}(s) &= \frac{1}{16} \sin y(s) + s - (s+1)y(s-1), \ 0 < s < 2, \\
y(s) &= s - \frac{1}{2}, \ \text{if} \ s \leq 0, \\
y(2) &= -\frac{1}{2},
\end{align*}
\]  

(1.4)

which is a special case of problems considered by Nevers and Schmitt [18];

3. The integro-differential equation

\[
\ddot{y}(s) = g(s) + \int_a^b y(u) f(s, du), \ a < s < b,
\]  

(1.5)

where the integral is a Stieltjes integral, and where \( f: \mathbb{R}^2 \rightarrow \mathbb{R}^1 \) is of bounded variation.

Concerning the theory of boundary value problems for functional differential equations see Cooke [2], El'sgol'ts [4], Fennell and Waltman [5, 6], Grimm and Schmitt [10, 11], Halanay [12], Halanay and Yorke [13], Hale [14], Kato [16], Norkin [19], Schmitt [21]. The numerical solution of boundary value problems for delay differential equations has been considered by Nevers and Schmitt [18], while the numerical solution of initial value problems for functional differential equations has been treated by Cryer and Tavernini [3], and Tavernini [22, 23].
The numerical solution of the two-point boundary value problem (1.3) by finite differences has been extensively studied (see, for example, Ciarlet et al [1], Henrici [15], and Keller [17]), and this work has guided the present paper.

Acknowledgements

We have benefited from the stimulus provided by discussions with Professor Lucio Tavernini.
2. PRELIMINARIES

Throughout the paper matrices and vectors will be understood to be \( n \times n \) matrices and \( n \)-vectors, respectively.

We set \( h = \frac{1}{(n+1)} \). It will often be assumed that \( h \leq h_0 \) where

\[
    h_0 = \min \left\{ \left( \frac{4}{|\beta|} \left( \frac{n^2 - |\beta|}{n^2 + |\beta|} \right) \right)^{\frac{1}{2}}, \left( \frac{1}{|\beta|} \right)^{\frac{1}{2}} \right\}, \quad -\pi^2 < \beta < 0, \tag{2.1}
\]

\[= \infty, \quad \beta \geq 0.\]

If \( Z = (Z_i) \) is an \( n \)-vector, then \( \|Z\| = \max_i |Z_i| \).

Throughout the remainder of this section, \( A = (a_{ij}) \) will denote an \( n \times n \) matrix. The following norms are used:

\[
    \|A\| = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|,
\]

\[
    \|A\|_s = \sum_{j=1}^{n} \max_{1 \leq i \leq n} |a_{ij}|,
\]

\[
    |A| = \max_{1 \leq i, j \leq n} |a_{ij}|.
\]

If \( a_{ij} \geq 0 \) we write \( A \geq 0 \) and say that \( A \) is non-negative. \( A \) is monotone if \( A^{-1} \) exists and \( A^{-1} \geq 0 \), and \( A \) is an M-matrix if \( A \) is monotone, \( a_{ii} > 0 \), and \( a_{ij} < 0 \) for \( i \neq j \). The following theorem (Ortega and Rheinboldt [20, p. 54]) will be useful:
Theorem 2.1

Assume that $\tilde{A}$ is an M-matrix, that $\tilde{\tilde{A}} \geq \tilde{A}$, and that $\tilde{\tilde{A}}$ satisfies the sign restrictions for M-matrices. Then $\tilde{A}$ is an M-matrix and $\tilde{A}^{-1} \geq \tilde{\tilde{A}}^{-1} \geq 0$.

If $\tilde{A}^{-1}$ exists we set $\tilde{A}^{-1} = (a_{ij}^*)$. We say that $\tilde{A}$ permits of an LU factorization if $\tilde{A} = LU$ where $L$ is a lower triangular matrix with unit diagonal and $U$ is an upper triangular matrix. We denote by $\ell(A)$ the strictly lower triangular matrix obtained by setting equal to zero the elements of $\tilde{A}$ above and on the diagonal.

We denote by $\Lambda$ the $n \times n$ matrix

$$
\Lambda = \begin{pmatrix}
1 & 1 & 1 & \cdots & 1 \\
0 & 1 & 1 & \cdots & 1 \\
& 0 & 1 & \cdots & 1 \\
& & & \ddots & \vdots \\
& & & & 0
\end{pmatrix},
$$

and set $A^\Lambda = A A A$: $A^\Lambda$ is obtained from $A$ by inverting the order of the rows and columns of $A$. The next lemma follows immediately from the definitions.

Lemma 2.2

(i) $\Lambda^2 = \Lambda$; (ii) $\|A^\Lambda\| = \|A\|$; (iii) If $A$ is non-singular, $(A^\Lambda)^{-1} = (A^{-1})^\Lambda$;

(iv) If $A$ is lower triangular (upper triangular) then $A^\Lambda$ is upper triangular (lower triangular); (v) $(A^\Lambda)^\Lambda = A$. 

We will need the following elementary lemmas:

**Lemma 2.3**

Assume that $0 \leq \psi \leq \frac{\pi}{2}$ and $t \geq 1$. Then $\sin(t\psi) \leq t \sin \psi$.

**Proof:** It clearly suffices to consider the case when $t \psi \leq \frac{\pi}{2}$. Set

$$g(t) = \sin (t \psi) - t \sin \psi.$$  

Then $g(1) = 0$ and

$$g'(t) = \psi \cos (t \psi) - \sin \psi,$$

$$\leq \psi \cos \psi - \sin \psi,$$

$$= \cos \psi (\psi - \tan \psi),$$

$$\leq 0.$$ 

The lemma follows.

**Lemma 2.4**

Let $0 \leq a \leq 1$. Then $\arccos (1 - a) \leq \frac{2a}{\sqrt{1 - \frac{a^2}{2}}}.$

**Proof:**

$$\arccos (1 - a) = \int_{1-a}^{1} \frac{dt}{\sqrt{\frac{1}{\left[1 - t^2\right]^2}}}.$$
Setting \( u = 1 - t \),

\[
\text{arc cos} \ (1 - a) = \int_{0}^{a} \frac{du}{\frac{1}{2} \left[ 2u - u^2 \right]^2},
\]

\[
= \int_{0}^{a} \frac{du}{\frac{1}{2} \left[ 2u \right]^2 \left[ 1 - \frac{u}{2} \right]^2},
\]

\[
\leq \frac{1}{\frac{1}{2} \left[ 1 - \frac{a}{2} \right]^2} \int_{0}^{a} \frac{du}{\frac{1}{2} \left[ 2u \right]^2},
\]

\[
= \left[ \frac{2a}{1 - \frac{a}{2}} \right]^{\frac{1}{2}}.
\]

Lemma 2.5

Assume that \( \psi \geq 0 \) and \( 0 < a \leq 1 \). Then

\[
\frac{\sinh \psi}{\cosh \psi \sinh a \psi} \leq \frac{1}{a}.
\]

Proof: Set

\[
f(\psi) = \frac{\sinh \psi}{\cosh \psi \sinh a \psi}.
\]
Then \( f(0+) = \frac{1}{a} \), and

\[
f'(\psi) [\cosh \psi \sinh a\psi]^2
= \cosh^2 \psi \sinh a\psi - \sinh^2 \psi \sinh a\psi - a \sinh \psi \cosh \psi \cosh a\psi,
= \sinh a\psi - a \sinh \psi \cosh \psi \cosh a\psi.
\]

But, \( \sinh a\psi \leq a \sinh \psi \) so that \( f'(\psi) \leq 0 \). The lemma follows.

**Lemma 2.6**

If \( A \) and \( B \) are \( n \times n \) matrices then \( \|A\| \leq n \|A\|; \|A\|_g \leq n \|A\|; \|AB\|_g \leq \|A\| \|B\|_g; \|A\| \leq \|A\|_g; |AB| \leq \|A\| \|B\|; \) and \( |AB| \leq |A| |B^T| \).

If \( H \) is an \( n \times n \) strictly lower triangular matrix then \( \|H\|_g \leq (n-1) |H| \).

**Proof:**

\[
\|AB\|_g = \sum_{j=1}^{n} \max_{1 \leq i \leq n} \left| \sum_{k=1}^{n} a_{ik} b_{kj} \right|
\leq \sum_{j=1}^{n} \max_{1 \leq i \leq n} \|A\| \max_{1 \leq k \leq n} |b_{kj}|,
= \|A\| \|B\|_g,
\]

as asserted. The other inequalities follow similarly.
3. LU FACTORIZATION OF JACOBI MATRICES

Let \( L \) be the tri-diagonal or Jacobi matrix,

\[
L = \begin{pmatrix}
    a_1 & -b_1 & 0 & 0 & \cdots & 0 \\
    -c_1 & a_2 & -b_2 & 0 & \cdots & 0 \\
    0 & -c_2 & a_3 & -b_3 & \cdots & 0 \\
    \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\
    0 & 0 & \cdots & \cdots & -b_{n-1} & 0 \\
    0 & 0 & \cdots & \cdots & -c_{n-1} & a_n
\end{pmatrix}
\]  \hspace{1cm} (3.1)

For \( 1 \leq k \leq n \) we denote by \( D_k \) the determinant of the \( k \times k \) submatrix of \( L \) formed by deleting the last \( n-k \) rows and columns; we set \( D_0 = 1 \) and \( D_{-1} = 0 \).

Lemma 3.1

Assume that \( D_k \neq 0, \ 1 \leq k \leq n \). Then \( L \) permits an LU factorization and

\[
L = (l_{ij}) \text{ and } U = (u_{ij}) \text{ are tri-diagonal matrices with coefficients}
\]

\[
l_{kk} = 1 \text{ and } u_{kk} = D_k/D_{k-1}, \ 1 \leq k \leq n; \]

\[
l_{k+1,k} = -c_k D_{k-1}/D_k \text{ and } u_{k,k+1} = -b_k, \ 1 \leq k \leq n-1;
\]

\[
l_{ij}^* = (D_{j-1}/D_{i-1}) \prod_{k=j}^{i-1} c_k, \ j \leq i;
\]

\[
u_{ij}^* = (D_{i-1}/D_{j}) \prod_{k=i}^{j-1} b_k, \ j \geq i;
\]

where

\[
\prod_{k=m+1}^{m} c_k = \prod_{k=m+1}^{m} b_k = 1
\]
Proof: Follows immediately from Gantmacher [8, p. 35].

Lemma 3.2

Let \( \mathbf{I} \) and \( \mathbf{\tilde{I}} \) be Jacobi matrices with non-negative diagonal elements and non-positive off-diagonal elements. Assume that \( \mathbf{I} \succeq \mathbf{\tilde{I}} \) and that \( \mathbf{\tilde{D}}_k > 0 \) for \( 1 \leq k \leq n \). Then: (i) \( \mathbf{I} \) and \( \mathbf{\tilde{I}} \) are M-matrices; (ii) \( \mathbf{I} \) and \( \mathbf{\tilde{I}} \) permit LU factorizations; (iii) \( \mathbf{U}, \mathbf{\tilde{U}}, \mathbf{L}, \) and \( \mathbf{\tilde{L}} \) are Jacobi M-matrices and \( \mathbf{U} \succeq \mathbf{\tilde{U}}, \mathbf{L} \succeq \mathbf{\tilde{L}} \); (iv) \( \mathbf{I}^{-1} \succeq \mathbf{\tilde{I}}^{-1} \succeq 0; \mathbf{\tilde{U}}^{-1} \succeq \mathbf{\tilde{U}}^{-1} \succeq 0; \mathbf{\tilde{L}}^{-1} \succeq \mathbf{\tilde{L}}^{-1} \succeq 0. 

Proof: We assert that

\[
\mathbf{D}_{k+1}/\mathbf{D}_k \succeq \mathbf{\tilde{D}}_{k+1}/\mathbf{\tilde{D}}_k > 0,
\]

for \( 0 \leq k \leq n-1 \). To see this we first note that \( \mathbf{D}_1/\mathbf{D}_0 = a_1/l \succeq \mathbf{\tilde{a}}_1/l = \mathbf{\tilde{D}}_1/\mathbf{\tilde{D}}_0 > 0 \), so that (3.2) holds for \( k = 0 \). Assume that (3.2) holds for \( 0 \leq k \leq m-1 \) where \( m \geq 1 \). Then, using the recurrence relations for \( \mathbf{D}_k \) and \( \mathbf{\tilde{D}}_k \) (Gantmacher and Krein [9, p. 77]), and remembering that \( a_k \succeq \mathbf{\tilde{a}}_k \succeq 0, 0 \leq b_k \succeq \mathbf{\tilde{b}}_k, \) and \( 0 \leq c_k \leq \mathbf{\tilde{c}}_k \),

\[
\mathbf{D}_{m+1}/\mathbf{D}_m = a_m - b_{m-1} c_{m-1} [\mathbf{D}_m/\mathbf{D}_{m-1}]^{-1},
\]

\[
\succeq \mathbf{\tilde{a}}_m - \mathbf{\tilde{b}}_{m-1} \mathbf{\tilde{c}}_{m-1} [\mathbf{\tilde{D}}_m/\mathbf{\tilde{D}}_{m-1}]^{-1},
\]

\[
\succeq \mathbf{\tilde{a}}_m - \mathbf{\tilde{b}}_{m-1} \mathbf{\tilde{c}}_{m-1} [\mathbf{\tilde{D}}_m/\mathbf{\tilde{D}}_{m-1}]^{-1},
\]

\[
= \mathbf{\tilde{D}}_{m+1}/\mathbf{\tilde{D}}_m,
\]

so that (3.2) holds for \( k = m \). Using induction, the assertion follows.
Since $D_k > 0$ and $\tilde{D}_k > 0$ for $1 \leq k \leq n$, $\mathcal{I}$ and $\mathcal{I}$ are "sign-regular" (Gantmacher and Krein [9, p. 94], and hence $\mathcal{I}$ and $\mathcal{I}$ are M-matrices.

Since $D_k > 0$ and $\tilde{D}_k > 0$ for $1 \leq k \leq n$, it follows from Lemma 3.1 that $\mathcal{I}$ and $\mathcal{I}$ permit LU factorizations. Using (3.2) and the explicit representations for $U$, $L$, etc. in Lemma 3.1 it is easily seen that $U \geq \tilde{U}$, and $L \geq \tilde{L}$.

Moreover, $U^{-1}$, $\tilde{U}^{-1}$, $L^{-1}$, $\tilde{L}^{-1} \geq 0$. Since the sign restrictions for M-matrices are satisfied, $U$, $\tilde{U}$, $L$, and $\tilde{L}$ are M-matrices. The remaining assertions of the lemma follow from Theorem 2.1.
4. PROPERTIES OF $I_\alpha$.

In this section we consider the special $n \times n$ Jacobi matrices $I_\alpha = (\gamma_{\alpha;ij})$,

$$
I_\alpha = \begin{pmatrix}
\alpha & -1 \\
-1 & \alpha & -1 \\
& -1 & \ddots & -1 \\
& & \ddots & \ddots & -1 \\
& & & -1 & \alpha \\
& & & & -1 & \alpha
\end{pmatrix}.
$$

The corresponding determinants and LU factors are denoted by $D_{\alpha,k}$, $L_\alpha = (l_{\alpha;ij})$, and $U_\alpha = (u_{\alpha;ij})$, respectively.

We shall use the functions

$$
\theta = \begin{cases}
\theta(\alpha) = \text{arc cosh}(\alpha/2), & \alpha > 2, \\
0, & \alpha = 2, \\
\text{arc cos} (\alpha/2), & 0 < \alpha < 2,
\end{cases}
$$

$$
\varphi = \varphi(\alpha) = (n+1)\theta(\alpha) = \frac{1}{h} \theta(\alpha),
$$

$$
d(\alpha, s) = \begin{cases}
\sinh[(s+1)\theta(\alpha)]/\sinh[\theta(\alpha)], & \alpha > 2, \\
s + 1, & \alpha = 2, \\
\sin[(s+1)\theta(\alpha)]/\sin[\theta(\alpha)], & 0 < \alpha < 2.
\end{cases}
$$

The following result is proved by Fischer and Usmani [7].
Lemma 4.1

Assume that $\alpha > 0$. Then: $D_{\alpha,k} = d(\alpha,k)$, for $-1 \leq k \leq n$;

$$
\gamma^*_{\alpha;ij} = \frac{d(\alpha,j-1)d(\alpha,n-i)}{d(\alpha,n)}, \quad 1 \leq j \leq i \leq n,
$$

$$
= \frac{d(\alpha,i-1)d(\alpha,n-j)}{d(\alpha,n)}, \quad 1 \leq i \leq j \leq n.
$$

Using Lemma 4.1 we obtain

Lemma 4.2

$$
|I_{\alpha}^{-1}| \leq \frac{\sinh(\phi/2)}{2 \sin \theta \cosh(\phi/2)} \leq \frac{(n+1)}{4}, \quad \alpha > 2,
$$

$$
\leq \frac{(n+1)}{4}, \quad \alpha = 2,
$$

$$
\leq \frac{\sin(\phi/2)}{2 \sin \theta \cos(\phi/2)} \leq \frac{(n+1)}{4 \cos(\phi/2)}, \quad 0 < \alpha < 2 \text{ and } 0 < \phi \leq \frac{\pi}{2},
$$

$$
\leq \frac{n+1}{2 \sqrt{2} \cos(\phi/2)}, \quad 0 < \alpha < 2 \text{ and } \frac{\pi}{2} \leq \phi < \pi.
$$

Proof: We begin by noting that if $\alpha \geq 2$ or $\alpha < 2$ and $\phi \leq \frac{\pi}{2}$ then $d(\alpha,s)$ is a monotone increasing function of $s$ so that

$$
|I_{\alpha}^{-1}| = \max_{1 \leq i \leq n} |\gamma^*_{\alpha;ii}|.
$$

We now consider the four cases separately.
Case 1: $\alpha > 2$. Then

$$\gamma^*_{\alpha; ii} = \frac{\sinh(\varphi-i\theta) \sinh i\theta}{\sinh \theta \sinh \varphi},$$

$$\leq \frac{\cosh \varphi - 1}{2 \sinh \theta \sinh \varphi},$$

$$= \frac{2 \sinh^2(\varphi/2)}{4 \sinh \theta \sinh (\varphi/2) \cosh(\varphi/2)},$$

$$= \frac{\sinh(\varphi/2)}{2 \sinh \theta \cosh (\varphi/2)}.$$

Appealing to Lemma 2.5 (or noting that $I_\alpha > I_2$ so that $I_\alpha^{-1} \leq I_2^{-1}$),

$$\gamma^*_{\alpha; ii} \leq \frac{1}{2} \frac{1}{\theta/(\varphi/2)},$$

$$\leq \frac{(n+1)}{4}.$$
Case 3: $\alpha < 2$ and $0 \leq \varphi \leq \frac{\pi}{2}$. Then

$$\gamma^*_{\alpha;ii} = \frac{\sin(\varphi-i\theta) \sin i \theta}{\sin \theta \sin \varphi},$$

$$= \frac{\cos(\varphi-2i\theta) - \cos \varphi}{2 \sin \theta \sin \varphi},$$

$$\leq \frac{1 - \cos \varphi}{2 \sin \theta \sin \varphi},$$

$$= \frac{2 \sin^2(\varphi/2)}{4 \sin \theta \sin(\varphi/2) \cos(\varphi/2)},$$

$$= \frac{\sin(\varphi/2)}{2 \sin \theta \cos(\varphi/2)}.$$

Appealing to Lemma 2.3,

$$\gamma^*_{\alpha;ii} \leq \frac{(\varphi/2) \sin \theta}{2 \sin \theta \cos(\varphi/2)},$$

$$= \frac{(n+1)}{4 \cos(\varphi/2)}.$$

Case 4: $0 < \alpha < 2$ and $\pi/2 \leq \varphi \leq \pi$. Then

$$\gamma^*_{\alpha;ij} = \frac{\sin p \theta \sin q \theta}{\sin \theta \sin \varphi},$$

where $p$ and $q$ are positive integers depending on $i$ and $j$ satisfying $p + q \leq n + 1$. Without loss of generality we may assume that $p \leq (n+1)/2$.

Using Lemma 2.3,
\[ \gamma^*_{\alpha;ij} \leq \frac{p \sin \theta \sin q \theta}{\sin \theta \sin \varphi}, \]
\[ \leq \frac{p}{\sin \varphi}, \]
\[ = \frac{p}{2 \sin(\varphi/2) \cos(\varphi/2)}, \]
\[ \leq \frac{p}{\sqrt{2} \cos(\varphi/2)}, \]
\[ \leq \frac{n+1}{2\sqrt{2} \cos(\varphi/2)}. \]

**Theorem 4.3**

Assume that \( \alpha \geq 2 + \frac{2}{h} \beta \) and that \( h \leq h_0 \). Then (i) \( I_\alpha \) is an M-matrix; (ii) \( D_{\alpha,k} > 0 \) for \( 1 \leq k \leq n \); (iii) \( |I_\alpha^{-1}| \leq K(\beta)/h \), where

\[ K(\beta) = \frac{1}{4}, \quad \beta \geq 0, \]
\[ = \left\{ 2 \sqrt{2} \cos \left( \frac{1}{2} \left( \frac{\pi^2 + |\beta|}{2} \right)^{\frac{1}{2}} \right) \right\}^{-1}, \quad 0 > \beta > -\pi^2. \]

**Proof:** It follows from (2.1) that \( \alpha \geq 1 \). Hence, using (4.4) and Lemma 4.1 we see that \( D_{\alpha,k} > 0 \) for \( 1 \leq k \leq n \). Applying Lemma 3.2 with \( J = J = I_\alpha \), it follows that \( I_\alpha \) is an M-matrix.

If \( \alpha \geq 2 \) then, from Lemma 4.2, \( |I_\alpha^{-1}| \leq 1/(4h) \). On the other hand, if \( \alpha < 2 \) then \( \beta < 0 \) so that from (2.1) and Lemma 2.4,
\[ \varphi = (n + 1) \theta , \]
\[ = (n+1) \arccos \left( \frac{a}{2} \right) , \]
\[ \leq (n + 1) \arccos \left( 1 + \frac{\beta h^2}{2} \right) , \]
\[ \leq \left( n + 1 \right) \left[ \frac{|\beta| h^2}{1 - |\beta| h^2 / 4} \right]^{\frac{1}{2}} , \]
\[ = \left[ \frac{|\beta|}{1 - |\beta| h^2 / 4} \right]^{\frac{1}{2}} , \]
\[ \leq \left[ \frac{\pi^2 + |\beta|}{2} \right]^{\frac{1}{2}} , \]
so that \( \varphi < \pi \). Hence, from Lemma 4.2,
\[ |I_\alpha^{-1}| \leq \frac{(n + 1)}{2\sqrt{2} \cos (\varphi/2)} \leq K(\beta)/h , \]
and the proof of the theorem is complete.
5. PERTURBATIONS OF MONOTONE MATRICES

Lemma 5.1

Assume that $\mathbf{H}$ is a strictly lower triangular matrix. Then $(\mathbf{I} - \mathbf{H})^{-1}$ and $(\mathbf{I} - \mathbf{H}^T)^{-1}$ exist and satisfy

$$|| (\mathbf{I} - \mathbf{H})^{-1} || \leq \exp[|| \mathbf{H} ||_s],$$

$$|| (\mathbf{I} - \mathbf{H}^T)^{-1} || \leq \exp[|| \mathbf{H}^T ||_s].$$

Proof: Let $\mathbf{H}^{(k)}$ denote the matrix obtained from $\mathbf{H}$ by setting equal to zero all the elements of $\mathbf{H}$ except those in the $k$-th column. Then

$$(\mathbf{I} + \mathbf{H}^{(n-1)}) \ldots (\mathbf{I} + \mathbf{H}^{(1)}) (\mathbf{I} - \mathbf{H}) = \mathbf{I}.$$ 

Hence,

$$|| (\mathbf{I} - \mathbf{H})^{-1} ||$$

$$= || (\mathbf{I} + \mathbf{H}^{(n-1)}) \ldots (\mathbf{I} + \mathbf{H}^{(1)}) ||,$$

$$\leq \prod_{k=1}^{n-1} || \mathbf{I} + \mathbf{H}^{(k)} ||,$$

$$\leq \prod_{k=1}^{n-1} [1 + || \mathbf{H}^{(k)} ||_s],$$

$$\leq \prod_{k=1}^{n-1} \exp[|| \mathbf{H}^{(k)} ||_s],$$

$$= \exp \left[ \sum_{k=1}^{n-1} || \mathbf{H}^{(k)} ||_s \right],$$

$$= \exp \left[ || \mathbf{H} ||_s \right],$$

as asserted, since $|| \mathbf{H} ||_s = \sum_{k=1}^{n-1} || \mathbf{H}^{(k)} ||_s$. 
Let $\tilde{H}^{(k)}$ denote the matrix obtained from $H^T$ by setting equal to zero all the elements of $H^T$ except those in the $k$-th column. Then

$$(I + \tilde{H}^{(1)}) \ldots (I + \tilde{H}^{(n-1)}) (I - H^T) = I.$$ 

Repeating the previous arguments, the second inequality follows.

**Theorem 5.2**

Assume that $A$, $P$, $L$, and $U$ are matrices such that:

(i) $P = (p_{ij})$ is non-negative and lower triangular;
(ii) $A = U L$;
(iii) $L = (l_{ij})$ is a lower triangular $M$-matrix;
(iv) $U$ is an upper triangular tri-diagonal $M$-matrix with unit diagonal. Then $(A + P)^{-1}$ exists and

$$\| (A + P)^{-1} \| \leq \| A^{-1} \| \exp \left[ \| P A^{-1} \| \kappa_s / (1 + \kappa) \right] / (1 + \kappa),$$

$$| (A + P)^{-1} | \leq | A^{-1} | \exp \left[ \| (PA^{-1})^T \| \kappa_s / (1 + \kappa) \right] / (1 + \kappa),$$

where

$$0 \leq \kappa \leq \min_i \left( l_{ii}^x p_{ii} \right).$$

**Proof:** Let $E = (e_{ij}) = P L^{-1}$. Then $A + P = (U + E) L$. Since $P$ and $L^{-1}$ are non-negative lower triangular matrices, the same is true of $E$. Moreover,

$$e_{ii} = p_{ii} l_{ii}^x \geq \kappa.$$

The off-diagonal elements of $U$ will be denoted by $-u_{ii}$. Thus, a typical $U + E$ is of the form
Now consider the process whereby $\mathbf{U} + \mathbf{E}$ is transformed into a lower triangular matrix, $\mathbf{I} + \mathbf{G}$ say, by using column operations to successively "kill off" the elements $-u_1$, $-u_2$, $\ldots$, $-u_{n-1}$ of $\mathbf{U} + \mathbf{E}$. It is easily seen that

$$\mathbf{I} + \mathbf{G} = (\mathbf{U} + \mathbf{E}) \mathbf{V},$$

$$\mathbf{V} = (\mathbf{I} + \mathbf{V}^{(1)}) \ldots (\mathbf{I} + \mathbf{V}^{(n-1)}),$$

$$0 \leq \mathbf{V}^{(k)} \leq \mathbf{Q}^{(k)}.$$ 

where $\mathbf{Q}^{(k)} = (q_{ij}^{(k)})$ is defined by

$$q_{ij}^{(k)} = u_k, \text{ if } i = k \text{ and } j = k + 1,$$

$$= 0, \text{ otherwise.}$$

Hence,

$$\mathbf{V} \leq (\mathbf{I} + \mathbf{Q}^{(1)}) \ldots (\mathbf{I} + \mathbf{Q}^{(n-1)}),$$

$$= \mathbf{U}^{-1}.$$ 

Therefore, remembering that $\mathbf{U}$ and $\mathbf{V}$ are upper triangular,

$$\mathbf{I} + \mathbf{G} = (\mathbf{U} + \mathbf{E}) \mathbf{V},$$

$$= \mathbf{I} + \mathbf{D} + \mathbf{L} (\mathbf{E} \mathbf{V}),$$

$$= \mathbf{I} + \mathbf{D} + \mathbf{L} (\mathbf{P} \mathbf{L}^{-1} \mathbf{V}),$$

$$= (\mathbf{I} + \mathbf{D}) (\mathbf{I} + \mathbf{H}), \text{ say,}$$
where

\[
D \geq \text{diag} \ (U + E) \geq (1 + \kappa)1,
\]

is a diagonal matrix, and

\[
H = (I + D)^{-1} \cdot (PL^{-1} U),
\]

\[
\leq \cdot (PL^{-1} U^{-1})/(1 + \kappa),
\]

\[
= \cdot (PA^{-1})/(1 + \kappa),
\]

is nonnegative and strictly lower triangular.

Now,

\[
(A + P) = (U + E) \cdot L,
\]

\[
= (I + G) \cdot V^{-1} \cdot L,
\]

\[
= (I + D)(I + H) \cdot V^{-1} \cdot L.
\]

Hence, \((A + P)^{-1}\) exists and

\[
(A + P)^{-1} = L^{-1} \cdot V \cdot (I + H)^{-1} \cdot (I + D)^{-1}.
\]

Using Lemma 5.1,

\[
\| (I + H)^{-1} \| \leq \exp \left[ \| H \|_s \right],
\]

\[
\leq \exp \left[ \| PA^{-1} \|_s / (1 + \kappa) \right],
\]

and

\[
\| (I + H^T)^{-1} \| \leq \exp \left[ \| H^T \|_s \right],
\]

\[
\leq \exp \left[ \| (PA^{-1})^T \|_s / (1 + \kappa) \right].
\]
Hence, using Lemma 2.6,

$$\| (A + P)^{-1} \|$$

\[ \leq \| L^{-1} U \| \| (I + H)^{-1} \| \| (I + D)^{-1} \|, \]

\[ \leq \| L^{-1} U^{-1} \| \| (I + H)^{-1} \| / (1 + \kappa), \]

\[ \leq \| A^{-1} \| \exp \left[ \| P A^{-1} \|_{s} / (1 + \kappa) \right] / (1 + \kappa), \]

and

$$\| (A + P)^{-1} \|$$

\[ \leq \| L^{-1} U \| \| (I + H)^{-1} (I + D)^{-1} \|_{T}, \]

\[ \leq \| L^{-1} U^{-1} \| \| (I + H)^{-1} \| \| (I + D)^{-1} \|, \]

\[ \leq \| A^{-1} \| \exp \left[ \| (P A^{-1})^{T} \|_{s} / (1 + \kappa) \right] / (1 + \kappa), \]

as asserted.
6. **THE NUMERICAL METHOD**

The interval \([0,1]\) is divided into \(n + 1\) subintervals each of length \(h\) the points of subdivision, or gridpoints, being denoted by \(t_{h,i} = ih, 0 \leq i \leq n + 1.\) The solution \(x\) of (1.1) is approximated at the \(n\) interior gridpoints.

The mappings \(\varphi_h : \mathbb{R}[0,1] \rightarrow \mathbb{R}^n\) and \(G_h : \mathbb{R}^n \rightarrow \mathbb{R}^n\) are defined by

\[
(\varphi_h y)_i = y(t_{h,i}), \quad 1 \leq i \leq n, \tag{6.1}
\]

\[
(G_h Y)_i = g(t_{h,i}, Y_i), \quad 1 \leq i \leq n, \tag{6.2}
\]

while \(I_\alpha\) is the \(n \times n\) matrix (4.1).

We assume that \(F_h : \mathbb{R}^n \rightarrow \mathbb{R}^n\) is an approximation to \(y\) such that

\[
F_h \varphi_h x - \varphi_h y x = \eta_h(x), \tag{6.3}
\]

where \(\|\eta_h(x)\| \rightarrow 0\) as \(h \rightarrow 0\). We observe that mappings \(F_h\) satisfying (6.3) are easily constructed for problems (1.3), (1.4), and (1.5).

Then the approximation to \(x\) at the \(n\) interior gridpoints is taken to be the solution \(Z_{h} \in \mathbb{R}^n\) of the system of \(n\) nonlinear equations

\[
\phi_h Z_h = I_2 Z_h + h^2 G_h Z_h + h^2 F_h Z_h = 0, \tag{6.4}
\]

provided that \(Z_h\) exists and is unique.

We set \(X_h = \varphi_h x\), and \(E_h = X_h - Z_h\). Since \(x\) is twice continuously differentiable,

\[
h^2 \varphi_h x'' = -I_2 X_h + h^2 \eta_h(x), \tag{6.5}
\]
where \( \| \varepsilon_h(x) \| \to 0 \) as \( h \to 0 \). Using (1.1), (6.3), and (6.5), we see that
\[
\phi_h \cdot \varepsilon_h = h^2 \varepsilon_h(x),
\] (6.6)
where \( \varepsilon_h(x) = \varepsilon_h(x) + \eta_h(x) \) so that \( \| \varepsilon_h(x) \| \to 0 \) as \( h \to 0 \).

In general, \( y \cdot x \) will have discontinuous derivatives at certain interior points of \([0,1]\). For example, for problem (1.4), \( y \cdot x \) has a discontinuous derivative at \( t = \frac{1}{2} \) (see (7.3)). To allow for this, we set
\[
\varepsilon_h(x) = \varepsilon_h^{(1)}(x) + \varepsilon_h^{(2)}(x).
\] (6.7)

We assume that \( \varepsilon_h^{(1)}(x) \) has at most \( m \) non-zero components, and also that
\[
\| \varepsilon_h^{(1)}(x) \| \leq K_1 h^{r-1}, \quad \| \varepsilon_h^{(2)}(x) \| \leq K_2 h^r,
\] (6.8)
where \( m, K_1, K_2, \) and \( r \) are constants. The idea is that \( \varepsilon_h^{(1)}(x) \) is the truncation error at the gridpoints where \( y \cdot x \) is non-smooth, while \( \varepsilon_h^{(2)}(x) \) is the truncation error at the remaining gridpoints.

Theorem 6.1

Assume that \( h \leq h_0 \), that \( F_h \) has a continuous Frechet derivative \( F_h' \), and that for all \( Y \in \mathbb{R}^n \), \( F_h'(Y) \) is a non-negative lower triangular matrix satisfying \( \| F_h'(Y) \| \leq M \), where \( M \) is a constant independent of \( h \) and \( Y \).

Then, \( \phi_h'(Y)^{-1} \) exists for all \( Y \in \mathbb{R}^n \) and
\[
\| \phi_h'(Y)^{-1} \| \leq K(\beta) \exp [M K(\beta)] h^{-2},
\]
where \( K(\beta) \) is defined in Theorem 4.3.
Also, there exists a unique solution $Z_h$ of (6.4) and the error $E_h$ satisfies

$$\|E_h\| \leq K(\beta) \exp \left[ M K(\beta) \right] \|e_h(x)\|.$$

If there is a constant $M_T$ such that $\|(F_h'(Y))^T\|_S \leq M_T$ for all $Y$, then

$$\|E_h\| \leq h^r K(\beta) \{mK_1 \exp[M_T K(\beta)] + K_2 \exp[M K(\beta)]\},$$

where $K_1$, $K_2$, $m$, and $r$, are as in (6.8).

Finally, if $g$ and $F_h$ are twice continuously differentiable, $Z_h$ may be computed by Newton's method,

$$Z^{(k+1)}_h = Z^{(k)}_h - \left[ \phi_h'(Z^{(k)}_h) \right]^{-1} \phi_h(Z^{(k)}_h), \quad (6.9)$$

provided that the initial approximation $Z^{(0)}_h$ is sufficiently good.

**Proof:** For any $Y \in \mathbb{R}^n$,

$$\phi_h'(Y) = \left[ I_2 + h^2 G_h'(Y) \right] + h^2 F_h'(Y),$$

$$= A + P, \text{ say.}$$

Set $\alpha = 2 + h^2 \beta$. Since $G_h'(Y) = \text{diag} \left( \frac{\partial g(t, h, Y_i)}{\partial Y_i} \right) \geq \beta I$, it follows that $A \geq I_\alpha$. Using Theorems 2.1 and 4.3 we see that $A^{-1}$ exists and that

$$|A^{-1}| \leq |I^{-1}_{2+h^2 \beta}| \leq K(\beta)/h. \text{ From Lemma 2.6 it follows that } ||A^{-1}|| \leq K(\beta)/h^2, \text{ and } ||A^{-1}||_S \leq K(\beta)/h^2.$$
From Theorem 4.3 we know that $D_{\alpha,k} > 0$ for $1 \leq k \leq n$. Therefore, using Lemma 3.2 with $\tilde{I} = I_\alpha$ and $\hat{I} = A$, we conclude that $A = L_1 U_1$ where $L_1$ and $U_1$ are tri-diagonal M-matrices which are, respectively, lower and upper triangular. Invoking Lemma 2.2,

$$A = \hat{A} = L_1^\Lambda U_1^\Lambda = (L_1 \; D)(D^{-1} U_1^\Lambda) = U L,$$

say, where $D = \text{diag} (U_1^\Lambda)$.

It is now easily seen that all the conditions of Theorem 5.2 are satisfied. Hence, using Lemma 2.6 and setting $\kappa = 0$, we see that $\Phi_h'(Y)^{-1}$ exists and satisfies

$$\| \Phi_h'(Y)^{-1} \|$$

$$\leq \| A^{-1} \| \exp \left[ \| P A^{-1} \|_S \right],$$

$$\leq \| A^{-1} \| \exp \left[ \| P \| \| A^{-1} \|_S \right],$$

$$\leq (K(\beta) h^{-2}) \exp \left[ (h^2 M)(K(\beta) h^{-2}) \right],$$

$$= K(\beta) \exp \left[ M K(\beta) / h^2 \right],$$

as asserted. Moreover, remembering that $A$ is symmetric,

$$| \Phi_h'(Y)^{-1} |$$

$$\leq | A^{-1} | \exp \left[ \| (P A^{-1})^T \|_S \right],$$

$$\leq | A^{-1} | \exp \left[ \| (A^{-1})^T \| \| P \|_S \right],$$

$$= K(\beta) \exp \left[ K(\beta) M_T \right]/h.$$
Since $\phi_h',(y)$ exists and is bounded for all $y$, it follows from the theorem of Hadamard (Ortega and Rheinboldt [20, p. 137]) that there exists a unique solution $Z_n$ of (6.4).

From (6.4) and (6.6)

$$\Phi_h X_n - \phi_h Z_n = h^2 e_h(x),$$

so that (Ortega and Rheinboldt [20, p. 71])

$$(\tilde{A} + \tilde{P}) E_n = h^2 e_h(x),$$

where

$$\tilde{A} = A_2 + h^2 \int_0^1 G_h(Z_n + t(X_n - Z_n)) dt,$$

$$\tilde{P} = h^2 \int_0^1 P_h(Z_n + t(X_n - Z_n)) dt.$$ 

Since $\tilde{A}$ and $\tilde{P}$ have the same properties as $A$ and $P$, respectively, the inverse $(\tilde{A} + \tilde{P})^{-1}$ exists and satisfies the same inequalities as $(A + P)^{-1}$.

Therefore,

$$\|E_n\| = \| (\tilde{A} + \tilde{P})^{-1} h^2 e_h(x) \|,$$

$$\leq K(\beta) \exp [M K(\beta)] \| e_h(x) \|. $$
Also,
\[
\| E_h \| \\
= \| (\tilde{A} + \tilde{P})^{-1} h^2 (\varepsilon_h^{(1)}(x) + \varepsilon_h^{(2)}(x)) \|,
\]
\[
\leq m h^2 \| (\tilde{A} + \tilde{P})^{-1} \| \| \varepsilon_h^{(1)}(x) \| + h^2 \| (\tilde{A} + \tilde{P})^{-1} \| \| \varepsilon_h^{(2)}(x) \|,
\]
\[
\leq K(\beta) h^r \{ m K_1 \exp [M_T K(\beta)] + K_2 \exp [M K(\beta)] \}.
\]

Finally, the assertion that Newton's method can be used if $g$ and $F_h$ are twice continuously differentiable, is a trivial consequence of the fact that $\phi_h'(Y)^{-1}$ exists and is bounded for all $Y$, and Kantorovitch's theorem on the convergence of Newton's method (Henrici [15, p. 367]).

**Theorem 6.2**

Assume that $F_h'(Y)$ is independent of $Y$, that
\[
h^2 \| I_2^{-1} \ F_h'(0) \| \leq p_1 < 1,
\]
that
\[
h^2 \| I_2^{-1} (G_h^h Y - G_h^h W) \| \leq p_2 \| Y - W \|.
\]
for all $Y, W \in R^n$, and that $p_3 = p_2/(1 - p_1) < 1$.

Then there exists a unique solution $Z_h$ of (6.4) which can be found by successive approximation,
\[ Z_h^{(k)} = \left( I_2 + h^2 F_h'(0) \right)^{-1} \left[ -h^2 F_h 0 - h^2 G_h Z_h^{(k-1)} \right], \]  

starting with any initial guess \( Z_h^{(1)} \).

The error \( E_h \) satisfies the inequalities

\[ \| E_h \| \leq \| e(x) \| / [4(1 - p_1)(1 - p_3)], \]

and

\[ \| E_h \| \leq h^r \left[ mK_1 + K_2 \right] / [4(1-p_1)(1-p_3)], \]

where \( m, r, K_1, \) and \( K_2 \), are as in (6.8).

**Proof:** From the assumptions it follows \( \left( I_2 + h^2 I_2^{-1} F_h'(0) \right)^{-1} \) exists and has norm less than \( 1/(1 - p_1) \). Since \( F_h'(Y) \) is independent of \( Y \), (6.4) may be rewritten in the equivalent form,

\[ I_2 Z_h + h^2 F_h'(0) Z_h = -h^2 F_h 0 - h^2 G_h Z_h. \]

The first part of the theorem is now an immediate consequence of the contraction mapping theorem (Ortega and Rheinboldt [20, p. 383]).

The estimates for \( E_h \) follow from the observation that

\[ \| E_h \| = \| X_h - Z_h \|, \]

\[ = \| (I_2 + h^2 F_h'(0))^{-1} h^2 \left[ G_h(Z_h) - G_h(X_h) + e_h(x) \right] \|, \]

\[ \leq p_2 \| E_h \| / (1 - p_1) + h^r I_2^{-1} e_h(x) / (1 - p_1), \]

and the fact that, from Lemma 4.2, \( I_2^{-1} \leq 1/4h \).
Remarks

1. There are many possible variations of Theorems 6.1 and 6.2. Many of the results quoted by Ortega and Rheinboldt [20] could be applied to (6.4). It is possible to obtain sharper bounds for \( \| I^{\alpha -1}_x \|, \| I^{\alpha -1}_s \|_s \), and \( | I^{\alpha -1}_x | \) (see Appendix A), and these can be used to sharpen the bounds in Theorem 6.1.

2. The following rather vague comments may be of some help in giving the reader a feel for Theorem 6.1. If \( \gamma \) is continuously differentiable and \( F_h \) is a "reasonable" approximation to \( \gamma \), then, noting (6.3), one sees that \( M \) will in general exist. The existence of \( M_T \) implies that the value of \( x(s) \) affects the values of \( x(t) \) for \( t > s \) in a "moderate" fashion. For example, \( M_T \) will not exist if the equation (1.1) is a delay differential equation with delay \( \Delta(t) = t - \frac{1}{2} \), since then \( t - \Delta(t) = \frac{1}{2} \) so that the value of \( x(\frac{1}{2}) \) will greatly affect the values of \( x(t) \) for \( t > \frac{1}{2} \).
7. A NUMERICAL EXAMPLE.

In this section we consider the numerical solution of (1.4), and compare our results with those of Nevers and Schmitt [18].

The existence of a unique solution \( y \) of (1.4) was established by Nevers and Schmitt [18].

Setting

\[
x(t) = -y(2t) - \frac{1}{2},
\]

we find that \( x \) satisfies (1.1) with

\[
g(t, x(t)) = -\frac{1}{4} \sin (x(t) + \frac{1}{2}) - (12t + 2),
\]

\[
(\gamma x)(t) = -(8t + 4)(1 - 2t), \quad t \leq \frac{1}{2},
\]

\[
= -(8t + 4)x \left( t - \frac{1}{2} \right), \quad \frac{1}{2} \leq t \leq 1.
\]

Taking \( n \) to be odd we define \( F_h \) as follows:

\[
(F_h \gamma)_{i} = -(8t_{h,i} + 4)(1 - 2t_{h,i}), \quad i \leq (n+1)/2,
\]

\[
= -(8t_{h,i} + 4)Y_{i-(n+1)/2}, \quad (n+1)/2 < i \leq n,
\]

so that \( \eta_h(x) = 0 \).

Since \( x \in \mathcal{C}[0,1] \), \( \gamma x \in \mathcal{C}[0,1] \) so that \( x \in \mathcal{C}^{(2)}[0,1] \). Noting (7.3) it follows that \( x \) is four times continuously differentiable on \([0, \frac{1}{2}]\) and \([\frac{1}{2}, 1]\), but that \( x^{(3)}(t) \) and \( x^{(4)}(t) \) may have jump discontinuities at the point \( t = \frac{1}{2} \). Let \( \tau_h(x) = \tau_h^{(1)}(x) + \tau_h^{(2)}(x) \), where

\[
\tau_h^{(1)}(x) = \tau_{h,i}(x), \quad i = (n+1)/2,
\]

\[
= 0, \text{ otherwise.}
\]
From (6.5) it follows that

\[ \| x^{(1)}(x) \| \leq \frac{h}{3} \sup_{0 \leq t \leq 1} |x^{(3)}(t)|, \]

\[ \| x^{(2)}(x) \| \leq \frac{h^2}{12} \sup_{0 \leq t \leq 1} |x^{(4)}(t)|, \]

so that (6.8) holds with \( m = 1, r = 2, K_1 = \frac{1}{3} \sup_{0 \leq t \leq 1} |x^{(3)}(t)|, \)

and \( K_2 = \frac{1}{12} \sup_{0 \leq t \leq 1} |x^{(4)}(t)|. \)

Since \( F'_h (y) \leq 0, \) we cannot apply Theorem 6.1.

In order to apply Theorem 6.2 we need the following lemma.

**Lemma 7.1**

If \( F'_h \) is as in (7.4) then, for all \( y \in \mathbb{R}^n, \)

\[ h^2 \| J_2^{-1} \| F'_h (y) \| \leq 25/36. \]

**Proof:** Set \( m = (n+3)/2. \) From (7.4),

\[ F'_h (y) = -4 \begin{pmatrix} O & O \\ D & O \end{pmatrix} \]

where \( D \) is the \((n+1-m) \times (n+1-m)\) diagonal matrix,

\[ D = \text{diag} (d_j) = \text{diag} (1 + 2jh), \quad m \leq j \leq n. \]
We denote by $S_i$ the $i$-th row-sum of $\left[ \mathbf{I}_2^{-1} \mathbf{F}_h(Y) \right]/(-4)$. Using Lemma 4.1,

$$S_i = \sum_{j=m}^n \gamma^*_2;ij \ d_j,$$

$$= \sum_{j=m}^n \frac{j(n+1-i)}{n+1} d_j + \sum_{j=m}^n \frac{i(n+1-j)d_j}{n+1}.$$

Clearly, $S_i \leq S_m$ if $i \leq m$, so we may restrict ourselves to the case $i \geq m$.

Then, if $ih = z$, and since $mh = \frac{1}{2} + \ h, \ nh = 1 - h$,

$$\sum_{j=1}^n j = n(n+1)/2,$$

$$\sum_{j=1}^n j^2 = n(n+1)(2n+1)/6,$$

$$S_i h^2 = (1-z) \sum_{j=m}^{i-1} jh^2(1+2jh) + z \sum_{j=1}^n h(1-jh)(1+2jh),$$

$$= (1-z) S_i^i + z S_i^i, \ \text{say}.$$
Now,

\[ S'_1 = h^2 \sum_{j=m}^{i-1} j + 2h^3 \left[ \sum_{j=1}^{i-1} j^2 - \sum_{j=1}^{m-1} j^2 \right], \]

\[ = h^2 \frac{(i-m)(m+i-1)}{2} + 2h^3 \left( \frac{i-1}{i}(2i-1)/6 - 2h^3 \frac{(m-1)(m)(2m-1)}{6} \right), \]

\[ = (z - \frac{1}{2} - h)(z + \frac{1}{2})/2 + (z-h)(z)(2z-h)/3 - \frac{1}{2}(\frac{1}{2} + h)(1 + h)/3, \]

\[ = \frac{2}{24} \left( (16z^3 + 12z^2 - 5) + h(-24z^2 - 12z - 12) + h^2 (8z - 4) \right). \]

\[ S''_1 = h \sum_{j=i}^{n} (1 + jh - 2h^2 j^2), \]

\[ = h \sum_{j=i}^{n} (1 + jh) - 2h^3 \left[ \sum_{j=1}^{n} j^2 - \sum_{j=1}^{i-1} j^2 \right], \]

\[ = h \frac{(n+1-i)(1+ih+1+nh)}{2} - 2h^3 \frac{n(n+1)(2n+1)}{6} + 2h^3 \frac{(i-1)(i)(2i-1)}{6}, \]

\[ = (1-z)(3z-h)/2 - (1-h)(2-h)/3 + (z-h)z(2z-h)/3, \]

\[ = (1-z)(3z-h)/2 + (z-1)z(2z-h)/3 + \frac{1-h}{3} [z(2z-h) - (2-h)], \]

\[ = (1-z)(3z-h)/2 + (z-1)z(2z-h)/3 + \frac{1-h}{3} [(z-1)(2z-h) + (2z-h) - (2-h)], \]

\[ = (1-z) \left( \frac{3z-h}{2} - \frac{2z^2-hz}{3} - \frac{(1-h)(2z+2-h)}{3} \right), \]

\[ = \frac{1-z}{6} \left[ 9 + 3z - 3h - 4z^2 + 2hz - 4z - 4 + 2h + 4hz + 4h - 2h^2 \right], \]

\[ = \frac{1-z}{6} \left[ (-4z^2 - z + 5) + h(6z + 3) - 2h^2 \right], \]
so that

\[
\frac{z}{1-z} S''_1 = \frac{1}{24} \left\{ -16z^3 - 4z^2 + 20z + h(24z^2 + 12z) - 8zh^2 \right\}.
\]

Hence,

\[
24h^2 S_1 / (1-z) = S'_1 + z S''_1 / (1-z),
\]

\[
= A_0(z) + h A_1(z) + h^2 A_2(z),
\]

where

\[
A_0(z) = 8z^2 + 20z - 5,
\]

\[
A_1(z) = -12,
\]

\[
A_2(z) = -4.
\]

Since \( A_1(z), A_2(z) < 0 \), it follows that

\[
h^2 S_1 \leq (1-z) A_0(z)/24,
\]

\[
= \varnothing(z), \text{ say}.
\]

Now,

\[
\varnothing(z) = (-8z^3 - 12z^2 + 25z - 5)/24,
\]

\[
\varnothing'(z) = -z^2 - z + \frac{25}{24}.
\]

Since \( \varnothing'(0.5) > 0, \varnothing'(1) < 0 \), and \( \varnothing'(0.64) < 0 \), \( \varnothing \) attains its maximum on \([\frac{1}{2}, 1]\) at a point \( z_0 < .64 \) satisfying \( \varnothing'(z_0) = -z_0^2 - z_0 + \frac{25}{24} = 0 \). Hence
\[ \varrho(z_0) = (1-z_0)(8z_0^2 + 20z_0 - 5)/24, \]
\[ = (1-z_0)(-8z_0 + \frac{25}{3} + 20z_0 - 5)/24, \]
\[ = (1-z_0)(18z_0 + 5)/36, \]
\[ = (-18z_0^2 + 13z_0 + 5)/36, \]
\[ = (18z_0 - \frac{75}{4} + 13z_0 + 5)/36, \]
\[ = (31z_0 - \frac{55}{4})/36, \]
\[ < ((31)(.64) - \frac{55}{4})/36 \]
\[ < \frac{25}{(4)(36)}. \]

Combining the above results, the lemma follows.

From Theorem 4.3 it follows that \( \|I_2^{-1}\| \leq 1/(4h^2) \). However, it is well known (Henrici [15, p. 371]) that \( \|I_2^{-1}\| \leq 1/(8h^2) \). Using this stronger inequality we see that for all \( \mathbf{v}, \mathbf{w} \in \mathbb{R}^n \),
\[ \|h^2 I_2^{-1} (G_h \mathbf{v} - G_h \mathbf{w})\| \]
\[ \leq \max | \frac{\partial g(t, \mathbf{v})}{\partial \mathbf{v}} | \cdot \|\mathbf{v} - \mathbf{w}\| / 8, \]
\[ \leq \|\mathbf{v} - \mathbf{w}\|/32. \]

From this and Lemma 7.1 we see that Theorem 6.2 applies with \( p_1 = 25/36 \), \( p_2 = 1/32 \), and \( p_3 = 9/88 \). In particular, the iteration (6.10) can be used and \( \|\mathbf{E}_h\| = O(h^2) \).
A program, NEVERS, was written to implement the algorithm (6.10) for problems of the form

\[ y(t) = g(t, y(t)) + c(t) y(t - \Delta(t)), \quad 0 < t < 1, \]

\[ y(t) = u(t), \quad t \leq 0, \]

\[ y(t) = v(t), \quad t \geq 1, \]

where \( \Delta(t) \) may be positive or negative but must be a multiple of \( h \). A listing of the program is given in Appendix B.

NEVERS was used to compute the solution \( x(t) \) of (1.1) with \( g \) and \( f \) defined by (7.2) and (7.3). The computations were performed using double-precision arithmetic (16 decimals) on the UNIVAC 1108 at the University of Wisconsin. The initial approximation, \( Z^{(0)} \), was taken to be zero. The iteration (6.10) was used until \( \|Z^{(k+1)} - Z^{(k)}\| \leq 1.10^{-10} \); this always occurred when \( k \) was less than or equal to 6. The approximation was computed for \( n = 4, 8, 16, 32, 64, \) and 128. Total computation time, including compilation, was 55 seconds.

Let \( Y = -Z - 0.5 \). Noting (7.1), we see that \( Y \) is an approximation to the solution \( y(s) \) of (1.4). The components of \( Y \) corresponding to \( s = 0.5, 1.0, \) and 1.5, are given in Table 7.1. For comparison, Table 7.1 also contains the values computed by Nevers and Schmitt [18].

The following observations concerning the numerical results may be made. Firstly, the results of Table 7.1 confirm the assertion of Theorem 6.2 that

\[ \| E_h \| = O(h^2). \]

Secondly, the rapid convergence of (6.10) is due to the fact that
p_3 = 1. Thirdly, as can be seen from Table 7.1, the approximations \( Y_h \) decrease monotonely as \( h \to 0 \); this is connected with the fact that \( G'_h(Y) \leq 0 \) and that \( \left[ I + h^{2} J_{2}^{-1} F'_h(0) \right]^{-1} \geq 0 \).
<table>
<thead>
<tr>
<th>n+1</th>
<th>s = .5</th>
<th>s = 1.0</th>
<th>s = 1.5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Y</td>
<td>ΔY</td>
<td>Y</td>
</tr>
<tr>
<td>4</td>
<td>-1.471368</td>
<td></td>
<td>-1.927188</td>
</tr>
<tr>
<td>8</td>
<td>-1.524873</td>
<td>-53505</td>
<td>-2.042571</td>
</tr>
<tr>
<td>16</td>
<td>-1.538884</td>
<td>-14011</td>
<td>-2.072713</td>
</tr>
<tr>
<td>32</td>
<td>-1.542430</td>
<td>-3546</td>
<td>-2.080336</td>
</tr>
<tr>
<td>64</td>
<td>-1.543319</td>
<td>-889</td>
<td>-2.082247</td>
</tr>
<tr>
<td>128</td>
<td>-1.543542</td>
<td>-223</td>
<td>-2.082725</td>
</tr>
<tr>
<td>Nevers &amp; Schmitt</td>
<td>-1.543053</td>
<td></td>
<td>-2.081821</td>
</tr>
</tbody>
</table>

Table 7.1
Numerical solution of (1.4).
APPENDIX A

Further properties of $I_\alpha$.

During the present investigation we obtained bounds for $\|I_\alpha^{-1}\|$, $\|I_\alpha^{-1}\|$, and $\|U_\alpha^{-1}\|$ which are sharper than those available in the literature. As it turned out, these bounds were not needed. We include them in the present appendix since they may be useful to other workers. We use the notation of section 4.

Fischer and Usmani [7] prove

**Theorem A.1**

$$\|I_\alpha^{-1}\| \leq \frac{\sinh \varphi - 2 \sinh [(\varphi - \theta)/2]}{(\alpha - 2) \sinh \varphi}, \quad \alpha > 2,$$

$$\leq (n + 1)^2 / 8, \quad \alpha = 2,$$

$$\leq \frac{n}{|\sin \theta \sin \varphi|}, \quad \alpha < 2.$$  

The following theorem gives a different bound for $\|I_\alpha^{-1}\|$ when $\alpha < 2$.

**Lemma A.2**

Let $0 < \alpha < 2$ and $\varphi < \pi$. Then

$$\|I_\alpha^{-1}\| < \frac{(n+1)^2}{8 \cos(\varphi/2)}.$$

**Proof:** Since $\varphi < \pi$, it follows from Lemma 4.1 that $I_\alpha^{-1} > 0$. Set

$$S_1 = \sum_{j=1}^{n} \gamma \ast_{\alpha;ij} = \sum_{j=1}^{n} \gamma \ast_{\alpha;ij}.$$
Then (Fischer and Usmani [7, p. 132]),

\[ S_1 = \frac{D_{\alpha, n} - (D_{\alpha, n-1} + D_{\alpha, i-1})}{(\alpha - 2) D_{\alpha, n}} \]

\[ = \frac{\sin \phi - \sin (\phi-i \theta) - \sin (i \theta)}{(\alpha - 2) \sin \phi} \]

\[ = \frac{2 \sin (\phi/2) [\cos (\phi/2) - \cos((\phi/2) - i \theta)]}{2(\alpha - 2) \sin (\phi/2) \cos (\phi/2)} \]

\[ = \frac{\sin [(\phi-i \theta)/2] \sin [(i \theta)/2]}{2 \sin^2 (\theta/2) \cos (\phi/2)} \]

Using Lemma 2.3,

\[ S_1 \leq \frac{(n + 1 - i)i}{2 \cos(\phi/2)} \]

\[ \leq \frac{(n + 1)^2}{8 \cos(\phi/2)} \]

as asserted.

**Lemma A.3**

\[ \| L_\alpha^{-1} \| = \frac{1}{2} + \frac{1}{2} \frac{\tanh(n \theta/2)}{\tanh(\theta/2)} \quad \alpha > 2; \]

\[ = (n+1)/2, \quad \alpha = 2; \]

\[ = \frac{1}{2} + \frac{1}{2} \frac{\tan(n \theta/2)}{\tan(\theta/2)} \quad 0 < \alpha < 2 \text{ and } \phi < n; \]

and
\[ ||U^{-1}_{\alpha}|| \leq \frac{\tanh \frac{\phi}{\theta}}{\theta}, \quad \alpha > 2; \]
\[ \leq \frac{1}{\sin \phi} + \frac{n+1}{\phi} \left( 1 + \left\lfloor \log \tan \left( \frac{\phi}{2} \right) \right\rfloor \right), \quad 0 < \alpha < 2 \text{ and } \phi < \pi; \]

where
\[ [u]_+ = u, \quad \text{if } u \geq 0, \]
\[ = 0, \quad \text{if } u < 0. \]

**Proof:** If \( A = (a_{ij}) \) and
\[ S_i = \sum_{j=1}^{n} |a_{ij}| \]
then
\[ ||A|| = \max_{1 \leq i \leq n} |S_i|. \]

Using Lemmas 3.1 and 4.1 we investigate the different possible cases.

**Case 1:** \( A = I_2^{-1} \). Then
\[ S_i = \sum_{j=1}^{i} \frac{D_2,j-1}{D_2,i-1}, \]
\[ = \sum_{j=1}^{i} \frac{i}{j}, \]
\[ = \frac{(i+1)}{2}, \]
so that

\[ \| L^{-1}_2 \| = (n+1)/2. \]

**Case 2:** \( A = L^{-1}_\alpha, \alpha > 2. \) Then

\[ S_1 = \sum_{j=1}^{i} D_{\alpha, j-1} / D_{\alpha, i-1}, \]

\[ = \frac{1}{\sinh i\theta} \sum_{j=1}^{i} \sinh j\theta, \]

\[ = \frac{1}{\sinh i\theta} \sum_{j=1}^{i} \frac{\cosh (j+\frac{1}{2})\theta - \cosh (j-\frac{1}{2})\theta}{2 \sinh(\theta/2)}, \]

\[ = \frac{\cosh (i+\frac{1}{2})\theta - \cosh \theta/2}{2 \sinh i\theta \sinh \theta/2}, \]

\[ = \frac{\cosh i\theta \cosh \theta/2 + \sinh i\theta \sinh \theta/2 - \cosh \theta/2}{2 \sinh i\theta \sinh \theta/2}, \]

\[ = \frac{1}{2} + \frac{(\cosh \theta/2)(\cosh i\theta - 1)}{2 \sinh i\theta \sinh \theta/2}, \]

\[ = \frac{1}{2} + \frac{(\cosh \theta/2)(2 \sinh^2 i\theta/2)}{2(\sinh \theta/2)(2 \sinh i\theta/2 \cosh i\theta/2)}, \]

\[ = \frac{1}{2} + \frac{1}{2} \frac{\tanh i\theta/2}{\tanh \theta/2}, \]

so that

\[ \| L^{-1}_\alpha \| = \frac{1}{2} + \frac{1}{2} \frac{\tanh(n\theta/2)}{\tanh (\theta/2)}. \]
Case 3: \( A = L^{-1}_\alpha \), \( \alpha < 2 \) and \( \varphi < \pi \). Then

\[
S_i = \sum_{j=1}^{i} \frac{D_{\alpha,j-1}}{D_{\alpha,i-1}},
\]

\[
= \frac{1}{\sin i\theta} \sum_{j=1}^{i} \sin j\theta,
\]

\[
= \frac{1}{\sin i\theta} \sum_{j=1}^{i} \frac{\cos (j - \frac{1}{2})\theta - \cos(j + \frac{1}{2})\theta}{2 \sin(\theta/2)}.
\]

\[
= \frac{\cos(\theta/2) - \cos(i + \frac{1}{2})\theta}{2 \sin i\theta \sin \theta/2},
\]

\[
= \frac{\cos(\theta/2) - \cos i\theta \cos \theta/2 + \sin i\theta \sin \theta/2}{2 \sin i\theta \sin \theta/2},
\]

\[
= \frac{1}{2} + \frac{\cos \theta/2 (1 - \cos i\theta)}{2 \sin \theta/2 \sin i\theta},
\]

\[
= \frac{1}{2} + \frac{(\cos \theta/2) \sin^2(i\theta/2)}{2 \sin(\theta/2) (2 \sin i\theta/2 \cos i\theta/2)},
\]

\[
= \frac{1}{2} + \frac{1}{2} \tan(i\theta/2),
\]

so that

\[
\|L^{-1}_\alpha\| = \frac{1}{2} + \frac{1}{2} \tan(n\theta/2).
\]
Case 4: \( A = U_2^{-1} \). Then

\[
S_i = \sum_{j=i}^{n} \frac{D_{2,i-1}/D_{2,j'}}{j+1},
\]

\[
= \sum_{j=i}^{n} \frac{i}{j+1},
\]

\[
\leq i \int_{i}^{n+1} \frac{1}{x} \, dx,
\]

\[
= i \log \left( \frac{n+1}{i} \right),
\]

\[
= f(i), \text{ say}.
\]

Since

\[
f'(x) = \log \left( \frac{n+1}{x} \right) - 1,
\]

\( f(x) \) attains its maximum in the interval \((0,n+1]\) at the point \( x = (n+1)/e \).

Hence

\[
\| U_2^{-1} \| \leq f \left( \frac{n+1}{e} \right) = (n+1)/e.
\]
Case 5: $A = U^{-1}_\alpha$, $\alpha > 2$. Then

$$S_i = \sum_{j=1}^{n} \frac{D_{\alpha,i-1}}{D_{\alpha,j}},$$

$$= \sum_{j=1}^{n} \frac{\sinh \theta}{\sinh(j+1)\theta},$$

$$\leq \sinh(i\theta) \int_{i}^{n+1} \text{csch}(t\theta) \, dt,$$

$$= \frac{\sinh(i\theta)}{\theta} \int_{i\theta}^{\Phi} \text{csch} \, dz,$$

$$= \frac{\sinh i\theta}{\theta} \log \left[ \frac{\tanh(\Phi/2)}{\tanh(i\theta/2)} \right],$$

$$= \frac{2}{\theta} f(i \theta/2),$$

where

$$f(z) = \frac{1}{2} \sinh 2z \log \left[ \frac{\tanh(\phi/2)}{\tanh z} \right].$$

Now: (i) $f(\phi/2) = 0$; (ii) $f(0) = 0$;

(iii) $f(z) > 0$ if $0 < z < \phi/2$. Hence, $f$ attains its maximum in the interval $[0, \phi/2]$ at an interior point $z = \zeta$, where $f'(\zeta) = 0$. Furthermore,

$$f'(z) = (\cosh 2z) \log \left[ \frac{\tanh(\phi/2)}{\tanh z} \right] - \frac{1}{2} \sinh 2z \frac{\text{sech}^2 z}{\tanh z},$$

$$= (\cosh 2z) \log \left[ \frac{\tanh(\phi/2)}{\tanh z} \right] - 1.$$
Hence, if \( f'(\zeta) = 0 \),

\[
\log \left( \frac{\tanh(\varphi/2)}{\tanh \zeta} \right) = \frac{1}{\cosh 2\zeta},
\]

and

\[
f(\zeta) = \frac{1}{2} \tanh 2\zeta.
\]

Therefore,

\[
\max_{0 \leq z < (\varphi/2)} f(z) \leq \frac{1}{2} \tanh \varphi.
\]

Consequently,

\[
\| U^{-1}_\alpha \| \leq [\tanh \varphi]/\vartheta.
\]

The above bound for \( \max f(z) \) is rather crude, but we have been unable to obtain a simple sharp bound.
Case 6: $A = \frac{U^{-1}}{\alpha}$, $0 < \alpha < 2$, $\varphi < \pi$. Then

$$S_i = \sum_{j=1}^{n} \frac{D_{\alpha,i-1}}{D_{\alpha,j}},$$

$$= \sum_{j=1}^{n} \frac{\sin i\theta}{\sin (j+1)\theta},$$

$$= \sin i\theta \left\{ \frac{1}{\sin(n+1)\theta} + \sum_{k=i+1}^{n} \frac{1}{\sin k\theta} + \sum_{k=i+1}^{n} \frac{1}{\sin k\theta} \right\},$$

$$\leq \sin i\theta \left\{ \frac{1}{\sin \varphi} + \left[ \frac{1}{\theta} \int_{\pi/2}^{\varphi} \frac{dt}{\sin t} \right]_+ + \left[ \frac{1}{\theta} \int_{i\theta}^{\pi/2} \frac{dt}{\sin t} \right]_+ \right\},$$

where

$$[u]_+ = \begin{cases} u, & \text{if } u \geq 0, \\ 0, & \text{if } u < 0, \end{cases}$$

and where we have used the fact that $\sin t$ is monotone increasing for $0 \leq t \leq \frac{\pi}{2}$ and monotone decreasing for $\frac{\pi}{2} \leq t \leq \pi$.

Since

$$\int \frac{dt}{\sin t} = \log \tan (t/2),$$

$$S_i \leq \frac{1}{\sin \varphi} + \left( n+1 \right) \left[ \frac{\log \tan \left( \varphi/2 \right)}{\varphi} \right]_+ + \frac{(n+1)}{\varphi} \left[ f(i\theta) \right]_+.$$
where
\[
\begin{align*}
  f(z) &= -\sin z \left[ \log \tan \left(\frac{z}{2}\right) \right], \\
        &= -\frac{2 \tan \left(\frac{z}{2}\right) \log \tan \left(\frac{z}{2}\right)}{1 + (\tan \left(\frac{z}{2}\right))^2}, \\
        &= \frac{2 \cot \left(\frac{z}{2}\right) \log \cot \left(\frac{z}{2}\right)}{1 + [\cot \left(\frac{z}{2}\right)]^2}, \\
        &= 2g[\cot \left(\frac{z}{2}\right)],
\end{align*}
\]
with
\[
g(u) = \frac{u \log u}{1 + u^2}.
\]

Consider \( g(u) \) on the interval \([1, \infty)\). We have that
\[
g'(u) = \frac{\frac{u^2}{(u^2+1)^2} \left[ 1 + \frac{2}{u^2} - \log u \right]}{u^2 - 1},
\]
from which it is easily seen that \( g' \) has only one zero, \( u = \delta \) say, and that \( e < \delta < e^{3/2} \). Since \( g(1) = g(\infty) = 0 \), \( g \) attains its maximum at \( u = \delta \).

Hence
\[
\max_{1 \leq i \leq n} [f(i\theta)]_+ \leq \max_{0 \leq z \leq \frac{\pi}{2}} f(z),
\]
\[
= 2 \max_{1 \leq u < \infty} g(u),
\]
\[
= 2 \max_{e \leq u \leq e^{3/2}} g(u),
\]
\[
\leq 2(\log e^{3/2}) \max_{e \leq u \leq \infty} \frac{u}{1 + u^2},
\]
\[
= \frac{3e}{1 + e^2},
\]
\[
< 1.
\]
Combining the above results,

\[ \| U_{\alpha}^{-1} \| \leq \frac{1}{\sin \varphi} + \frac{n+1}{\varphi} \left( 1 + \left[ \log \tan(\varphi/2) \right]_+ \right). \]
APPENDIX B

The program NEVERS.

A listing of NEVERS is given at the end of this appendix. Here we make a few comments to make the program more easily comprehensible to the reader.

NEVERS computes approximate solutions to problems of the form

\[
\begin{align*}
\ y''(t) &= g(t, y(t)) + c(t) y(t-\Delta(t)), \quad 0 < t < 1, \\
\ y(t) &= u(t), \quad t \leq 0, \\
\ y(t) &= v(t), \quad t \geq 1,
\end{align*}
\] (A.1)

where \( \Delta(t) \) may be positive or negative but must be a multiple of the stepsize \( h \).

Setting

\[ x(t) = y(t), \quad 0 \leq t \leq 1, \] (A.2)

(A.1) takes the form (1.1) with

\[
(\tilde{x})_x(t) = \begin{cases} 
    c(t) u(t - \Delta(t)), & \text{if } t - \Delta(t) \leq 0, \\
    c(t) x(t), & \text{if } 0 < t - \Delta(t) < 1, \\
    c(t) v(t - \Delta(t)), & \text{if } 1 \leq t - \Delta(t).
\end{cases}
\] (A.3)

The approximate solution \( Z_h \) is computed using the iteration (6.10). The correspondence between the variables of (6.10) and the arrays used in NEVERS is as follows:

\begin{align*}
\text{Equation (6.10)} & \quad \text{NEVERS} \\
I_2 + h^2 F'_h(0) & \quad A \\
-h^2 F_h(0) & \quad BC \\
Z_h^{(k+1)} & \quad X \\
Z_h^{(k)} & \quad XP
\end{align*}
The initial guess $z_h^{(1)}$ is taken to be $\varphi_h x_0$ where $x_0 \in \mathcal{S}[0,1]$ is provided by the user.

The functions $u(t)$, $v(t)$, $g(t,x)$, $c(t)$, $\Delta(t)$, and $x_0(t)$ must be provided by the user as procedures at the beginning of NEVERS. The iteration (6.10) is continued until $k = KMAX$ or $\|z_h^{(k)} - z_h^{(k-1)}\| \leq \epsilon$, and the approximations $z_h$ are computed for $2 \leq n \leq 2**IMAX$; in the listed program $KMAX = 20$, $\epsilon = 10^{-10}$, and $IMAX = 7$.

NEVERS makes use of the BUMP2 matrix package on the Univac 1108 at the University of Wisconsin. In addition to using the housekeeping subroutines MTADFM, MTMDEF, and MTMDFM, NEVERS uses the following BUMP2 subroutines:

- MTCNST - Set up a constant matrix
- MTMPRT - Print a matrix
- MTINVD - Invert a double-precision matrix
- MTMPY - Multiply two matrices
- MTSUB - Subtract two matrices
- MTNRMR - Compute the maximum row-sum norm of a matrix
- MTMOVE - Move a matrix

The listed program uses double-precision arithmetic but is written so that it can be converted to single-precision by changing only three of the program's statements.
PROGRAM NEVERS

******************************************************************************

PROGRAM TO SOLVE BOUNDARY VALUE PROBLEMS FOR EQUATIONS WITH
LINEAR DEPENDENT COEFFICIENT ON INTERVAL [0, 1]

Y(T) = u(T) FOR T \leq 0  Y(T) = v(T) FOR T \geq 1

DY(T)/DT + G(T, Y(T)) = C(T) FOR 0 < T < 1

SOLUTION IS COMPUTED BY ITERATION INITIAL GUESS IS X0(T)

PROGRAM USES BUMPS MATRIX PACKAGE ON IUSS 1108

SAMPLE PROBLEM IS PROBLEM OF NEVERS AND SCHMITT

******************************************************************************

IMPLICIT DOUBLE PRECISION (A-H, O-7)
DIMENSION A(129), B(129), X(129), R(129), XDF(129)
INTEGER FORM(2), XDIF(129)

II(T) = -2 * T
V(T) = 0 * T
G(T, X) = -sin(x + 5)/4 - (12 * t + 2)
C(T) = -(8 * T)*4)
DELTA(T) = 0.5 + n * T

NA = 129
IMAX = 7

TO RUN IN SINGLE PRECISION CHANGE TYP, FORM(1), IMPLICIT DOUBLE
PRECISION, AND MGINVD (STATEMENT NUMBER 670)

******************************************************************************

NA = 129
IMAX = 7

PROBLEM RUN WITH STPSIZE H = 1/4 *** 1/2**IMAX
NA = 5**2 FOR SUMS OF ARRAYS IN 1**IMAX.

******************************************************************************

KMAX = 20
EPS = 1.E-10

****** ITERATIONS CONTINUED UNTIL DIFFERENCE BETWEEN SUCCESSIVE
APPROXIMATIONS LESS THAN EPS
******

****** NUMBER OF ITERATIONS EQUAL TO KMAX
******

**PRINT 50, NA, IMAX, KMAX
**FORMAT(11, NA, IMAX, KMAX, EPS = 1, 9.5)
**WRITE(6, FORM) EPS

**CALL MTADFM(1, A, NA, NA, TYP)
**CALL MTADFM(5, B, R, X, XDF, X, NA, 1, TYP)

**START OF OUTER LOOP
**
**DO 9000 I = 2, 1MAX
**N = (7**I) - 1
**NEN = 1.
**H = HEN/(N+1)
**H2 = H**2
PRINT 100,1,N
FORMAT (1*1,N*N,H=1,8F15)
WRITE(A,FORMH)
CALL MTMDEF(A,N,N,(GFX1))
CALL MTMDEFM(A,RC,R,YD,N,1,(GFX1))
C******************************************************************************
C SET UP MATRIX A AND CONSTANT RIGHT HAND SIDE RC
C******************************************************************************
?FR=0.
CALL MTCNST(A,?FR,TD)
CALL MTCNST(RC,?FR,TD)
DO 500 J=1,N
T=J*H
A(J,J)=?.

JP1=J+1
IF ( (JP1) .LE. N) A(J,JP1)=-1.0
IF ( (JP1) .GT. N) BC(J)=RC(J)+V(T+H)
C
JM1=J-1
IF ( (JM1) .GE. 1) A(J,JM1)=1.0
IF ( (JM1) .LT. 1) BC(J)=RC(J)+H(V(T-H)
C
JD=J-INT( (DLTA(T)+.5*H)/H)
TD=T-DLTA(T)
CD=C(T)*H
IF ( (JD) .LT. 1) AND ( (JD) .LE. N) ) A(J,JD)=A(J,JD)+CD
IF ( (JD) .LT. 1) RC(J)=RC(J)-CD*V(TD)
IF ( (JD) .GT. N) RC(J)=RC(J)-CD*V(TD)
500 CONTINUE
C******************************************************************************
C IF N.LE.4 PRINT MATRIX A FOR DEBUGGING PURPOSES
C******************************************************************************
C IF ( N.LE.4 ) CALL MTMPRINT(A,FORMH,0,1 MATRIX A = 1 )
C******************************************************************************
C INVERT A AND SET UP INITIAL GUESS XP
C******************************************************************************
600 CALL MTINVD( A,A,?S700)
GO TO 800
700 STOP 700
800 CONTINUE
DO 830 J=1,N
T=J*H
830 XP(J)=X0(T)
C******************************************************************************
C IFREATF
C******************************************************************************
K=0
850 K=K+1
DO 900 J=1,N
T=J*H
900 A(J)=RC(J)-G(T,XP(J)*H2
CALL MTMDFM(A,RC,R,YD,N,1)
C******************************************************************************
C IF N.LE.4 AND K=1 PRINT VARIOUS MATRICES FOR DEBUGGING
C

*** INDEXED LOOP ***
IF ( M.GT.4 ) ..OR.. ( K.GT.1 ) GO TO 925
CALL MTRMT(A,FORM,0,1 INVERSE OF MATRIX A ...)
CALL MTRMT(RC,FORM,0,1 VECTOR RC ...
CALL MTRMT(XP,FORM,0,1 INITIAL VECTOR XP ...
CALL MTRMT(R,FORM,0,1 VECTOR R ...
CALL MTRMT(X,FORM,0,1 VECTOR X ...
225 CONTINUE
CALL MTSUB(X,XP,XDIF)
CALL MTRMRA(XDIF,RNXDIF)
PRINT 950,K
950 FORMAT ( 'ITERATION NUMBER, ROW NORM OF XDIF = ', F5)
WRITE(6, FORM) RNXDIF
IF ( RNXDIF .LT. EPS ) GO TO 9900
IF ( K.GE.KMAX ) GO TO 8800
CALL MTRMT(X,XP)
GO TO 850
8800 PRINT 8801
8801 FORMAT ( 'ITERATION MAXIMUM REACHED' )
GO TO 8950
8900 PRINT 8901
8901 FORMAT ( 'ITERATIONS CONVERGED' )
GO TO 8950
8950 CALL MTRMT(X,FORM,0,1 VECTOR X ...
9000 CONTINUE
STOP
END
REFERENCES


