Recursive Collocation for the Numerical Solution of Stiff Ordinary Differential Equations*

By H. Brunner

Abstract. The exact solution of a given stiff system of nonlinear (homogeneous) ordinary differential equations on a given interval I is approximated, on each subinterval σ_k corresponding to a partition τ_I of I, by a linear combination U_k(x) of exponential functions. The function U_k(x) will involve only the "significant" eigenvalues (in a sense to be made precise) of the approximate Jacobian for σ_k. The unknown vectors in U_k(x) are computed recursively by requiring that U_k(x) satisfy the given system at certain suitable points in σ_k (collocation), with the additional condition that the collection of these functions \{U_k\} represent a continuous function on I satisfying the given initial conditions.

1. Introduction. Consider a system of nonlinear ordinary differential equations,

\begin{align}
Y'(x) &= F(x, Y(x)), \quad x \in I = [0, a] \quad (Y \in \mathbb{R}^n, n > 1), \\
Y(0) &= Y_0.
\end{align}

with the prescribed initial condition

(1.2) \quad F(x, 0) = 0 \quad \text{for all } x \in I

(we note that this definition differs somewhat from the definition of homogeneity used by Hahn [3, p. 278]).

Let G(x, Y) = \frac{\partial F}{\partial Y}_{(x,Y)} denote the Jacobian of (1.1a), with G(x, Y) real in

T = \{(x, Y): x \in I, ||Y|| < \infty \}. We assume that the elements \{g_{ij}\} of G(x, Y) satisfy

\begin{align}
g_{ij} &\in C(T), \quad i, j = 1, \ldots, n.
\end{align}

For simplicity in notation, we restrict our discussion to the case where the eigenvalues \{\lambda_\nu(x): \nu = 1, \ldots, n\} of G(x, Y(x)) (with Y(x) being the exact solution of the initial-value problem (1.1)) are simple for x \in I. Furthermore, we shall adopt the following notation:

(a) the real eigenvalues will be ordered such that

\begin{align}
\lambda_1(x) < \lambda_{i-1}(x) < \cdots < \lambda_i(x) < 0,
\end{align}

and

Received June 22, 1973.

AMS (MOS) subject classifications (1970). Primary 65L05.

Key words and phrases. System of stiff nonlinear homogeneous ordinary differential equations, approximate solution by sums of exponential functions, recursive collocation, one-step methods.

* This research was supported by the National Research Council of Canada (Grant No. A-4805).

Copyright © 1974, American Mathematical Society

475
(b) the complex eigenvalues \( \lambda_\nu(x) = u_\nu(x) + iv_\nu(x) \), with \( v_\nu(x) > 0 \), will be arranged such that

\[
  u_r(x) \leq u_{r-1}(x) \leq \cdots \leq u_{i+1}(x) < 0,
\]

where \( r = (n + l)/2 \).

Since \( G(x, Y(x)) \) is assumed to be real-valued for \( x \in I \), \( \lambda_\nu(x) \) will also be an eigenvalue of \( G(x, Y(x)) \).

Although the present paper will emphasize the numerical solution of stiff systems (1.1a), we note that the above assumption that \( \Re(\lambda_\nu(x)) < 0 \), \( x \in I \), \( \nu = 1, \cdots, r \), is not essential for the following discussion. In fact, many problems arising in chemistry, for example, yield \( \lambda_\nu(x) = 0 \), \( x \in I \), for at least one value of \( \nu \). Compare also the numerical example presented in Section 3.

We recall that the system (1.1a) is called stiff in the interval \( I \) if the eigenvalues \( \{ \lambda_\nu(x) : \nu = 1, \cdots, n \} \) satisfy, for all \( x \in I \), the two conditions

(i) \[ \Re(\lambda_\nu(x)) < 0, \quad \nu = 1, \cdots, n, \]

and

(ii) \[ \max_{\nu} [\Re(-\lambda_\nu(x))] \gg \min_{\nu} [\Re(-\lambda_\nu(x))] \]

(see, for example, Lambert [4, p. 232]).

It is the purpose of this paper to introduce an adaptive method for the numerical solution of a stiff system of the form (1.1) (subject to the assumptions stated above). This method is based on the following idea: The exact solution \( Y(x) \) of (1.1) shall be approximated on \( I \) by a continuous function \( U(x) \in \mathbb{R}^n \) which is chosen such as to reflect the "structure" (i.e., the stiffness) of the given system. In other words, on a given subinterval \( \sigma_k \) of \( I \) (defined by a partition \( \pi_N \) of \( I \), \( U(x) \) shall have a representation \( U_k(x) \) which will be a linear combination of exponential terms involving only the "significant" eigenvalues (to be made precise in the following section) of the approximate Jacobian \( G(x, U(x)) \) for that interval. For a given partition \( \pi_N \) of \( I \), the \( N \) representations \( \{ U_k(x) : k = 0, 1, \cdots, N - 1 \} \) of \( U(x) \) will be generated recursively, via the requirements of continuity at the points of \( \pi_N \) and of collocation at suitable points of \( \sigma_k \), \( k = 0, 1, \cdots, N - 1 \). We note that this method of recursive collocation may be regarded as an implicit one-step method to generate approximate values to \( Y(x) \) at the points of the partition \( \pi_N \) of \( I \).

2. Computation of \( U(x) \). Let \( N \geq 1 \), and let \( \pi_N \) denote a partition of \( I \):

\[ \pi_N : 0 = \xi_0 < \xi_1 < \cdots < \xi_N = a. \]

Set \( \sigma_k = \{ x : \xi_k \leq x \leq \xi_{k+1} \} \), \( k = 0, 1, \cdots, N - 1 \). We denote by \( U(x) \) an approximate solution of (1.1), with \( U(x) = U_k(x), x \in \sigma_k \). Furthermore, let \( \{ \lambda_{k,\nu} : \nu = 1, \cdots, n \} \) be the eigenvalues of \( G(\xi_k, U_k(\xi_k)) \) and, as above, suppose that:

(i) for \( k = 0, 1, \cdots, N - 1 \), the eigenvalues \( \{ \lambda_{k,\nu} \} \) are simple;

(ii) for a given \( k \), the real eigenvalues are ordered such that \( \lambda_{k,1} < \lambda_{k,1-1} < \cdots < \lambda_{k,1} < 0 \); and

(iii) for a given \( k \), the complex eigenvalues \( \lambda_{k,\nu} = u_{k,\nu} + iv_{k,\nu} \), with \( v_{k,\nu} > 0 \), satisfy
where \( r = (n + l)/2 \).

For \( x \in \sigma_k \), \( U(x) \) shall possess the representation

\[
U_k(x) = \sum_{r=1}^{\omega(k)} A_{k,r} \varepsilon_{k,r}(x) + \sum_{\mu=1}^{r(k)} (B_{k,\mu} c_{k,\mu}(x) + C_{k,\mu} s_{k,\mu}(x)) \cdot \varepsilon_{k,\mu}(x),
\]

with

\[
\varepsilon_{k,r}(x) = \exp[\text{Re}(\lambda_{k,r}) \cdot (x - \xi_k)],
\]

\[
c_{k,\mu}(x) = \cos[\nu_{k,\mu} \cdot (x - \xi_k)],
\]

\[
s_{k,\mu}(x) = \sin[\nu_{k,\mu} \cdot (x - \xi_k)].
\]

Here, \( \omega(k) \) and \( r(k) \) are such that only the "significant" eigenvalues \( \{\lambda_{k,r}\} \) are taken into account: for a given \( M_k \),

\[
M_k \geq \min_{r=1, \ldots, r} [\text{Re}(-\lambda_{k,r})] > 0,
\]

\( \omega(k) \) and \( r(k) \) are given by

\[
\text{Re}(-\lambda_{k,r}) > M_k \quad \text{for } r = \omega(k) + 1, \ldots, l; \lambda_{k,r} \text{ real},
\]

\[
\text{Re}(-\lambda_{k,r}) > M_k \quad \text{for } r = \tau(k) + 1, \ldots, r; \lambda_{k,r} \text{ complex}.
\]

Clearly, \( \max(\omega(k), r(k) - l) \geq 1 \). (If the upper limit in one of the two sums in (2.1) is smaller than the lower one, then this sum is assigned the value zero.)

We note here that, for large systems of ordinary differential equations, an alternate definition of the term "significant" eigenvalue may be more appropriate, especially if the eigenvalues occur in clusters. In order not to complicate the description of the method of recursive collocation, we refer to the remark I of Section 4.

From (2.1), we obtain

\[
U_l(x) = \sum_{r=1}^{\omega(k)} \tilde{A}_{k,r} \cdot \varepsilon_{k,r}(x) + \sum_{\mu=1}^{\tau(k)} (\tilde{B}_{k,\mu} c_{k,\mu}(x) + \tilde{C}_{k,\mu} s_{k,\mu}(x)) \cdot \varepsilon_{k,\mu}(x),
\]

where

\[
\tilde{A}_{k,r} = \lambda_{k,r} A_{k,r}, \quad \nu = 1, \ldots, \omega(k),
\]

\[
\tilde{B}_{k,\mu} = u_{k,\mu} B_{k,\mu} + \nu_{k,\mu} C_{k,\mu},
\]

\[
\tilde{C}_{k,\mu} = -\nu_{k,\mu} B_{k,\mu} + u_{k,\mu} C_{k,\mu}, \quad \mu = l + 1, \ldots, r(k).
\]

For a given value of \( k \), the unknown vectors \( \{A_{k,r}\}, \{B_{k,\mu}, C_{k,\mu}\} \) (which are in \( R^n \)) will be computed recursively by imposing the following conditions on the approximate solution \( U(x) \):

(A) if \( U(x) = (u^{(1)}(x), \ldots, u^{(n)}(x))^T \), then

\[
u^{(j)}(x) \in C(I), \quad j = 1, \ldots, n; \]

(B) for given points \( \{x_{k,j}\} \subset \sigma_k, \xi_k < x_{k,1} < \cdots < x_{k,m-1} \leq \xi_{k+1} \), the representation \( U_k(x) \) of \( U(x) \) in \( \sigma_k \) is to satisfy (1.1a) at these points (collocation points), i.e.,
\[ U'_k(x_{k,i}) = F(x_{k,i}, U_k(x_{k,i})), \quad j = 1, \ldots, m - 1. \]

Here, \( m = m(k) = \omega(k) + 2(\tau(k) - 1) \geq 1 \). In the case \( m = 1 \), the collocation condition (2.5) is empty.

Relation (2.5) constitutes a system of \((m - 1)n\) nonlinear equations for the \((mn)\) unknown components of \( \{A_{k,i}\}, \{B_{k,u}, C_{k,u}\} \), while (2.4), which may be rewritten as

\[ (2.6) \quad U_k(\xi_k) = \sum_{i=1}^{w(k)} A_{k,i} + \sum_{u=1}^{r(k)} B_{k,u} = U_{k-1}(\xi_k), \]

yields \( m \) additional (linear) equations for these unknown components. If the given system (1.1a) is stiff (i.e., its right-hand side possesses a large Lipschitz constant in \( I \)), direct functional iteration in (2.5) will not converge for widely spaced collocation points. Hence, one is forced to apply Newton’s method (or one of its various modifications; see also [5]) to the nonlinear system (2.5). This situation is similar to the one encountered when solving a stiff system (1.1) by an implicit Runge-Kutta or linear multistep method.

So far, a practically useful error analysis (i.e., an error analysis for finite values of \( N \), with the differences \( h_k = \xi_{k+1} - \xi_k \) not tending to zero) is not yet available. However, the application of the method of recursive collocation to numerous stiff systems (the majority of which is taken from [1]) has furnished approximate values of uniformly good accuracy, both for small and large values of \( h_k \). In the following section, we present one of these examples to illustrate this point.

### 3. Numerical Example

The following example of a very stiff system of nonlinear (homogeneous) ordinary differential equations arose in a problem of chemistry (see Gear [2], Bjurel [1]):

\[ y'_1(x) = -0.013 y_2(x) - 1000 y_1(x)y_2(x) - 2500 y_1(x)y_2(x), \quad 0 \leq x \leq 50, \]

\[ y'_2(x) = -0.013 y_3(x) - 1000 y_1(x)y_2(x), \]

\[ y'_3(x) = -2500 y_1(x)y_2(x), \quad 0 \leq x \leq 50, \]

with \( y_1(0) = 0, y_2(0) = 1, y_3(0) = 1 \).

Its Jacobian (computed along the exact solution) possesses three distinct real eigenvalues. A selection of these eigenvalues are contained in Table I.

The approximate solution \( U(x) \) (involving the approximate eigenvalues \( \{\lambda_{k,1}\} \) and \( \{\lambda_{k,2}\} \) corresponding to \( \lambda_1(x) \) and \( \lambda_2(x) \)) was computed for different step sizes (i.e., spacings of the collocation points) in the interval \([0, 1]\); for \( x \in (1, 50) \), the uniform step size \( h = 1.0 \) was chosen. The resulting systems of nonlinear equations (2.5) was solved by Newton’s method; here it was usually sufficient to perform one iteration step (using the previous set of coefficients as initial values) to come within a tolerance of \( 10^{-10} \).

In Table II, we list a sample of numerical results. Observe that, in column (a), the values given are those at the collocation points (i.e., the values generated as if the method of recursive collocation were used as a one-step method), whereas in columns (b) and (c), the collocation points for the interval \([0, 1]\) are \( \xi_1 = 0.5, \xi_2 = 1.0, \) and \( \xi_3 = 1.0, \) respectively.
Approximate values of similarly good accuracy were obtained for rather large step sizes. As an example, we chose for the spacing of the collocation points over the entire interval [0, 50] the value \( h = 5.0 \). The corresponding values for \( U(x) \) at \( x = 50 \) are as follows:

\[
U(50) = \begin{pmatrix} -1.893 \cdot 10^{-8} \\ 0.5974750 \\ 1.4025231 \end{pmatrix},
\]

compared with the exact values

\[
Y(50) = \begin{pmatrix} -1.893 \cdot 10^{-8} \\ 0.5976547 \\ 1.4023434 \end{pmatrix}.
\]

The error at \( x = 50 \) is

\[
E(50) = Y(50) - U(50) = \begin{pmatrix} -0.00071 \cdot 10^{-6} \\ 0.0001797 \\ -0.0001797 \end{pmatrix}.
\]

All the computations were performed on the CDC 6400 (single precision) at Dalhousie University Computer Centre.

4. Remarks. We conclude with some additional remarks regarding the practical implementation of the method of recursive collocation described in Section 2.

I. If the given system (1.1a) is large, one will, in general, in (2.1) not select the significant eigenvalues of \( G(\xi_k, U_k(\xi_k)) \) by means of the criterion (2.2), as briefly mentioned in Section 2. If some (real or complex) eigenvalues occur in one or several clusters, one will choose an appropriate eigenvalue ("centre" of cluster) to be representative for the particular cluster under consideration. (This choice of a representative eigenvalue somewhat resembles the choice of the parameter values in the method of exponential fitting; see Liniger and Willoughby [5, p. 56].) Furthermore, the eigenvalues of the approximate Jacobian \( G(\xi_k, U_k(\xi_k)) \) need not be computed to a great accuracy and, in the case where all the eigenvalues are real, the power
<table>
<thead>
<tr>
<th>$x$</th>
<th>$\gamma(x)$</th>
<th>$U(x)$</th>
<th>$V(x)$</th>
<th>$U(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.0)</td>
<td>(0.1)</td>
<td>(0.2)</td>
<td>(0.3)</td>
</tr>
<tr>
<td></td>
<td>(-3.709 \times 10^{-6})</td>
<td>(-3.699 \times 10^{-6})</td>
<td>(-3.704 \times 10^{-6})</td>
<td>(-3.700 \times 10^{-6})</td>
</tr>
<tr>
<td></td>
<td>(0.9990706)</td>
<td>(0.9990703)</td>
<td>(0.9981425)</td>
<td>(0.9972149)</td>
</tr>
<tr>
<td></td>
<td>(1.0009257)</td>
<td>(1.0009260)</td>
<td>(1.0018538)</td>
<td>(1.0027814)</td>
</tr>
</tbody>
</table>

Column (a): \(\pi_N\): \(\xi_i = jh\) \((h = 0.1), \ j = 0, 1, \ldots, 10; \)
\(\xi_i = (j - 9)h\) \((h = 1.0), \ j = 11, \ldots, 59.\)

Column (b): \(\pi_N\): \(\xi_i = jh\) \((h = 0.5), \ j = 0, 1, 2; \)
\(\xi_i = (j - 1)h\) \((h = 1.0), \ j = 3, \ldots, 51.\)

Column (c): \(\pi_N\): \(\xi_i = jh\) \((h = 1.0), \ j = 0, 1, \ldots, 50.\)
method (inverse iteration) may be used to obtain approximations to the significant
eigenvalues \( \{ \lambda_k \} \) (significant in the sense of criterion (2.2) or of the alternate criterion
stated above).

II. If the system (1.1a) is not homogeneous, then the local behavior of the exact
solution of (1.1) will not be described by the eigenvalues of the corresponding Jacobian.
As an example, consider a nonlinear system with a forcing term \( f(x) \) (see Bjurel [1]):

\[
\begin{align*}
y'_1(x) &= -(55 + y_3(x))y_1(x) + 65 y_2(x) + f(x), \\
y'_2(x) &= 0.0785 (y_1(x) - y_2(x)), \\
y'_3(x) &= 0.1 y_1(x),
\end{align*}
\]

with the initial conditions \( y_1(0) = 1, y_2(0) = 1, y_3(0) = 0 \). At \( x = 0 \), the eigenvalues
of the exact Jacobian are given by

\[
\lambda_1 = -55.091, \quad \lambda_{2,3} = 0.062 \pm 0.011 i.
\]

Here, for \( x \in \sigma_0 \), the approximate solution \( U(x) \) will in general be chosen such as
to include a (differentiable) function \( \varphi(x) \) reflecting the influence of the forcing
function \( f(x) \), i.e.,

\[
(4.1) \quad U_0(x) = (B_{0,1}c_{0,1}(x) + C_{0,1} s_{0,1}(x)) \cdot \varepsilon_{0,1}(x) + D_{0,1} \cdot \varphi(x),
\]

with

\[
\begin{align*}
c_{0,1}(x) &= \cos(0.011 x), \\
s_{0,1}(x) &= \sin(0.011 x), \\
\varepsilon_{0,1}(x) &= \exp(0.062 x).
\end{align*}
\]

For the subsequent intervals \( \{ \sigma_1, \cdots, \sigma_{N-1} \} \), the process is then continued in a
similar manner.

On the other hand, if the forcing function \( f(x) \) varies only slowly in \( I \), it will be
suggestive to reduce the number of unknowns in (4.1) by replacing this expression by

\[
U_0(x) = (B_{0,1}c_{0,1}(x) + C_{0,1} s_{0,1}(x)) \cdot \varepsilon_{0,1}(x),
\]

i.e., by following the approach described in Section 2. However, this obviously will
yield reasonable results only if \( m = \omega(k) + 2(\tau(k) - 1) \geq 2 \) in (2.5).