

Correction in the Dominant Space: A Numerical Technique for a Certain Class of Stiff Initial Value Problems

By P. Alfeld and J. D. Lambert

Abstract. Consider a stiff linear initial value problem $y' = A(x)y + g(x)$, where the eigenvalues of $A(x)$ may be separated into two sets, one of which dominates the other. The dominant eigenvalues and corresponding right and left eigenvectors may be computed by the power method. A technique is proposed which consists of taking one forward step by a conventional multistep method and then making a correction entirely in the subspace spanned by the eigenvectors corresponding to the dominant eigenvalues. A number of alternative corrections are proposed and discussed. It is shown that the technique is stable provided that the product of the steplength and each of the subdominant eigenvalues lies within the region of absolute stability of the multistep method. The application of the technique to nonlinear problems is discussed, and numerical results are reported.

1. Introduction. It is well known that the initial value problem

$$y' = f(x, y), \quad y(a) = \eta, \quad y, f \in \mathbf{R}^m,$$

causes serious computational difficulties if it is stiff, that is, if the eigenvalues of the Jacobian $\partial f/\partial y$ are widely distributed in the left half-plane. The presence of eigenvalues with very large negative real part makes it necessary severely to restrict the steplength if the numerical method employed has a finite region of absolute stability. On the other hand, if the method has a region of absolute stability which is infinite in the left half-plane (e.g. A -stability [3], $A(\alpha)$ -stability [11], A_0 -stability [2], stiff stability [7]), one inevitably finds that the method is implicit, and the resulting implicit difference equation must be solved at each step, not by simple iteration (which fails to converge unless the steplength is again severely restricted), but by the expensive Newton iteration. In the case of a linear problem, there are corresponding computational problems with matrix inversion when stiffness is present.

In a preliminary paper [8], one of us proposed a technique for dealing with a special form of stiff problem. Consider the linear problem

$$y' = A(x)y + g(x), \quad y(a) = \eta,$$

where the eigenvalues $\lambda^{(i)}(x)$ of $A(x)$ satisfy, in some interval of x ,

- (i) $\lambda^{(1)}(x)$ real and negative,
- (ii) $|\lambda^{(1)}(x)| \gg \max_{2 \leq i \leq m} |\lambda^{(i)}(x)|$.

The method proposed in [8], which is based on a geometrical argument, consists of

taking one step from x_n to x_{n+1} by a conventional explicit one-step method, which we call the basic method, and then adding a correction $\xi_{n+1}^{(1)}c^{(1)}(x_{n+1})$, where $c^{(1)}(x)$ is the normalized eigenvector corresponding to $\lambda^{(1)}(x)$. The correction factor $\xi_{n+1}^{(1)}$ is a function of $\lambda^{(1)}(x_{n+1})$, so that it is necessary to compute $\lambda^{(1)}(x)$ and $c^{(1)}(x)$ at each point of the discretization; this can be economically done by the power method, which is particularly efficient for a matrix with the structure outlined above. The effect of this correction is that the restriction on the steplength imposed by stability requirements is simply that which would pertain if the eigenvalue $\lambda^{(1)}$ were not present, and will therefore not be severe.

In this paper we consider generalizations of this idea. Firstly, we extend the class of problems to the case where the eigenvalues of $A(x)$ can be separated into two sets, one of which dominates the other, and assume that we are able to compute the dominant eigenvalues and corresponding right and left eigenvectors. We then consider a correction in the dominant space, that is, the subspace spanned by the dominant eigenvectors. A number of alternative correction factors is discussed, and it transpires that the choice corresponding to that made in [8] is not the best. Finally, we extend the class of basic methods from one-step to multistep; again, we find advantage in this, explicit linear multistep methods emerging as the best choice. The extension of the technique to the fully nonlinear case is also discussed, and numerical results reported. It should be noted that the technique is primarily intended for use in an interval in which transients corresponding to the dominant eigenvalues no longer contribute to the exact solution.

2. Correction in the Dominant Space. The techniques to be developed in this paper are applicable to the nonlinear initial value problem

$$(2.1) \quad y' = f(x, y), \quad y(a) = \eta, \quad y, f \in \mathbf{R}^m,$$

but a more natural development of the ideas can be given in the context of the linear variable coefficient problem

$$(2.2) \quad y' = A(x)y + g(x), \quad y(a) = \eta, \quad y, g \in \mathbf{R}^m,$$

where $A(x)$ is an $m \times m$ matrix. For the major part of this paper we shall adopt (2.2) as the standard problem, but in Section 5 the applicability of the techniques to (2.1) will be discussed. It is assumed throughout that, for all x , $A(x)$ has a complete set of eigenvectors. We assume the usual discretization $\{x_n | x_n = a + nh, n = 0, 1, \dots\}$, and denote by y_n an approximation to $y(x_n)$, the exact solution of the initial value problem.

Definition 1. Let $\lambda^{(i)}$ ($= \lambda^{(i)}(x)$), $i = 1, 2, \dots, m$, be the eigenvalues of $A(x)$. Then the problem (2.2) and the matrix $A(x)$ are said to be *separably stiff* in an interval I of x if there exists a constant integer s , $1 \leq s < m$, such that, for all $x \in I$, $\lambda^{(i)}$, $i = 1, 2, \dots, s$, are real, distinct, and negative, and

$$\min_{1 \leq i \leq s} |\lambda^{(i)}| \gg \max_{s+1 \leq i \leq m} |\lambda^{(i)}|.$$

Remark. We are considering problems whose stiffness arises from the existence of a set (which we assume to be small) of troublesome eigenvalues $\lambda^{(i)}, i = 1, 2, \dots, s$, which are real, negative, and well separated from the rest; note, however, that we do not require all the eigenvalues to have negative real parts.

We denote by $c_n^{(i)} (= c^{(i)}(x_n))$ and $d_n^{(i)} (= d^{(i)}(x_n)), i = 1, 2, \dots, m$, the right and left eigenvectors of $A(x_n)$. These are normalized according to the following rules:

- (i) $\langle c_n^{(i)}, d_n^{(i)} \rangle := (c_n^{(i)})^T d_n^{(i)} = 1, i = 1, 2, \dots, m.$
- (ii) $\|c_n^{(i)}\| := \langle c_n^{(i)}, c_n^{(i)} \rangle^{1/2} = 1, i = 1, 2, \dots, m.$
- (iii) The first nonvanishing component of $c_0^{(i)}$ is positive.
- (2.3) (iv) If ${}^t c_n^{(i)}$ denotes the t th component of $c_n^{(i)}, t = 1, 2, \dots, m$, and $\hat{r}^{(n)} c_n^{(i)}$ is the component of $c_n^{(i)}$ with maximum modulus (the first such component if the maximum is not unique), then

$$\operatorname{sgn} \{ \hat{r}^{(n)} c_{n+1}^{(i)} \} = \operatorname{sgn} \{ \hat{r}^{(n)} c_n^{(i)} \}, \quad i = 1, 2, \dots, m; n = 0, 1, 2, \dots$$

The set of right and left eigenvectors of $A(x_n)$ is then uniquely defined. Note that $\langle c_n^{(i)}, d_n^{(j)} \rangle = \delta_{ij}$, the Kronecker delta.

Definition 2. Let $\lambda_n^{(i)} = \lambda^{(i)}(x_n)$, and let the right and left eigenvectors $c_n^{(i)}$ and $d_n^{(i)}$ of $A(x_n)$ be normalized by (2.3). The *dominant* and *subdominant eigensystems* of $A(x_n)$ are defined to be $\{ \lambda_n^{(i)}, c_n^{(i)}, d_n^{(i)} \mid i = 1, 2, \dots, s \}$ and $\{ \lambda_n^{(i)}, c_n^{(i)}, d_n^{(i)} \mid i = s + 1, s + 2, \dots, m \}$, respectively. The subspaces spanned by $\{ c_n^{(i)} \mid i = 1, 2, \dots, s \}$ and $\{ c_n^{(i)} \mid i = s + 1, s + 2, \dots, m \}$ are defined to be the *dominant* and *subdominant spaces* at x_n , respectively.

Remark. The technique requires the explicit computation of the dominant eigensystem for each n . This can be efficiently performed by some variant of the power method; see, for example, Clint and Jennings [1]. Note that in the case $s = 1$, separable stiffness implies rapid convergence of the power method; moreover, provided that the eigensystem of $A(x)$ does not change too rapidly with x , the computed value for $c_n^{(1)}$ can be expected to be a good starting vector for the power method applied to $A(x_{n+1})$.

Since $\langle d^{(i)}(x), z(x) \rangle, z(x) \in \mathbb{R}^m$, is the component of $z(x)$ in the $c^{(i)}(x)$ direction, we may write the exact solution $y(x)$ of (2.2) in the form

$$(2.4) \quad y(x) = \sum_{i=1}^m \phi^{(i)}(x) c^{(i)}(x), \quad \phi^{(i)}(x) = \langle d^{(i)}(x), y(x) \rangle.$$

We also write

$$(2.5) \quad y'(x) = \sum_{i=1}^m \hat{\phi}^{(i)}(x) c^{(i)}(x), \quad \hat{\phi}^{(i)}(x) = \langle d^{(i)}(x), y'(x) \rangle.$$

The technique we propose consists of using a conventional discrete variable method, which we shall call the *basic method*, to advance the solution from x_n to x_{n+1} , and then applying a correction before proceeding to the next step. We consider basic methods which satisfy the following requirement:

Definition 3. A discrete variable k -step method is said to be *rational* if, when

applied to the test problem $y' = Ay$, A is a constant matrix, it yields a linear k -step difference equation of the form

$$y_{n+1} = \sum_{j=0}^{k-1} P_j(hA)y_{n-j},$$

where $P_j(\cdot)$, $j = 0, 1, \dots, k-1$, are rational functions.

Remark. Most conventional methods, such as Runge-Kutta (RK)—both explicit and implicit—linear multistep (LM) and predictor-corrector (PC) (provided these are applied in a mode $P(EC)^lE$ and not $P(EC)^l$) are rational methods. When the method is explicit, the rational functions $P_j(\cdot)$ are, in fact, polynomials.

The basic method will always be chosen to be a convergent rational method, and its application to advance the solution from x_n to x_{n+1} will be denoted by

$$(2.6) \quad y_{n+1} = By_n.$$

This notation does not imply that the basic method is necessarily one-step; nor does it imply that it is necessarily explicit, though in practice we shall always choose it to be explicit, in order to avoid the difficulties over matrix inversion (iteration, in the non-linear case) encountered with stiff systems. The correction process is designed to overcome the poor stability properties usually associated with explicit methods in a stiff context. The technique—correction in the dominant space (CDS)—then advances the numerical solution of (2.2) from x_n to x_{n+1} by the following scheme:

$$(2.7(i)) \quad \tilde{y}_{n+1} = By_n,$$

$$(2.7(ii)) \quad y_{n+1} = \tilde{y}_{n+1} + \sum_{i=1}^s \xi_{n+1}^{(i)} c_{n+1}^{(i)},$$

where the $\xi_{n+1}^{(i)}$, $i = 1, 2, \dots, s$, are scalar correction factors, which can be determined in a number of alternative ways to be discussed in the next section.

Definition 4. The CDS scheme (2.7) is said to be *dominantly stable* if all solutions of the resulting difference equation in $\{y_n\}$ when the method is applied to the test equation $y' = Ay$, A a constant separably stiff matrix, tend to zero as n tends to infinity, for all steplengths h such that $h\lambda^{(i)} \in \mathcal{R}_B$, $i = s+1, s+2, \dots, m$, where \mathcal{R}_B is the region of absolute stability of the basic method, and $\lambda^{(i)}$, $i = s+1, s+2, \dots, m$, are the subdominant eigenvalues of A .

Remark. A dominantly stable CDS scheme is not unconditionally stable in the sense of A -stability; however, the restrictions that stability imposes on the steplength will be just those that would arise if the dominant eigenvalues were not present, that is, if the problem were no longer stiff. While the scheme calls for the computation of the dominant eigensystem at each step, it should be noted that when the basic method is explicit (which will always be the case in practice), (2.7) is a completely explicit process. Note also that in certain circumstances (e.g. $s = 1$), excessive (separable) stiffness will actually increase the efficiency with which the dominant eigensystem can be computed.

The following theorem makes it easy to establish the dominant stability of the

various CDS schemes to be proposed later:

THEOREM 1. *A CDS scheme for which the basic method is a rational method is dominantly stable if and only if $\langle d^{(i)}, y_n \rangle \rightarrow 0$ as $n \rightarrow \infty, i = 1, 2, \dots, s$.*

Proof. Applying a CDS scheme, with a k -step rational basic method, to $y' = Ay$, A a constant separably stiff matrix, yields

$$\begin{aligned} \tilde{y}_{n+1} &= \sum_{j=0}^{k-1} P_j(hA)y_{n-j}, \quad P_j(\cdot) \text{ a rational function,} \\ y_{n+1} &= \tilde{y}_{n+1} + \sum_{i=1}^s \xi_{n+1}^{(i)} c^{(i)}, \end{aligned}$$

where, since A is constant, $c^{(i)}$ no longer depends on n . Since A has a complete set of eigenvectors, we may write

$$y_t = \sum_{i=1}^m \gamma_t^{(i)} c^{(i)}, \quad \gamma_t^{(i)} = \langle d^{(i)}, y_t \rangle, \quad t = 0, 1, 2, \dots$$

Thus,

$$\begin{aligned} y_{n+1} &= \sum_{i=1}^m \gamma_{n+1}^{(i)} c^{(i)} = \sum_{j=0}^{k-1} P_j(hA) \left\{ \sum_{i=1}^m \gamma_{n-j}^{(i)} c^{(i)} \right\} + \sum_{i=1}^s \xi_{n+1}^{(i)} c^{(i)} \\ &= \sum_{i=1}^m \left\{ \sum_{j=0}^{k-1} P_j(h\lambda^{(i)}) \gamma_{n-j}^{(i)} \right\} c^{(i)} + \sum_{i=1}^s \xi_{n+1}^{(i)} c^{(i)}, \end{aligned}$$

since $P_j(\cdot)$ is a rational function. It follows that

$$(2.8) \quad \gamma_{n+1}^{(i)} = \sum_{j=0}^{k-1} P_j(h\lambda^{(i)}) \gamma_{n-j}^{(i)}, \quad i = s + 1, s + 2, \dots, m.$$

Consider the basic method applied to the test equation $z' = \lambda z$, λ a complex scalar constant, $z \in \mathbf{R}^1$; it yields the difference equation

$$(2.9) \quad z_{n+1} = \sum_{j=0}^{k-1} P_j(h\lambda) z_{n-j}.$$

Now, for all $z_0, z_n \rightarrow 0$ as $n \rightarrow \infty$ if and only if $h\lambda \in \mathcal{R}_B$, where \mathcal{R}_B is the region of absolute stability of the basic method. On comparing (2.8) with (2.9), it follows that, for $i = s + 1, s + 2, \dots, m, \gamma_{n+1}^{(i)} \rightarrow 0$ as $n \rightarrow \infty$ if and only if $h\lambda^{(i)} \in \mathcal{R}_B, i = s + 1, s + 2, \dots, m$. Hence, $y_{n+1} \rightarrow 0$ as $n \rightarrow \infty$ if and only if $h\lambda^{(i)} \in \mathcal{R}_B, i = s + 1, s + 2, \dots, m$, and $\gamma_{n+1}^{(i)} (= \langle d^{(i)}, y_{n+1} \rangle) \rightarrow 0$ as $n \rightarrow \infty, i = 1, 2, \dots, s$; hence the result.

3. Choice of Correction Factors. Recall the initial value problem

$$(3.1) \quad y' = f(x, y) := A(x)y + g(x), \quad y(a) = \eta,$$

and the general CDS scheme

$$(3.2) \quad \tilde{y}_{n+1} = By_n, \quad y_{n+1} = \tilde{y}_{n+1} + \sum_{i=1}^s \xi_{n+1}^{(i)} c_{n+1}^{(i)}.$$

In this section we discuss various alternative ways of choosing the scalar correction factors $\xi_{n+1}^{(i)}$, $i = 1, 2, \dots, s$.

I. *Projection into the Subdominant Space.*

$$(3.3) \quad \xi_{n+1}^{(i)} = -\langle d_{n+1}^{(i)}, y_{n+1} \rangle.$$

With this choice, y_{n+1} will have zero components in each of the directions of the dominant eigenvectors; it follows from Theorem 1 that the resulting CDS scheme is dominantly stable. However, since $y(x)$ will not, in general, have zero components in the dominant directions, there will be a persistent error, E_I , where

$$(3.4) \quad E_I := \sum_{i=1}^s \langle d_{n+1}^{(i)}, y(x_{n+1}) - y_{n+1} \rangle c_{n+1}^{(i)} = \sum_{i=1}^s \phi^{(i)}(x_{n+1}) c_{n+1}^{(i)},$$

from (2.4). This persistent error is wholly in the dominant space, is independent of the basic method and the steplength, and depends only on the initial value problem. It will not, in general, be small, and for this reason this particular choice of correction factors is not recommended. It is included here because of its relevance to later choices we shall propose. Note that the local error in the subdominant space consists of the subdominant components of the local error of the basic method. (This will be true for all choices of correction factors.)

II. *Minimization of Gradient.* One interpretation of stiffness is that the function f is ill conditioned with respect to y . Thus, if $y_n = y(x_n)$, one finds with a conventional explicit method, that, while y_{n+1} may be an acceptable approximation to $y(x_{n+1})$, $f(x_{n+1}, y_{n+1})$ is a very poor approximation to $f(x_{n+1}, y(x_{n+1}))$. In the steady-state phase, integral curves which are neighbors to the solution curve coalesce extremely rapidly, with the result that the gradient on neighboring curves is very much larger than on the solution curve. One is thus motivated to make a correction which avoids regions of high gradient; we thus choose the correction factors $\xi_{n+1}^{(i)}$, $i = 1, 2, \dots, s$, to minimize $\|f(x_{n+1}, y_{n+1})\|$. We write f_{n+1} for $f(x_{n+1}, y_{n+1})$, and \tilde{f}_{n+1} for $f(x_{n+1}, \tilde{y}_{n+1})$. From (3.1) and (3.2) it follows that

$$(3.5) \quad f_{n+1} = \tilde{f}_{n+1} + \sum_{i=1}^s \lambda_{n+1}^{(i)} \xi_{n+1}^{(i)} c_{n+1}^{(i)},$$

whence

$$\begin{aligned} \|f_{n+1}\|^2 &= \|\tilde{f}_{n+1}\|^2 + 2 \sum_{i=1}^s \lambda_{n+1}^{(i)} \xi_{n+1}^{(i)} \langle c_{n+1}^{(i)}, \tilde{f}_{n+1} \rangle \\ &\quad + \sum_{i=1}^s \sum_{j=1}^s \lambda_{n+1}^{(i)} \lambda_{n+1}^{(j)} \xi_{n+1}^{(i)} \xi_{n+1}^{(j)} \langle c_{n+1}^{(i)}, c_{n+1}^{(j)} \rangle. \end{aligned}$$

After some manipulation, one finds that the values $\xi_{n+1}^{(i)}$, $i = 1, 2, \dots, s$, which minimize $\|f_{n+1}\|^2$ are given by

$$(3.6) \quad M\xi = -\tilde{F},$$

where

$$\xi = [\lambda_{n+1}^{(1)} \xi_{n+1}^{(1)}, \lambda_{n+1}^{(2)} \xi_{n+1}^{(2)}, \dots, \lambda_{n+1}^{(s)} \xi_{n+1}^{(s)}]^T,$$

$$\tilde{F} = [\langle c_{n+1}^{(1)}, \tilde{f}_{n+1} \rangle, \langle c_{n+1}^{(2)}, \tilde{f}_{n+1} \rangle, \dots, \langle c_{n+1}^{(s)}, \tilde{f}_{n+1} \rangle]^T,$$

and M is an $s \times s$ matrix with i, j th element $\langle c_{n+1}^{(i)}, c_{n+1}^{(j)} \rangle$. Thus, each of the s correction factors will depend on all of the dominant eigenvectors, a feature which we shall call *interdependence*. Note that, in the case $s = 1$, (3.6) reduces to

$$\xi_{n+1}^{(1)} = -\langle c_{n+1}^{(1)}, \tilde{f}_{n+1} \rangle / \lambda_{n+1}^{(1)},$$

which yields the scheme originally proposed by Lambert [8].

Note also that, in the special case where $A(x)$ is symmetric, the eigenvectors $c_{n+1}^{(i)}$ are orthogonal, and the matrix M in (3.6) becomes the identity matrix; the scheme is then no longer interdependent. Since, for general $A(x)$, the right eigenvectors are orthogonal to the left eigenvectors, it is natural to ask if there exists a variant of this approach in which inner products of right eigenvectors are replaced by inner products of right and left eigenvectors; it so happens that this arises naturally in the approach of the next section.

III. Gradient Projection Into the Subdominant Space. Following the motivation preceding Eq. (3.5), we can choose the correction factors not to minimize the norm of the gradient, but to remove entirely the components of the gradient in the dominant directions, since it is these components that are responsible for the ill-conditioning. Thus, for $i = 1, 2, \dots, s$, we set

$$\begin{aligned} 0 &= \langle d_{n+1}^{(i)}, f_{n+1} \rangle \\ &= \langle d_{n+1}^{(i)}, \tilde{f}_{n+1} \rangle + \sum_{j=1}^s \lambda_{n+1}^{(j)} \xi_{n+1}^{(j)} \langle d_{n+1}^{(i)}, c_{n+1}^{(j)} \rangle, \quad \text{by (3.5)} \\ &= \langle d_{n+1}^{(i)}, \tilde{f}_{n+1} \rangle + \lambda_{n+1}^{(i)} \xi_{n+1}^{(i)}. \end{aligned}$$

Hence, we choose the correction factors to be

$$(3.7) \quad \xi_{n+1}^{(i)} = -\langle d_{n+1}^{(i)}, \tilde{f}_{n+1} \rangle / \lambda_{n+1}^{(i)}, \quad i = 1, 2, \dots, s.$$

The scheme is now noninterdependent. Moreover, applying the test equation $y' = Ay$, A a constant separably stiff matrix, we obtain that

$$\begin{aligned} \langle d^{(i)}, y_{n+1} \rangle &= \left\langle d^{(i)}, \tilde{y}_{n+1} - \sum_{j=1}^s \frac{\langle d^{(j)}, \tilde{f}_{n+1} \rangle}{\lambda^{(j)}} c^{(j)} \right\rangle \\ &= \langle d^{(i)}, \tilde{y}_{n+1} \rangle - \langle d^{(i)}, \tilde{f}_{n+1} \rangle / \lambda^{(i)} \\ &= 0, \end{aligned}$$

since $\tilde{f}_{n+1} = A\tilde{y}_{n+1}$. Thus the scheme is, by Theorem 1, dominantly stable.

The choice (3.7) for the correction factors means that f_n will have zero components in the dominant directions; we cannot expect the same to be true, in general, for $f(x_n, y(x_n)) = y'(x_n)$, and we must again expect a persistent error. Applying the CDS

scheme (3.2) with correction factors (3.7) to the initial value problem (3.1) gives

$$y_{n+1} = \tilde{y}_{n+1} - \sum_{j=1}^s \frac{1}{\lambda_{n+1}^{(j)}} \langle d_{n+1}^{(j)}, A_{n+1} \tilde{y}_{n+1} + g_{n+1} \rangle c_{n+1}^{(j)},$$

where we have written A_{n+1} for $A(x_{n+1})$ and g_{n+1} for $g(x_{n+1})$. Hence,

$$\begin{aligned} \langle d_{n+1}^{(i)}, y_{n+1} \rangle &= \langle d_{n+1}^{(i)}, \tilde{y}_{n+1} \rangle - \langle d_{n+1}^{(i)}, A_{n+1} \tilde{y}_{n+1} \rangle / \lambda_{n+1}^{(i)} - \langle d_{n+1}^{(i)}, g_{n+1} \rangle / \lambda_{n+1}^{(i)} \\ &= -\langle d_{n+1}^{(i)}, g_{n+1} \rangle / \lambda_{n+1}^{(i)} \\ &= -\langle d_{n+1}^{(i)}, y'(x_{n+1}) - A_{n+1} y(x_{n+1}) \rangle / \lambda_{n+1}^{(i)}, \quad \text{from (3.1),} \\ &= -\hat{\phi}^{(i)}(x_{n+1}) / \lambda_{n+1}^{(i)} + \langle d_{n+1}^{(i)}, y(x_{n+1}) \rangle, \quad \text{by (2.5).} \end{aligned}$$

Hence the persistent error E_{III} is given by

$$(3.8) \quad E_{\text{III}} := \sum_{i=1}^s \langle d_{n+1}^{(i)}, y(x_{n+1}) - y_{n+1} \rangle c_{n+1}^{(i)} = \sum_{i=1}^s \frac{\hat{\phi}^{(i)}(x_{n+1})}{\lambda_{n+1}^{(i)}} c_{n+1}^{(i)}.$$

On comparing (3.8) with (3.4), it is clear that although E_{I} will not, in general, be small, due to the presence of the dominant eigenvalues in the denominator, E_{III} will indeed be small, provided that $y(x)$ does not have an abnormally large gradient; and this will normally be the case, at least in the steady state region. Note that the persistent error, being independent of the basic method and the discretization, will not accumulate in the dominant space as the computation proceeds.

For problems for which the persistent error (3.8) is not acceptably small, a *further improvement* can be carried out. The aim is to approximate the persistent error (3.8) without calling for evaluation of the ill-conditioned function f . Suppose a solution for the problem (3.1) is required on the interval $[x_0, x_N]$, and that the basic method employed is a k -step method. (The necessary additional starting values y_1, y_2, \dots, y_{k-1} are assumed to have been computed by some implicit method with adequate stability.) Let $\tau = [k/2]$. The further improvement procedure then consists of the following steps:

(i) Calculate the sequence $\{y_n \mid n = k, k+1, \dots, N-k+\tau\}$ using the CDS scheme (3.2) with correction factors given by (3.7).

(ii) Let $\pi_n(x)$ be the vector of unique polynomials of degree k which interpolate componentwise the values $y_{n-\tau}, y_{n-\tau+1}, \dots, y_{n-\tau+k}$. For $n = k, k+1, \dots, N$, compute $\dot{y}_n := \pi_n'(x_n)$. The vector \dot{y}_n is taken as an approximation to $y'(x_n)$, and $\langle d_n^{(i)}, \dot{y}_n \rangle$ as an approximation to $\hat{\phi}_i(x_n)$. (See Eq. (2.5).)

(iii) Compute the improved solution $\{Y_n \mid n = k, k+1, \dots, N\}$ defined by

$$Y_n := y_n + \sum_{i=1}^s \frac{1}{\lambda_n^{(i)}} \langle d_n^{(i)}, \dot{y}_n \rangle c_n^{(i)}.$$

Note that further improvement is an entirely *a posteriori* process, and so does not affect the dominant stability of the scheme. Note also that if one is interested in obtaining an accurate solution only in some subinterval of $[x_0, x_N]$ (for example, only in the steady-state region), then it is enough to compute the improved solution only in that subinterval.

It can be shown that, after further improvement, the component of the residual persistent error in the $c_n^{(i)}$ direction is proportional to $(\lambda_n^{(i)})^{-2}$; for reasonably stiff problems this can certainly be ignored.

As an alternative, it is of course possible to apply the improvement technique at each step. Moreover, other forms of improvement are possible, but the one described above has been chosen, for the numerical examples, because it is the best conditioned.

IV. *Reduction to Scalar Problem.* The choice (3.3) of correction factors in I led to the persistent error given by (3.4). The persistent error could be avoided if it were possible to make the choice

$$(3.9) \quad \xi_{n+1}^{(i)} = -\langle d_{n+1}^{(i)}, \tilde{y}_{n+1} \rangle + \phi^{(i)}(x_{n+1}), \quad i = 1, 2, \dots, s.$$

However, we do not know $\phi^{(i)}(x_{n+1})$, the component of $y(x_{n+1})$ in the $c_{n+1}^{(i)}$ direction. An approximation to it can be obtained as follows. Recall that $\phi^{(i)}(x) := \langle d^{(i)}(x), y(x) \rangle$. Consider the function

$$(3.10) \quad \psi_{n+1}^{(i)}(x) := \langle d_{n+1}^{(i)}, y(x) \rangle.$$

Then clearly $\psi_{n+1}^{(i)}(x_{n+1}) = \phi^{(i)}(x_{n+1})$. The function $\psi_{n+1}^{(i)}(x)$ satisfies the initial value problem

$$(3.11) \quad \psi_{n+1}^{(i)'}(x) = \langle d_{n+1}^{(i)}, f(x, y(x)) \rangle, \quad \psi_{n+1}^{(i)}(x_n) = \langle d_{n+1}^{(i)}, y(x_n) \rangle.$$

We consider a neighboring initial value problem, whose solution, $\kappa_{n+1}^{(i)}(x)$, is an approximation to $\psi_{n+1}^{(i)}(x)$. We do this by replacing $y(x)$ in the first of (3.11) by

$$y_n + (\kappa_{n+1}^{(i)}(x) - \langle d_{n+1}^{(i)}, y_n \rangle) c_{n+1}^{(i)},$$

since $\kappa_{n+1}^{(i)}(x)$ approximates in the interval $[x_n, x_{n+1}]$ the component of y in the $c_{n+1}^{(i)}$ direction. We also replace $y(x_n)$ in the second of (3.11) by y_n to obtain the initial value problem

$$(3.12) \quad \begin{aligned} \kappa_{n+1}^{(i)'}(x) &= \langle d_{n+1}^{(i)}, f(x, y_n + (\kappa_{n+1}^{(i)}(x) - \langle d_{n+1}^{(i)}, y_n \rangle) c_{n+1}^{(i)}) \rangle, \\ \kappa_{n+1}^{(i)}(x_n) &= \langle d_{n+1}^{(i)}, y_n \rangle. \end{aligned}$$

A straightforward expansion shows that, if we assume that $y_n = y(x_n)$, then

$$\kappa_{n+1}^{(i)}(x_{n+1}) - \phi^{(i)}(x_{n+1}) = O(h^3).$$

It follows that in solving the problem (3.12) numerically it is adequate to use a method with local truncation error $O(h^3)$. It transpires that dominant stability of the eventual CDS scheme can be achieved provided that an A -stable method is used to solve (3.12); we therefore use the Trapezoidal Rule. Note that in order to evaluate all s correction factors, s initial value problems of the form (3.12) have to be solved; but these s problems are uncoupled, and the implicitness of the Trapezoidal Rule does not involve us in any matrix inversions. Note also that we use the Trapezoidal Rule with step h for only one step, and accept the resulting approximation to $\kappa_{n+1}^{(i)}(x_{n+1})$, which we shall denote by $\kappa_{n+1}^{(i)}$, as an approximation for $\phi^{(i)}(x_{n+1})$ in (3.9). Solving (3.12) by the

Trapezoidal Rule gives

$$\begin{aligned} \kappa_{n+1}^{(i)} - \langle d_{n+1}^{(i)}, y_n \rangle &= \frac{h}{2} \langle d_{n+1}^{(i)}, A_{n+1} \{ y_n + (\kappa_{n+1}^{(i)} - \langle d_{n+1}^{(i)}, y_n \rangle) c_{n+1}^{(i)} \} + g_{n+1} \rangle \\ &+ \frac{h}{2} \langle d_{n+1}^{(i)}, A_n y_n + g_n \rangle. \end{aligned}$$

Now, $y_n - \langle d_{n+1}^{(i)}, y_n \rangle c_{n+1}^{(i)}$ is a vector with zero component in the $c_{n+1}^{(i)}$ direction. The same must be true for the vector $A_{n+1}(y_n - \langle d_{n+1}^{(i)}, y_n \rangle c_{n+1}^{(i)})$, and it follows that the inner product of this latter vector with $d_{n+1}^{(i)}$ is zero. Hence,

$$\kappa_{n+1}^{(i)} - \langle d_{n+1}^{(i)}, y_n \rangle = \frac{h}{2} \{ \lambda_{n+1}^{(i)} \kappa_{n+1}^{(i)} + \langle d_{n+1}^{(i)}, g_{n+1} \rangle + \langle d_{n+1}^{(i)}, f(x_n, y_n) \rangle \}$$

or

$$(3.13) \quad \kappa_{n+1}^{(i)} = \langle d_{n+1}^{(i)}, y_n + \frac{1}{2} h (f_n + g_{n+1}) \rangle / (1 - \frac{1}{2} h \lambda_{n+1}^{(i)}), \quad i = 1, 2, \dots, s,$$

where $f_n = A_n y_n + g_n$. The correction factors are now given by

$$(3.14) \quad \xi_{n+1}^{(i)} = -\langle d_{n+1}^{(i)}, \tilde{y}_{n+1} \rangle + \kappa_{n+1}^{(i)}, \quad i = 1, 2, \dots, s,$$

where the $\kappa_{n+1}^{(i)}$ are given by (3.13).

In order to establish the dominant stability of the resulting CDS scheme, it is necessary, from Theorem 1, to show that $\langle d^{(i)}, y_n \rangle \rightarrow 0$ as $n \rightarrow \infty$, $i = 1, 2, \dots, s$, in the case when $f = Ay$. In this case, (3.13) reduces to

$$\begin{aligned} \kappa_{n+1}^{(i)} (= \kappa^{(i)}) &= \langle d^{(i)}, (I + \frac{1}{2} h A) y_n \rangle / (1 - \frac{1}{2} h \lambda^{(i)}) \\ &= \frac{1 + \frac{1}{2} h \lambda^{(i)}}{1 - \frac{1}{2} h \lambda^{(i)}} \langle d^{(i)}, y_n \rangle, \quad i = 1, 2, \dots, s. \end{aligned}$$

From (2.7) and (3.14), the CDS scheme reduces to

$$y_{n+1} = \tilde{y}_{n+1} + \sum_{j=1}^s (\kappa^{(j)} - \langle d^{(j)}, \tilde{y}_{n+1} \rangle) c^{(j)},$$

whence

$$\langle d^{(i)}, y_{n+1} \rangle = \kappa^{(i)} = \frac{1 + \frac{1}{2} h \lambda^{(i)}}{1 - \frac{1}{2} h \lambda^{(i)}} \langle d^{(i)}, y_n \rangle, \quad i = 1, 2, \dots, s.$$

It follows that, for all positive h , $\langle d^{(i)}, y_n \rangle \rightarrow 0$ as $n \rightarrow \infty$, $i = 1, 2, \dots, s$.

Other forms of reduction to scalar are possible. Thus since

$$\phi^{(i)}(x) = \langle d^{(i)}(x), y(x) \rangle,$$

we could consider in place of the initial value problem (3.11) the problem

$$\phi^{(i)'}(x) = \langle d^{(i)'}(x), y(x) \rangle + \langle d^{(i)}(x), f(x, y(x)) \rangle,$$

$$\phi^{(i)}(x_n) = \langle d_n^{(i)}, y(x_n) \rangle,$$

and proceed as before, using numerical differentiation to estimate the derivatives of $d^{(i)}(x)$. In practice, the resulting CDS scheme is less accurate than the one determined by (3.14).

4. Choice of Basic Method. Numerical experimentation has shown that the computational results given by a CDS scheme can be adversely affected by an ill-considered choice of basic method. Consider the test problem $y' = Ay$. Since the basic method is always chosen to be a rational method, the CDS scheme may be written

$$(4.1) \quad y_{n+1} = \sum_{j=0}^{k-1} P_j(hA)y_{n-j}, \quad y_{n+1} = \tilde{y}_{n+1} + \sum_{i=1}^s \xi_{n+1}^{(i)} c^{(i)},$$

where k is the stepnumber of the basic method and, in the case of an explicit basic method, the functions $P_j(\cdot)$ are polynomials.

Consider, for example, the choice III for the correction factors. By (3.7), we have

$$\begin{aligned} \xi_{n+1}^{(i)} &= - \left\langle d^{(i)}, A \sum_{j=0}^{k-1} P_j(hA)y_{n-j} \right\rangle / \lambda^{(i)} \\ &= - \sum_{j=0}^{k-1} P_j(h\lambda^{(i)})(d^{(i)}, y_{n-j}), \quad i = 1, 2, \dots, s. \end{aligned}$$

Recalling that, for $i = 1, 2, \dots, s$, $|\lambda^{(i)}|$ can be expected to be large, it is clear that if the polynomials $P_j(\cdot)$ have high degree, then extremely large correction factors can be expected. This can cause amplification of round-off error in the second stage of (4.1), and can even result in overflow. For an explicit R -stage RK method, $k = 1$ and $P_0(\cdot)$ has degree R . For a PC method in mode $P(EC)^lE$, $P_j(\cdot)$ has degree $l + 1$, whereas for an explicit LM method, $P_j(\cdot)$ has degree one. We thus conclude that, in the general case, the best choice we can make for the basic method is an explicit LM method. Similar arguments in favor of LM methods can be made for CDS schemes employing the other correction factors discussed in Section 3, and numerical experimentation fully corroborates our conclusion.

There may, of course, exist separably stiff problems with special structure for which we can do better. An example would be where the subdominant eigenvalues all have very small negative real parts and imaginary parts which are not small, and $g(x)$ is a slowly varying function. For such a problem, after the transient stage corresponding to the dominant eigenvalues is negligible, the solution will be a slowly damped oscillation. It may then be appropriate to choose as basic method one of the special methods of Gautschi [6], based on trigonometric interpolation.

5. Extension to Nonlinear Problems. We now turn to the nonlinear problem (2.1),

$$y' = f(x, y), \quad y(a) = \eta, \quad y, f \in \mathbf{R}^m.$$

A natural extension of the CDS techniques so far described consists of replacing the eigensystem of $A(x)$ in (2.2) by the eigensystem of the Jacobian $\partial f(x, y)/\partial y$ of (2.1), namely

$$\{\lambda^{(i)}(x, y), c^{(i)}(x, y), d^{(i)}(x, y) \mid i = 1, 2, \dots, s, s + 1, \dots, m\}.$$

Previously, we made extensive use of the eigensystem of $A(x)$ evaluated at x_{n+1} . If

we replace this by the eigensystem of $\partial f/\partial y$ evaluated at (x_{n+1}, y_{n+1}) , then we clearly introduce complicated implicitness in y_{n+1} , calling for involved iteration schemes. Numerical experimentation with such schemes has, however, indicated that it is adequate to use the eigensystem evaluated at $(x_{n+1}, \tilde{y}_{n+1})$. We introduce the following notation:

$$\tilde{\lambda}_n^{(i)} := \lambda^{(i)}(x_n, \tilde{y}_n), \quad \tilde{c}_n^{(i)} := c^{(i)}(x_n, \tilde{y}_n), \quad \tilde{d}_n^{(i)} := d^{(i)}(x_n, \tilde{y}_n),$$

$$i = 1, 2, \dots, m.$$

The equation (2.7) describing the general CDS scheme is now replaced by

$$\tilde{y}_{n+1} = B y_n, \quad y_{n+1} = \tilde{y}_{n+1} + \sum_{i=1}^s \xi_{n+1}^{(i)} \tilde{c}_{n+1}^{(i)}.$$

Adaptations to the nonlinear case of the various correction factors considered in Section 3 are given below.

I. *Projection Into the Subdominant Space.* Formula (3.3) for the correction factor is unchanged, except that in (3.3) $d_{n+1}^{(i)}$ is replaced by $\tilde{d}_{n+1}^{(i)}$.

II. *Minimization of Gradient.* It is no longer possible to solve the minimization problem analytically, but any appropriate numerical algorithm can be applied to find the correction factors $\xi_{n+1}^{(i)}$, $i = 1, 2, \dots, s$, which minimize

$$\|f_{n+1}\|^2 = \left\| f \left(x_{n+1}, \tilde{y}_{n+1} + \sum_{i=1}^s \xi_{n+1}^{(i)} \tilde{c}_{n+1}^{(i)} \right) \right\|^2.$$

In view of the remarks made in Section 3, we do not pursue this approach further.

III. *Gradient Projection Into the Subdominant Space.* We now have to solve for the correction factors $\xi_{n+1}^{(i)}$, $i = 1, 2, \dots, s$, the equations

$$(5.1) \quad 0 = \langle \tilde{d}_{n+1}^{(i)}, f_{n+1} \rangle = \left\langle \tilde{d}_{n+1}^{(i)}, f \left(x_{n+1}, \tilde{y}_{n+1} + \sum_{i=1}^s \xi_{n+1}^{(i)} \tilde{c}_{n+1}^{(i)} \right) \right\rangle,$$

$$i = 1, 2, \dots, s.$$

This is achieved by the following iteration

$$[\xi_{n+1}^{(i)}]^{[0]} = \xi_n^{(i)},$$

$$(5.2) \quad [\xi_{n+1}^{(i)}]^{[t+1]} = [\xi_{n+1}^{(i)}]^{[t]} - \frac{1}{\tilde{\lambda}_{n+1}^{(i)}} \left\langle \tilde{d}_{n+1}^{(i)}, f \left(x_{n+1}, \tilde{y}_{n+1} + \sum_{i=1}^s [\xi_{n+1}^{(i)}]^{[t]} \tilde{c}_{n+1}^{(i)} \right) \right\rangle,$$

$$i = 1, 2, \dots, s, t = 0, 1, 2, \dots$$

A straightforward calculation shows that in the linear case, $f(x, y) = A(x)y + g(x)$, this iteration converges in one step to give the correction factor defined by (3.7). The convergence of (5.2) is controlled, not by the stiffness of the system, but by the rate of change of the eigensystem. If (5.1) has a unique solution convergence is achieved for

a steplength which is sufficiently small in relation to this rate of change (and this would be a normal requirement even for a nonstiff system). In practice, convergence is attained in about three iterations, without any undue restriction on steplength. The further improvement technique described in Section 3 can be applied without modification.

It is possible to construct examples for which the system of equations (5.1) has no solution, or has multiple solutions. For this reason, although successful results can be obtained for some nonlinear problems, this choice of correction factor cannot be recommended for the general nonlinear separably stiff problem.

IV. *Reduction to Scalar Problem.* The analysis of Section 3-IV, up to Eq. (3.12), holds for the nonlinear case, but now it is no longer possible to write explicit expressions for the correction factors which correspond to solving the initial value problem (3.12) by one step of the Trapezoidal Rule. We shorten the notation by writing

$$\langle \tilde{d}_{n+1}^{(i)}, f(x, y_n + (z - \langle \tilde{d}_{n+1}^{(i)}, y_n \rangle) \tilde{c}_{n+1}^{(i)}) \rangle =: F(x, z), \quad z \in \mathbf{R}^1.$$

The initial value problem corresponding to (3.12) is then

$$(5.3) \quad \kappa_{n+1}^{(i)}(x) = F(x, \kappa_{n+1}^{(i)}(x)), \quad \kappa_{n+1}^{(i)}(x_n) = \langle \tilde{d}_{n+1}^{(i)}, y_n \rangle.$$

Applying one step of the Trapezoidal Rule to this problem and, as in Section 3-IV, indicating the approximation to $\kappa_{n+1}^{(i)}(x_{n+1})$ so obtained by $\kappa_{n+1}^{(i)}$, yields the nonlinear equation

$$(5.4) \quad \begin{aligned} \kappa_{n+1}^{(i)} - \langle \tilde{d}_{n+1}^{(i)}, y_n \rangle - \frac{1}{2}h[F(x_{n+1}, \kappa_{n+1}^{(i)}) + F(x_n, \langle \tilde{d}_{n+1}^{(i)}, y_n \rangle)] \\ =: \Phi(\kappa_{n+1}^{(i)}) = 0. \end{aligned}$$

Finally, (5.3) is solved by the following iteration:

$$(5.5) \quad \begin{aligned} [\kappa_{n+1}^{(i)}]^{[0]} &= \kappa_n^{(i)}, \\ [\kappa_{n+1}^{(i)}]^{[t+1]} &= [\kappa_{n+1}^{(i)}]^{[t]} - \Phi([\kappa_{n+1}^{(i)}]^{[t]}) / (1 - \frac{1}{2}h\tilde{\lambda}_{n+1}^{(i)}), \\ & \quad i = 1, 2, \dots, s, t = 0, 1, 2, \dots \end{aligned}$$

Again, it is straightforward to show that in the linear case, $f(x, y) = A(x)y + g(x)$, the iteration (5.5) converges in one step to give formula (3.13). The convergence of (5.5), like that of (5.2), is essentially dependent on the rate of change of the eigensystem, and not on the stiffness. Note however that whereas (5.2) constitutes a simultaneous iteration in all s unknowns $\xi_{n+1}^{(i)}, i = 1, 2, \dots, s$, the iteration (5.5) can be performed separately on each of the s unknowns $\kappa_{n+1}^{(i)}, i = 1, 2, \dots, s$. Once convergence of (5.5) has been achieved, the correction factors are given by

$$\xi_{n+1}^{(i)} = -\langle \tilde{d}_{n+1}^{(i)}, \tilde{y}_{n+1} \rangle + \kappa_{n+1}^{(i)}, \quad i = 1, 2, \dots, s.$$

(Cf. Eq. (3.14).)

6. Numerical Examples. Two separably stiff examples are considered, one linear and one nonlinear; in each case $m = 3, s = 1$. The CDS schemes employed are labelled II, III, III*, IV, corresponding to the following choices for the correction factors:

II. Minimization of gradient.

III. Gradient projection.

III*. Gradient projection with further improvement.

IV. Reduction to scalar.

The basic method is chosen to be the fourth-order (explicit) Adams-Bashforth method, with steplength 0.1. The following measures of the dominant and subdominant errors are computed:

$$E_D := \max_{n=k, k+1, \dots, N} |\langle d_n^{(1)}, y(x_n) - y_n \rangle|,$$

$$E_S := \max_{n=k, k+1, \dots, N} \|y(x_n) - y_n - \langle d_n^{(1)}, y(x_n) - y_n \rangle c_n^{(1)}\|_\infty.$$

These errors are quoted for $N = 21$.

Example 1 (Linear).

$$y' = A(x)(z - z(x)) + z'(x), \quad y(0) = z(0).$$

(Note that the general linear problem (2.2) can always be put in this form.)

$$A(x) = \frac{1}{v(x) - 1} \begin{bmatrix} \alpha v(x) - \beta & \beta - \alpha & (\beta - \alpha)v(x) \\ (\gamma - \beta)v(x) & \beta v(x) - \gamma & \beta - \gamma \\ (\alpha - \gamma)v^2(x) & (\gamma - \alpha)v(x) & \gamma v(x) - \alpha \end{bmatrix},$$

where $\alpha = -10,000, \beta = -1/2, \gamma = -1/3, v(x) = 45x/23 - 5$. The exact solution is $y(x) = z(x) = e^{x/10} [-2, 6, 10]^T$. (Recall that the CDS schemes are intended for use in the steady-state region only; thus we have chosen initial conditions which avoid transients.)

The eigensystem of $A(x)$ is:

$$\lambda^{(1)} = \alpha; \quad c^{(1)} = a(x)[1, 0, v(x)]^T; \quad d^{(1)} = b(x)[v(x), -1, -1/v(x)]^T,$$

$$\lambda^{(2)} = \beta; \quad c^{(2)} = a(x)[1, v(x), 0]^T; \quad d^{(2)} = b(x)[-1, 1, 1/v(x)]^T,$$

$$\lambda^{(3)} = \gamma; \quad c^{(3)} = a(x)[0, 1, -v(x)]^T; \quad d^{(3)} = b(x)[v(x), -1, -1]^T,$$

where

$$a(x) = 1/\sqrt{1 + v^2(x)}, \quad b(x) = \sqrt{1 + v^2(x)}/(v(x) - 1).$$

(In the numerical test, the dominant eigensystem was not computed from the above information, but is obtained at each step by the power method.)

The errors E_D and E_S for this example are as follows:

CDS scheme	E_D	E_S
II	$1.15_{10} - 4$	$2.60_{10} - 2$
III	$6.12_{10} - 5$	$8.58_{10} - 3$
III*	$2.35_{10} - 6$	$8.58_{10} - 3$
IV	$7.55_{10} - 10$	$6.86_{10} - 8$

Note that further improvement (III*) substantially improves the dominant error, and leaves the subdominant error unchanged. The superiority of method IV (reduction to scalar) is clear.

Example 2 (Nonlinear).

$$y' = u(x, y) - u(x, z(x)) + z'(x), \quad y(0) = z(0),$$

where $y = [y_1, y_2, y_3]^T$,

$$u(x, y) = \left[\frac{1}{3} \gamma y_1^3 + w(x)y_2 - w(x)y_3, \frac{1}{3} \beta y_2^3 + w(x)y_3, \frac{1}{3} \alpha y_3^3 \right]^T,$$

$$\alpha = -10,000, \quad \beta = -\frac{1}{2}, \quad \gamma = \frac{-1}{3}, \quad w(x) = -160(x - 1.25).$$

The exact solution is $y(x) = z(x) = e^{x/10} [1, 1, 1]^T/3$. The eigenvalues of the Jacobian are:

$$\lambda^{(1)} = \alpha y_1^2, \quad \lambda^{(2)} = \beta y_2^2, \quad \lambda^{(3)} = \gamma y_3^2.$$

The expressions for the eigenvectors of the Jacobian are too complicated to be worth quoting, but an indication of the rate of change of the dominant eigenvector is given by the following numerical values:

x	0	0.4	0.9	1.4	1.0	2.3
$c^{(1)}$	0.204	0.124	0.044	-0.016	-0.060	-0.086
	-0.173	-0.111	-0.042	0.016	0.064	0.095
	0.963	0.986	0.998	1.000	0.996	0.992

(Once again, in the numerical test the power method was used to find the dominant eigensystem at each step.)

The errors E_D and E_S for this example are as follows:

CDS scheme	E_D	E_S
II	$2.96_{10} - 5$	$1.32_{10} - 2$
III	$2.99_{10} - 5$	$1.30_{10} - 2$
III*	$1.34_{10} - 6$	$1.30_{10} - 2$
IV	$4.50_{10} - 10$	$1.04_{10} - 7$

Once again, the superiority of reduction to scalar is clear.

7. Conclusions. The CDS schemes proposed in this paper have some features which appear to be novel in the context of stiffness. Provided one is prepared to compute the dominant eigensystem at each step of the computation, then the remainder of the process is, for a linear separably stiff problem, completely explicit, and no matrix inversions are required. For a nonlinear separably stiff problem, there is some implicitness (in a space of dimension s , not m), but this presents none of the computational difficulties usually associated with implicit methods for stiff problems; in particular, no Newton iterations are required. The computational effort in finding the dominant eigensystem can (for example, the case $s = 1$) diminish as the stiffness increases. Moreover, the dominant components of the local error also, in general, diminish as the stiffness increases, since, for all the correction factors recommended, these components contain terms which are inversely proportional to the dominant eigenvalues.

Three alternative methods of defining correction factors have been proposed, and these have been referred to as minimization of gradient (MG), gradient projection into the subdominant space (GP), and reduction to scalar problem (RS). The MG and GP schemes are closely related, and indeed coincide for the linear problem $y' = A(x) + g(x)$, when $A(x)$ is symmetric; they also coincide asymptotically as $\lambda^{(1)} \rightarrow -\infty$. In the numerical examples quoted, and in other experiments, MG and GP have produced very comparable accuracy. The advantage of MG is that it does not require computation of the left eigenvectors, while the advantage of GP is that the structure of the dominant error is much simpler, and enables us to devise a further improvement process which considerably reduces the dominant error; such a process cannot be applied to MG. The disadvantage of both MG and GP is that they produce persistent errors which are independent of the steplength and the basic method, and do not accumulate in the dominant space. These errors do, however, pollute the subdominant error, to a certain extent, due to a phenomenon we call *interprojection*, which arises whenever the eigensystem is not constant. Since the basic method uses back values y_{n-j} , the errors in the local dominant directions $c_{n-j}^{(i)}$, $i = 1, 2, \dots, s$, will be partially projected into the new subdominant directions $c_{n+1}^{(i)}$, $i = s + 1, s + 2, \dots, m$. To control this effect it is necessary to choose the steplength to be suitably small in relation to the rate of change of the eigensystem. The last method, RS, although conceptually less simple and computationally somewhat more expensive than MG or GP, has turned out, in the numerical examples quoted and in other experiments, easily to be the most accurate and robust scheme. It does not suffer from any persistent error, and in the case of a nonlinear problem with $s > 1$, the iterations involved remain scalar (cf. Eqs. (5.4) and (5.2)); its only disadvantage would appear to be that its accuracy is limited by that of the approximating initial value problem (3.12) or (5.3). We thus recommend RS as the best CDS scheme.

There are, however, two situations in which MG and GP can play a useful role. If a nonlinear problem is being solved by RS and trouble is experienced in attaining convergence of the iteration (5.4), then a better initial approximation can be obtained by using MG or GP, rather than by reducing the steplength. This device has been used successfully in numerical experiments. The second situation concerns the case $s > 1$.

In order to minimize the gradient or to project it into the subdominant space it is clearly not necessary to have any particular basis of the dominant space; any basis would do. Stewart [10] proposes a method for computing an orthonormal basis for the invariant dominant subspace of a given matrix which appears to be more efficient than methods which compute the dominant eigenvectors directly.

Finally, we note an interesting relationship that exists between a CDS scheme based on GP and the technique, known as SAPS, developed by Dahlquist ([4], [5], [9]). That technique also requires separable stiffness, and considers the differential-algebraic system

$$f_1(u, v) = 0,$$

$$u \in \mathbf{R}^q, v \in \mathbf{R}^p.$$

$$\frac{dv}{du} = f_2(u, v),$$

It transpires that, after some manipulation, it is possible to write the CDS-GP scheme in the above form, both for the linear and the nonlinear case. However, SAPS proceeds by a mixture of analytical and numerical techniques, while CDS-GP is purely numerical.

Acknowledgements. The first author wishes to acknowledge the financial support afforded by the German National Scholarship Foundation and the University of Dundee during the period when this work was carried out. Both authors wish to acknowledge useful remarks made at an early stage of the development by Dirk Laurie of the National Research Institute for Mathematical Sciences, Pretoria.

Department of Mathematics
University of Dundee
Dundee DD1 4HN, Scotland

1. M. CLINT & A. JENNINGS, "A simultaneous iteration method for the unsymmetric eigenvalue problem", *J. Inst. Math. Appl.*, v. 8, 1971, pp. 111–121. MR 45 #6174.
2. C. W. CRYER, "A new class of highly-stable methods: A_0 -stable methods," *BIT*, v. 13, 1973, pp. 153–159. MR 48 #1469.
3. G. G. DAHLQUIST, "A special stability problem for linear multistep methods," *BIT*, v. 3, 1963, pp. 27–43. MR 30 #715.
4. G. G. DAHLQUIST, "A numerical method for some ordinary differential equations with large Lipschitz constants," *Information Processing 68*. Vol. 1 (Proc. IFIP Congress, Edinburgh, 1968), A. J. H. Morell, Editor, North-Holland, Amsterdam, 1969, pp. 183–186. MR 41 #2937.
5. G. G. DAHLQUIST, "Problems related to the numerical treatment of stiff differential systems," *ACM Proc. International Computing Symposium*, North-Holland, Amsterdam, 1974.
6. W. GAUTSCHI, "Numerical integration of ordinary differential equations based on trigonometric polynomials," *Numer. Math.*, v. 3, 1961, pp. 381–397. MR 25 #1647.
7. C. W. GEAR, *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs, N. J., 1971. MR 47 #4447.
8. J. D. LAMBERT, "The numerical integration of a special class of stiff differential systems," *Proc. Fifth Manitoba Conf. on Numerical Mathematics*, Univ. of Manitoba, Oct. 1975.
9. L. ODEN, *An Experimental and Theoretical Analysis of the SAPS Method for Stiff Ordinary Differential Equations*, Report NA 71.28, Dept. of Inform. Proc., Royal Inst. of Tech., Stockholm, 1971.
10. G. W. STEWART, *Simultaneous Iteration for Computing Invariant Subspaces of Non-Hermitian Matrices*, Gatlinburg VI; Springer-Verlag, 1974.
11. O. B. WIDLUND, "A note on unconditionally stable linear multistep methods," *BIT*, v. 7, 1967, pp. 65–70. MR 35 #6373.