

The Application of Linear Multistep Methods to Singular Initial Value Problems*

By Frank R. de Hoog and Richard Weiss**

Abstract. A theory for linear multistep schemes applied to the initial value problem for a nonlinear first order system of differential equations with a singularity of the first kind is developed. Predictor-corrector schemes are also considered.

The specific examples given are systems derived from partial differential equations in the presence of symmetry.

1. Introduction. Ordinary differential equations with singular coefficients are often obtained from partial differential equations when symmetry is present. Although the reduction of these equations to a first order system can be achieved in a variety of ways, the form

$$(1.1a) \quad y' - My/t = f(t, y), \quad 0 \leq t \leq 1,$$

$$(1.1b) \quad b(y(0), y(1)) = 0,$$

where M is a constant matrix, and f , b are smooth, can generally be obtained. This is illustrated in the appendix. When, as is typically the case, M has no eigenvalues which have a positive real part or are purely imaginary, shooting techniques can be applied to (1.1). These require the solution of the initial value problem for (1.1a). Despite the presence of a singularity at $t = 0$, the application of a standard multistep scheme in Keller and Wolfe [5] to an equation similar to (1.1a) has led to satisfactory numerical results. It is therefore natural to ask if multistep schemes can generally be applied. This question is considered here.

It is found that the root condition is no longer sufficient for stability and that an additional condition involving the growth factors of the scheme and the eigenvalues of M must be satisfied. This condition always holds when the growth factors are real and nonnegative. Stability and consistency guarantee convergence in the usual way. Predictor-corrector schemes are also considered, and we find that the stability now depends on both the predictor and corrector. However, all Adams predictor-corrector combinations are still stable.

Received December 29, 1975; revised July 27, 1976.

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

Key words and phrases. Initial value problem, singularity of the first kind, linear multistep method, stability, convergence, predictor-corrector scheme.

* Sponsored by the United States Army under Contract No. DAAG29-75-C-0024 and by the National Science Foundation under Grant No. MPS72-00381 A01.

** This work was performed while one of the authors (R. Weiss) was at the Mathematics Research Center in Madison, Wisconsin.

2. **Preliminaries.** In this section we briefly discuss the initial value problem

$$(2.1) \quad \begin{aligned} y' - \frac{M}{t}y &= f(t, y), \quad 0 \leq t \leq 1, y \in C^1[0, 1], \\ y(0) &= \gamma, \end{aligned}$$

where y and f are n -vectors and M is a constant $n \times n$ matrix. We denote by R a projection onto the invariant subspace of M spanned by the eigenvectors corresponding to the eigenvalue zero and set $Q = I - R$. We assume that

- A2.1. M has no eigenvalues which are purely imaginary or have a positive real part,
- A2.2. the initial vector satisfies $\gamma \in \ker M$, i.e. $Q\gamma = 0$,
- A2.3. $f(t, y)$ is continuous with respect to t and uniformly Lipschitz continuous with respect to y for $0 \leq t \leq 1$ and all y .

A simple modification of the existence proof for the linear initial value problem given in de Hoog and Weiss [2, Section 3] then yields

THEOREM 2.1. *If A2.1, A2.2 and A2.3 are satisfied, (2.1) has a unique continuously differentiable solution $y(t)$. Furthermore, if f is p times continuously differentiable, $y \in C^{p+1}[0, 1]$.*

Although A2.3 is required for the existence of a solution on the entire interval $[0, 1]$, all subsequent results could also be obtained under the weaker assumption that a solution exists, and f is Lipschitz continuous in a neighborhood of this solution. However, as discussed in de Hoog and Weiss [2], A2.1 and A2.2 are necessary for problem (2.1) to be well defined.

3. **Linear Multistep Methods.** Linear multistep schemes for the problem

$$y' = \varphi(t, y), \quad y(0) = \gamma, \quad 0 \leq t \leq 1,$$

have the form

$$(3.1) \quad y_{j+k} = \sum_{\nu=0}^{k-1} \alpha_{\nu} y_{j+\nu} + \sum_{\nu=0}^k \beta_{\nu} \varphi(t_{j+\nu}, y_{j+\nu}), \quad j = 0, \dots, I-k,$$

where $t_j = jh$, $j = 0, \dots, I = 1/h$, y_j denotes the approximation to $y(t_j)$ and y_0, \dots, y_{k-1} are given. In the present case,

$$\varphi(t, y) = My/t + f(t, y);$$

and hence, $\varphi(0, y)$ is not well defined. However, since

$$y'(0) = (I - M)^{-1}f(0, y(0)),$$

we take

$$\varphi(0, y) = (I - M)^{-1}f(0, y).$$

Consistency and stability criteria for the case when no singularity is present are expressed in terms of the polynomials

$$\rho(\zeta) = \zeta^k - \sum_{\nu=0}^{k-1} \alpha_\nu \zeta^\nu \quad \text{and} \quad \sigma(\zeta) = \sum_{\nu=0}^k \beta_\nu \zeta^\nu.$$

In particular the root condition, which states that no roots of $\rho(\zeta)$ have modulus greater than one and that roots of modulus one (subsequently denoted ζ_1, \dots, ζ_r) are simple, is both necessary and sufficient for stability. The consistency conditions are

$$\rho(1) = 0, \quad \rho'(1) = \sigma(1).$$

Although it turns out that the root condition is no longer sufficient for stability in the present case, it is of course necessary.

4. Stability and Convergence. In this section we investigate the stability and convergence of the linear multistep scheme (3.1) for the initial value problem (2.1). Most of the essential questions concerning stability may be answered by considering the scalar problem

$$(4.1) \quad y' - \frac{\lambda}{t}y = g(t).$$

The extension to the general situation will turn out to be rather straightforward.

Therefore, we begin by studying the difference equation

$$(4.2) \quad (1 - \lambda\beta_k/(j+k))v_{j+k} = \sum_{\nu=0}^{k-1} (\alpha_\nu + \lambda\beta_\nu/(j+\nu))v_{j+\nu} + hg_{j+k},$$

$$j = 0, \dots, I-k,$$

where the g_{j+k} , $j = 0, \dots, I-k$ and v_0, \dots, v_{k-1} are given and $v_0/0$ stands for some given value v'_0 . (With the particular choice of $g_{j+k} = \sum_{\nu=0}^k \beta_\nu g(t_{j+\nu})$, (4.2) is precisely the scheme (3.1) applied to (4.1).) As it is clear that in (4.2) v'_0 is propagated in the same way as the v_0, \dots, v_{k-1} , there is no need to distinguish explicitly between v_0 and v'_0 . This will result in a slightly simpler notation.

Clearly, (4.2) has a unique solution provided

$$(4.3) \quad \lambda\beta_k \neq j, \quad j = k, \dots, I.$$

Writing (4.2) as a one step scheme and noting that

$$(1 - \lambda\beta_k/(j+k))^{-1} = 1 + \frac{\lambda\beta_k}{j+1} + O\left(\frac{1}{(j+1)^2}\right),$$

we obtain

$$(4.4) \quad z_{j+1} = S_{j+1}z_j + hr_j, \quad j = 0, \dots, I-k,$$

where

$$z_j = (v_{j+k-1}, \dots, v_j)^T, \quad r_j = ((1 - \lambda\beta_k/(j+k))^{-1}g_{j+k}, 0, \dots, 0)^T,$$

$$S_j = S + \frac{\lambda}{j}T + \frac{1}{j^2}U_j,$$

$$S = \begin{bmatrix} \alpha_{k-1} & \alpha_{k-2} & \cdots & \alpha_1 & \alpha_0 \\ & 1 & 0 & & \\ & & \ddots & & \mathbf{0} \\ & & & 1 & \\ & \mathbf{0} & & & 0 \\ & & & & \ddots \\ & & & & & 1 & 0 \end{bmatrix},$$

$$T = \begin{bmatrix} \beta_{k-1} + \beta_k \alpha_{k-1} & \cdots & \beta_0 + \beta_k \alpha_0 \\ & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ & 0 & \cdots & 0 \end{bmatrix},$$

and

$$\|U_j\| \leq \text{const}, \quad j = 1, \dots, I - k + 1.$$

(Here $\|\cdot\|$ denotes the usual maximum norm.)

The solution of (4.4) is

$$(4.5) \quad z_j = \prod_{l=1}^j S_l z_0 + h \sum_{\nu=0}^{j-1} \prod_{l=\nu+2}^j S_l r_\nu, \quad j = 1, \dots, I - k + 1,$$

where

$$\prod_{l=i}^j S_l = S_j S_{j-1} \cdots S_i.$$

To obtain estimates for (4.5), we require the following two lemmas.

LEMMA 4.1. Let $A_l, B_l, l = 1, 2, \dots$, be square matrices satisfying

$$0 < c_1 (i/j)^\alpha \leq \left\| \prod_{l=i}^j A_l \right\| \leq C_1 (i/j)^\alpha, \quad 1 \leq i \leq j, \quad c_1, C_1 = \text{const},$$

and

$$\|B_l\| \leq C_2, \quad l = 1, 2, \dots, \quad C_2 = \text{const},$$

where α is a real number. Then there exist constants c_3, C_3 and i_0 such that

$$(4.6) \quad 0 < c_3 (i/j)^\alpha \leq \left\| \prod_{l=i}^j \left(A_l + \frac{1}{l^2} B_l \right) \right\| \leq C_3 (i/j)^\alpha, \quad i_0 \leq i \leq j.$$

Proof. On defining

$$Y_{ij} = \prod_{l=i}^j \left(A_l + \frac{1}{l^2} B_l \right), \quad j = i, i + 1, \dots,$$

we obtain the recurrence relation

$$Y_{i,j+1} - A_{j+1} Y_{ij} = \frac{1}{(j+1)^2} B_{j+1} Y_{ij}, \quad j = i, i + 1, \dots$$

Hence,

$$(4.7) \quad Y_{ij} = \prod_{l=i+1}^j A_l Y_{ii} + \sum_{\nu=i}^{j-1} \prod_{l=\nu+2}^j A_l \frac{1}{(\nu+1)^2} B_{\nu+1} Y_{i\nu},$$

$j = i + 1, i + 2, \dots$

Without loss of generality we may assume that $C_1 \geq 1$. We now use induction on j to show that

$$\|Y_{ij}\| \leq 2C_1(i/j)^\alpha, \quad j = i, i + 1, \dots,$$

whenever $i \geq i_1 \geq 3^{|\alpha|} C_2(1 + 2C_1)$. This is easily verified for $j = i$, and if we assume that

$$(4.8) \quad \|Y_{im}\| \leq 2C_1(i/m)^\alpha, \quad m = i, \dots, j - 1,$$

then (4.7) yields

$$\|Y_{ij}\| \leq C_1(i/j)^\alpha + C_1 \left(\frac{i+1}{j} \right)^\alpha \frac{1}{i^2} C_2 + 2C_1^2 C_2 \sum_{\nu=i}^{j-1} \left(\frac{\nu+2}{j} \right)^\alpha \frac{1}{(\nu+1)^2} \left(\frac{i}{\nu} \right)^\alpha.$$

Hence,

$$\|Y_{ij}\| \leq (i/j)^\alpha C_1(1 + 2^{|\alpha|} C_2/i^2 + 2 \cdot 3^{|\alpha|} C_1 C_2/i)$$

and since $i \geq i_1$, we obtain (4.8) for $m = j$. This establishes the upper bound in (4.6). The lower bound is derived in a similar way. \square

LEMMA 4.2. *Let the root condition be satisfied. Then there exist positive constants c_4, C_4 and i_2 such that*

$$c_4(i/j)^\alpha \leq \left\| \prod_{l=i}^j S_l \right\| \leq C_4(i/j)^\alpha, \quad i_2 \leq i \leq j,$$

where

$$(4.9) \quad \alpha = - \max_{1 \leq \mu \leq r} \operatorname{Re} \left\{ \frac{\lambda \sigma(\xi_\mu)}{\xi_\mu \rho'(\xi_\mu)} \right\}.$$

Proof. Clearly, for large l the eigenvalues of S_l are close to the eigenvalues of S (the roots of $\rho(\xi)$). In particular, since the root condition is satisfied, it follows from Henrici [4, p. 237] that the eigenvalues of S_l corresponding to ξ_1, \dots, ξ_r (the eigenvalues of modulus one) satisfy

$$(4.10) \quad \xi_{\mu l} = \xi_\mu(1 + \lambda \tau_\mu/l) + O(1/l^2), \quad \mu = 1, \dots, r; l = 1, 2, \dots,$$

where $\tau_\mu = \sigma(\xi_\mu)/(\xi_\mu \rho'(\xi_\mu))$ are the growth factors of the scheme. Let $S = EDE^{-1}$

where D is the Jordan form of S , with

$$D = \begin{bmatrix} D_{1,0} & \mathbf{0} \\ \mathbf{0} & D_{2,0} \end{bmatrix}, \quad D_{1,0} = \begin{bmatrix} \xi_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \xi_r \end{bmatrix}.$$

A standard perturbation argument now yields

$$S_l = \left(E + \frac{1}{l} F + \frac{1}{l^2} G_l \right) D_l \left(E + \frac{1}{l} F + \frac{1}{l^2} G_l \right)^{-1}$$

for sufficiently large l , say $l \geq l_0 \geq 1$, where F is an $n \times n$ matrix,

$$D_l = \begin{bmatrix} D_{1l} & \mathbf{0} \\ \mathbf{0} & D_{2l} \end{bmatrix}, \quad D_{1l} = \begin{bmatrix} \xi_{1l} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \xi_{rl} \end{bmatrix},$$

$$(4.11) \quad \|D_{2,0} - D_{2l}\| \leq \text{const}(1/l)$$

and $\|G_l\| \leq \text{const}$. Clearly,

$$(4.12) \quad S_l = \left(E + \frac{1}{l+1} F \right) D_l \left(E + \frac{1}{l} F \right)^{-1} + \frac{1}{l^2} B_l =: A_l + \frac{1}{l^2} B_l,$$

where $\|B_l\| \leq \text{const}$. Since

$$\prod_{i=1}^j A_i = \left(E + \frac{1}{j+1} F \right) \left(\prod_{i=1}^j D_i \right) \left(E + \frac{1}{i} F \right)^{-1},$$

it follows that

$$\begin{aligned} c_5 \max \left\{ \left\| \prod_{i=1}^j D_{1i} \right\|, \left\| \prod_{i=1}^j D_{2i} \right\| \right\} \\ \leq \left\| \prod_{i=1}^j A_i \right\| \leq C_5 \max \left\{ \left\| \prod_{i=1}^j D_{1i} \right\|, \left\| \prod_{i=1}^j D_{2i} \right\| \right\}, \end{aligned}$$

where c_5, C_5 are positive constants. Now the D_{1l} are diagonal matrices and so

$$\left\| \prod_{i=1}^j D_{1i} \right\| = \max_{1 \leq \mu \leq r} \prod_{i=1}^j |\xi_{\mu i}|.$$

Hence, from (4.10) and (4.9),

$$0 < c_6 \prod_{i=1}^j (1 - \alpha/l) \leq \left\| \prod_{i=1}^j D_{1i} \right\| \leq C_6 \prod_{i=1}^j (1 - \alpha/l), \quad c_6, C_6 = \text{const},$$

when i is sufficiently large (say $i \geq i_3 \geq l_0$). Also,

$$\prod_{i=1}^j (1 - \alpha/l) = \frac{\Gamma(j+1-\alpha)\Gamma(i)}{\Gamma(i-\alpha)\Gamma(j+1)},$$

and the asymptotic results for the Gamma function (see, for instance, Luke [6, p. 33]) therefore yield

$$(4.13) \quad c_7(i/j)^\alpha \leq \left\| \prod_{l=i}^j D_{1l} \right\| \leq C_7(i/j)^\alpha, \quad i \geq i_4,$$

where c_7, C_7 are positive constants and i_4 is sufficiently large. As the eigenvalues of $D_{2,0}$ are smaller than one in modulus, there exists an induced norm, subsequently denoted by $||| \cdot |||$, such that $|||D_{2,0}||| = a < 1$. It therefore follows from (4.11) that for large l , say $l \geq l_1$,

$$|||D_{2l}||| \leq \tilde{a} < 1,$$

which yields

$$\left\| \prod_{l=i}^j D_{2l} \right\| \leq \text{const}(\tilde{a})^{j+1-i}, \quad i \geq l_1.$$

Since this term decreases exponentially, it is clear from (4.13) that for sufficiently large i ,

$$c_5 \left\| \prod_{l=i}^j D_{1l} \right\| \leq \left\| \prod_{l=i}^j A_l \right\| \leq C_5 \left\| \prod_{l=i}^j D_{1l} \right\|.$$

The result now follows from (4.13), (4.12) and Lemma 4.1. \square

From (4.5) and Lemma 4.2 we see that when $\alpha < 0$, the contribution from the starting values will grow algebraically with j ; and thus, such a scheme is not stable. To illustrate this, we apply the midpoint scheme ($k = 2, \alpha_1 = \beta_0 = \beta_2 = 0, \alpha_0 = 1, \beta_1 = 2; \xi_1 = \tau_1 = 1, \xi_2 = \tau_2 = -1$) to the problem $y' - \lambda y/t = 0$. It is easily verified that the resulting difference equation

$$y_{j+2} - y_j = \frac{2\lambda}{j+1} y_{j+1}, \quad y_0 = 0, \quad y_1 = \delta,$$

has the solution

$$y_j = j^{-\lambda}(-1)^{j+1}\delta, \quad j = 1, 2, \dots,$$

if $\lambda = -1$ or -2 .

While it is possible to develop a complete theory which also covers the case when $\alpha < 0$, the obvious shortcomings of such schemes make them of academic interest only. We therefore assume in the sequel that $\alpha \geq 0$.

We now return to (4.2) and consider the case when

$$(4.14) \quad g_{j+k} = c_{j+k} + d_{j+k}/t_{j+k}.$$

The second term will be required when analyzing the propagation of the starting vectors in systems and the error in predictor-corrector algorithms.

LEMMA 4.3. *Let $\alpha \geq 0$ and assume that (4.3) and the root condition are satisfied. Then the solution of (4.2) with g_{j+k} given by (4.14) satisfies*

$$|v_j| \leq \text{const} \left\{ j^{-\alpha} \max_{0 \leq \nu \leq k-1} |v_\nu| + t_j \max_{k \leq l \leq j} |c_l| + (\ln(j+1))^\eta \max_{k \leq l \leq j} |d_l| \right\},$$

$$j = k, \dots, I,$$

where $\eta = 1$ when $\alpha = 0$ and $\eta = 0$ otherwise.

LEMMA 4.4. *Let (4.3) and the root condition be satisfied. Then (4.16) has a unique solution for all starting vectors $v_\nu, \nu = 0, \dots, k - 1$. Furthermore, if*

(i) $\lambda = 0,$

$$\|v_j\| \leq \text{const} \{ V + t_j C_j + (\ln(j + 1))^{n-1} (V_1 + D_j) \}, \quad j = k, \dots, I;$$

(ii) $\alpha = 0, \text{Re } \lambda < 0,$

$$\|v_j\| \leq \text{const} \{ V + t_j C_j + (\ln(j + 1))^{n-1} V_1 + (\ln(j + 1))^n D_j \}, \quad j = k, \dots, I;$$

(iii) $\alpha > 0, \text{Re } \lambda < 0,$

$$\|v_j\| \leq \text{const} \{ V + t_j C_j + D_j \}, \quad j = k, \dots, I.$$

Proof. If $n = 1$, we have the scalar equation considered in Lemma 4.3. We now proceed inductively. Suppose the result holds for $n = 1, \dots, p - 1$, and consider the case when $n = p$. If we use the notation $v_j = (v_{j1}, \dots, v_{jp})^T$, it is clear that $\tilde{v}_j := (v_{j2}, \dots, v_{jp})^T$ satisfies an equation of the form (4.16) with $n = p - 1$. Furthermore, v_{j1} is given by

$$\begin{aligned} v_{j+k,1} - \sum_{\nu=0}^{k-1} \alpha_\nu v_{j+\nu,1} - \lambda \sum_{\nu=0}^k \beta_\nu v_{j+\nu,1} / (j + \nu) \\ = hc_{j+k,1} + \left\{ \lambda d_{j+k,1} + d_{j+k,2} + \sum_{\nu=0}^k \beta_\nu v_{j+\nu,2} (j+k) / (j + \nu) \right\} / (j+k), \end{aligned}$$

$$j = 0, \dots, I - k.$$

Using the estimate for $\|\tilde{v}_j\|$ to obtain a bound for the right-hand side of this equation and applying Lemma 4.3 yields the result for $n = p$. \square

The estimates in Lemma 4.4 can be shown to be sharp in the sense that all the logarithmic terms shown do occur. Note that when $\text{Re } \lambda < 0$ and $\alpha = 0$, the initial vectors grow logarithmically while this is not the case when $\alpha > 0$. Since schemes with $\alpha > 0$ can be realized, we subsequently restrict ourselves to this case.

We now consider the difference equation

$$(4.17) \quad v_{j+k} - \sum_{\nu=0}^{k-1} \alpha_\nu v_{j+\nu} - \sum_{\nu=0}^k \beta_\nu M v_{j+\nu} / (j + \nu) = h(c_{j+k} + M d_{j+k} / t_{j+k}),$$

$$j = 0, \dots, I - k.$$

Let $E = [e_1, \dots, e_n]$ be a basis in which M reduces to Jordan form, i.e. $M = EDE^{-1}$, where

$$D = \begin{bmatrix} J_1 & & & \\ & \ddots & & \\ & & \ddots & \\ & & & J_\chi \end{bmatrix};$$

and the $J_l, l = 1, \dots, \chi$, have the same structure as (4.15). We denote by X the related projection onto the span of the generalized eigenvectors corresponding to the eigenvalue

zero (not including the eigenvectors themselves) and set $Y = I - X$.

Multiplying (4.17) by E^{-1} yields

$$u_{j+k} - \sum_{\nu=0}^{k-1} \alpha_{\nu} u_{j+\nu} - \sum_{\nu=0}^k \beta_{\nu} D u_{j+\nu} / (j + \nu) = h(E^{-1} c_{j+k} + DE^{-1} d_{j+k} / t_{j+k}),$$

$$j = 0, \dots, I - k,$$

where $u_j = E^{-1} v_j$. But this equation now consists of χ separate equations of the form (4.16), and we can therefore apply Lemma 4.4 to obtain

LEMMA 4.5. *Let (4.3) be satisfied for all eigenvalues λ of M . Then (4.17) has a unique solution for all starting vectors v_0, \dots, v_{k-1} . Furthermore, if the root condition is satisfied and α defined by (4.9) is positive for all nonzero eigenvalues of M , then*

$$\|v_j\| \leq \text{const} \left\{ \max_{0 \leq \nu \leq k-1} \|Y v_{\nu}\| + (\ln(j+1))^{m-1} \max_{0 \leq \nu \leq k-1} \|X v_{\nu}\| \right. \\ \left. + t_j \max_{k \leq l \leq j} \|c_l\| + \max_{k \leq l \leq j} \|Y d_l\| + (\ln(j+1))^{m-1} \max_{k \leq l \leq j} \|X d_l\| \right\},$$

$$j = k, \dots, I,$$

where m is the dimension of the largest Jordan box in D associated with a zero eigenvalue.

For differential equations without a singularity ($M = 0$), the properties of the difference equation

$$v_{j+k} - \sum_{\nu=0}^{k-1} \alpha_{\nu} v_{j+\nu} = h g_{j+k}$$

completely determine the stability of the entire scheme. We have seen that this is no longer true in the present case. Now it is (4.17) which characterizes the behavior of the difference equation

$$v_{j+k} - \sum_{\nu=0}^{k-1} \alpha_{\nu} v_{j+\nu} - \sum_{\nu=0}^k \beta_{\nu} M v_{j+\nu} / (j + \nu) \\ = h \sum_{\nu=0}^k \beta_{\nu} Z_{j+\nu} v_{j+\nu} + h c_{j+k} + M d_{j+k} / (j + k), \quad j = 0, \dots, I - k.$$

THEOREM 4.1. *Let the hypotheses of Lemma 4.5 be satisfied and*

$$\|Z_j\| \leq \text{const}, \quad j = 0, \dots, I.$$

Then there exists a constant $h_0 > 0$ such that (4.19) has a unique solution for all starting vectors $v_{\nu}, \nu = 0, \dots, k - 1$, whenever $h \leq h_0$. Furthermore, this solution satisfies the estimate (4.18).

Proof. Consider the right-hand side of (4.19) as the inhomogeneous term in (4.17), and apply Lemma 4.5. When t_j is sufficiently small, the Banach Lemma yields the result. The standard translation argument is then used to obtain the estimate on the whole interval. \square

We are now ready to discuss the convergence of linear multistep schemes. A

scheme is called consistent of order q if q is the largest integer such that

$$w(t_{j+k}) - \sum_{\nu=0}^{k-1} \alpha_{\nu} w(t_{j+\nu}) - h \sum_{\nu=0}^k \beta_{\nu} w'(t_{j+\nu}) = O(h^{q+1}), \quad j = 0, \dots, I - k,$$

for all sufficiently smooth functions $w(t)$.

THEOREM 4.2. *Assume that*

- (i) (4.3) is satisfied for all eigenvalues λ of M ,
- (ii) the root condition is satisfied,
- (iii) α (defined by (4.9)) is positive for all nonzero eigenvalues of M ,
- (iv) the starting vectors satisfy,

$$y_{\nu} = y(t_{\nu}) + O(h^s), \quad \nu = 0, \dots, k - 1,$$

- (v) the scheme is consistent of order q ,
- (vi) f is q times continuously differentiable.

Then, there exists a constant $h_0 > 0$ such that the scheme (3.1) applied to (2.1) has a unique solution for all $h \leq h_0$, which satisfies

$$\|y(t_j) - y_j\| \leq \text{const} \{h^s (\ln(j + 1))^{\eta} + h^q\}, \quad j = k, \dots, I,$$

where $\eta = m - 1$ if $m \geq 1$ and 0 otherwise.

Proof. Existence and uniqueness of a solution follow from a contraction mapping argument similar to that used in Theorem 4.1.

From Theorem 2.1, $y(t)$ is $q + 1$ times continuously differentiable. The local truncation error is therefore $O(h^{q+1})$. (Note that the definition of consistency does not depend on the particular differential equation to be solved.) Since f is globally Lipschitz continuous, the stability demonstrated in Theorem 4.1 (where we take $d_{j+k} = 0$) now yields the result in the usual fashion. \square

5. Predictor-Corrector Algorithms. For implicit methods ($\beta_k \neq 0$) a nonlinear system of equations for the unknown y_{j+k} must be solved at each step. A convenient way to do this is to predict a vector using an explicit scheme,

$$(5.1) \quad y_{j+k}^{(0)} = \sum_{\nu=0}^{k-1} \{(\alpha_{\nu}^* + \beta_{\nu}^* M / (j + \nu)) y_{j+\nu} + h \beta_{\nu}^* f(t_{j+\nu}, y_{j+\nu})\};$$

and then apply an iteration procedure to solve

$$(5.2) \quad y_{j+k} - \sum_{\nu=0}^{k-1} \alpha_{\nu} y_{j+\nu} = \sum_{\nu=0}^k \beta_{\nu} (M y_{j+\nu} / (j + \nu) + h f(t_{j+\nu}, y_{j+\nu})),$$

with the predicted $y_{j+k}^{(0)}$ as the starting iterate. The schemes (5.1) and (5.2) are called the predictor and corrector, respectively.

An obvious iteration scheme is

$$(5.3) \quad (I - \beta_k M / (j + k)) y_{j+k}^{(\kappa+1)} = \sum_{\nu=0}^{k-1} \{(\alpha_{\nu} + \beta_{\nu} M / (j + \nu)) y_{j+\nu} + h \beta_{\nu} f(t_{j+\nu}, y_{j+\nu})\} + h \beta_k f(t_{j+k}, y_{j+k}^{(\kappa)}), \quad \kappa = 0, 1, \dots$$

It is clear that for this procedure a predictor-corrector theory can be developed which

in all aspects parallels that for differential equations without a singularity. This scheme however, has the disadvantage that each iteration requires the solution of a linear system of equations.

If we ignore the presence of the singularity and proceed as in the regular case, we obtain the usual predictor-corrector algorithms. This is now examined in detail when correcting only once. The resulting scheme is

$$\begin{aligned}
 (5.4) \quad y_{j+k} - \sum_{\nu=0}^{k-1} \alpha_{\nu} y_{j+\nu} &= \sum_{\nu=0}^{k-1} M(\beta_{\nu}/(j+\nu) + \beta_k \alpha_{\nu}^*/(j+k)) y_{j+\nu} \\
 &+ \beta_k M \sum_{\nu=0}^{k-1} \beta_{\nu}^* M y_{j+\nu} / (j+k)(j+\nu) \\
 &+ h \sum_{\nu=0}^{k-1} (\beta_{\nu} + \beta_k \beta_{\nu}^* M / (j+k)) f(t_{j+\nu}, y_{j+\nu}) \\
 &+ h \beta_k f(t_{j+k}, y_{j+k}^{(0)}), \quad j = 0, \dots, I-k.
 \end{aligned}$$

From Lemma 4.1 it follows that Theorem 4.1 remains valid when a term of the form

$$(5.5) \quad \sum_{\nu=0}^{k-1} G_{\nu} y_{j+\nu} / (j+k)(j+\nu)$$

with $\|G_{\nu}\| \leq \text{const}$, $\nu = 0, \dots, k-1$, is added to the left-hand side of (4.19). Hence, the stability condition for (5.4) is

$$(5.6) \quad \text{Re} \frac{\lambda \tilde{\sigma}(\xi_{\mu})}{\xi_{\mu} \rho'(\xi_{\mu})} < 0, \quad \mu = 1, \dots, r,$$

for all nonzero eigenvalues λ of M , where

$$\tilde{\sigma}(\xi) = \sigma(\xi) - \beta_k \rho^*(\xi).$$

The following convergence result can be established.

THEOREM 5.1. *Assume that*

- (i) $\rho(\xi)$ satisfies the root condition,
- (ii) (5.6) holds,
- (iii) the starting vectors satisfy

$$\|y_{\nu} - y(t_{\nu})\| \leq \text{const } h^s, \quad \nu = 0, \dots, k-1,$$

(iv) the schemes (5.1) and (5.2) are consistent of order p and q respectively, with $p \leq q$,

(v) $f(t, y)$ is u times continuously differentiable, where $u = \min\{p+1, q\}$.

Then the solution of (5.4) satisfies

$$\|y_j - y(t_j)\| \leq \text{const} \{ (h^s + h^{p+1})(\ln(j+1))^{\eta} + h^q \}, \quad j = k, \dots, I,$$

where $\eta = m - 1$ if $m \geq 1$ and 0 otherwise.

Proof. The residual obtained when substituting the exact solution into (5.4) has the form

$$hr_{j+k} = h \left(r_{1,j+k} + \frac{M}{t_{j+k}} r_{2,j+k} \right),$$

where

$$r_{1,j+k} = O(h^{p+1} + h^q), \quad r_{2,j+k} = O(h^{p+1}).$$

Theorem 4.1 (modified by the inclusion of a term of the form (5.5)) now yields the result. \square

A theory for correcting more than once can of course be established but, due to the singularity, further corrector iterations do not generally lead to improved convergence.

6. Remarks. It turns out that for most problems encountered in practice the eigenvalues of M are simple. This means that the logarithmic terms in the error estimates do not occur. In addition, the eigenvalues are usually real (typically 0 and -1 or 0 and -2 for problems obtained from partial differential equations when angular or spherical symmetry is present); and hence, the stability criterion

$$(6.1) \quad \operatorname{Re} \{ \lambda \sigma(\xi_\mu) / (\xi_\mu \rho'(\xi_\mu)) \} < 0, \quad \mu = 1, \dots, r, \operatorname{Re} \lambda < 0,$$

is satisfied if the growth factors

$$\tau_\mu = \sigma(\xi_\mu) / (\xi_\mu \rho'(\xi_\mu)), \quad \mu = 1, \dots, r,$$

have positive real parts. However, (6.1) always holds when the growth factors are real and positive. In particular, this is true for all schemes which have only one root of modulus one ($r = 1$), and are consistent, i.e. $\tau_1 = \sigma(1) / \rho'(1) = 1$.

It is clear that the solvability condition, i.e. the requirement that (4.3) holds for all eigenvalues λ of M , is always satisfied when $\beta_k \geq 0$.

For predictor-corrector pairs stability depends on both the predictor and corrector. Of course, if $\rho(\xi)$ has only the principal root, the combination is always stable. However, it is a curious fact that a stable predictor and corrector may lead to an unstable combination and vice versa.

One way to calculate starting values is via the implicit Runge-Kutta schemes considered in de Hoog and Weiss [3]. The resulting nonlinear equations may be solved by Newton's method or an iteration analogous to (5.3). Explicit Runge-Kutta schemes do not provide high order approximations in the general case although they can often be successfully applied to problems of practical interest. This will be reported in a subsequent paper. Of course, when $f(t, y)$ has a simple form, a truncated Taylor series expansion may also provide an effective starting procedure.

7. Numerical Results. In this section, we illustrate our theory by applying a predictor-corrector scheme to the problem

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \frac{1}{t} \begin{pmatrix} 0 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} y_2 \\ -y_1^5 \end{pmatrix}, \quad 0 \leq t \leq 1, y_1(0) = 1, y_2(0) = 0.$$

The solution is

$$y_1(t) = 1/\sqrt{1 + t^2/3}, \quad y_2(t) = y_1'(t).$$

As predictor and corrector, we choose the fourth order Adams-Bashforth and Adams-Moulton method, respectively (cf. Henrici [4, p. 194 and p. 199]).

The starting vectors for $t_j = jh, j = 1, 2, 3$, were obtained by an implicit Runge-Kutta (or collocation) scheme. To describe this procedure, the notation of de Hoog and Weiss [3] will now be used. We choose $m = 3$ and u_1, u_2, u_3 as the Radau points. The resulting nonlinear equations (see de Hoog and Weiss [3, Eq. (3.2)]) were solved by an iteration analogous to (5.3), with $y_{j,0}$ as the initial iterate for $y_{jk}, k = 1, 2, 3$. Five iterations were used for each j . From Theorem 5.3 in de Hoog and Weiss [3], the starting vectors, i.e. $y_{j3}, j = 0, 1, 2$, are at least fourth order accurate.

Numerical results for different values of t and h are given in Tables 1 and 2 and confirm the fourth order convergence predicted by Theorem 5.1.

The computations were done in double precision arithmetic on the UNIVAC 1110 of the University of Wisconsin-Madison.

TABLE 1
 $y_1(t) - y_{1,t/h}$

	$h = 0.1$	$h = 0.05$	$h = 0.025$	$h = 0.0125$
$t = 0.2$	-1.4852 E-8	-7.2106 E-9	9.5452 E-10	1.0861 E-10
0.4	-3.0586 E-7	6.1558 E-8	6.2040 E-9	4.2439 E-10
0.6	1.3630 E-6	1.7350 E-7	1.1713 E-8	7.1299 E-10
0.8	3.5068 E-6	2.4856 E-7	1.4131 E-8	8.0114 E-10
1.0	4.7504 E-6	2.5205 E-7	1.2568 E-8	6.6377 E-10

TABLE 2
 $y_2(t) - y_{2,t/h}$

	$h = 0.1$	$h = 0.05$	$h = 0.025$	$h = 0.0125$
$t = 0.2$	1.9629 E-8	2.2337 E-7	1.4183 E-8	6.0377 E-10
0.4	5.7639 E-6	3.2171 E-7	1.0908 E-8	3.7556 E-10
0.6	5.3537 E-6	1.0288 E-7	-1.7377 E-9	-3.1119 E-10
0.8	4.0519 E-7	1.8407 E-7	-1.5768 E-8	-1.0471 E-9
1.0	-4.7107 E-6	-3.9426 E-7	-2.4754 E-8	-1.4855 E-9

Appendix. Keller and Wolfe [5] consider the equations

$$L\alpha(t) = \rho[\alpha(t)\gamma(t) + Pt^2], \quad L\gamma(t) = \rho[t^2 - \alpha^2(t)]$$

subject to

$$\alpha(0) = 0, \quad \gamma(0) = 0, \quad \alpha(1) = 1, \quad \frac{d\gamma}{dt}(1) = \nu\gamma(1),$$

where L is the differential operator

$$L \cdot = t \frac{d}{dt} \left[\frac{1}{t} \frac{d}{dt} (t \cdot) \right],$$

and ρ, P, ν are constants. This problem arises in the study of the buckling mechanism of the cap portion of a spherical shell. We illustrate here how these equations may be transformed to a system of the form (1.1).

Let $y_1(t) = \alpha(t)/t$, $y_2(t) = \alpha'(t)$, $y_3(t) = \gamma(t)/t$ and $y_4(t) = \gamma'(t)$. Then we obtain

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}' = \frac{1}{t} \begin{bmatrix} -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} + \begin{bmatrix} 0 \\ t\rho(y_1 y_3 + P) \\ 0 \\ t\rho(1 - y_1^2) \end{bmatrix},$$

which has the required form. Note that the condition $Qy(0) = 0$ required in A2.2 is

$$y_1(0) = y_2(0), \quad y_3(0) = y_4(0),$$

which is consistent with

$$\alpha(0) = 0, \quad \lim_{t \rightarrow 0} \frac{\alpha(t)}{t} = \alpha'(0), \quad \gamma(0) = 0, \quad \lim_{t \rightarrow 0} \frac{\gamma(t)}{t} = \gamma'(0).$$

Similar transformations can be applied to other problems; such as the shell equations derived in Bauer, Reiss and Keller [1] and the Ginzburg-Landau equations discussed in Meissner and Tholfsen [7].

Computer Centre
The Australian National University
P.O. Box 4
Canberra, A.C.T. 2600, Australia

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

1. L. BAUER, E. L. REISS & H. B. KELLER, "Axisymmetric buckling of hollow spheres and hemispheres," *Comm. Pure Appl. Math.*, v. 23, 1970, pp. 529-568. MR 43 #4335.
2. F. R. de HOOG & R. WEISS, "Difference methods for boundary value problems with a singularity of the first kind," *SIAM J. Numer. Anal.*, v. 13, 1976, pp. 775-813.
3. F. R. de HOOG & R. WEISS, *Collocation Methods for Singular Boundary Value Problems*, MRC Technical Summary Report #1547, 1975.
4. P. HENRICI, *Discrete Variable Methods in Ordinary Differential Equations*, Wiley, New York, 1962. MR 24 #B1772.
5. H. B. KELLER & A. W. WOLFE, "On the nonunique equilibrium states and buckling mechanism of spherical shells," *J. Soc. Indust. Appl. Math.*, v. 13, 1965, pp. 674-705. MR 32 #656.
6. Y. L. LUKE, *The Special Functions and Their Approximations*, Vol. I, Academic Press, New York, 1969. MR 39 #3039.
7. H. MEISSNER & P. THOLFSEN, "Cylindrically symmetric solutions of the Ginzburg-Landau equations," *Phys. Rev.*, v. 169, 1968, pp. 413-416.