

Equivalent Forms of Multistep Formulas

By Robert D. Skeel*

Abstract. For uniform meshes it is shown that any linear k -step formula can be formulated so that only k values need to be saved between steps. By saving additional m values it is possible to construct local polynomial approximations of degree $k + m - 1$, which can be used as predictor formulas. Different polynomial bases lead to different equivalent forms of multistep formulas. In particular, local monomial bases yield Nordsieck formulas. An explicit one-to-one correspondence is established between Nordsieck formulas and k -step formulas of order at least k , and a strong equivalence result is proved for all but certain pathological cases. Equivalence is also shown for P(EC)* formulas but not for P(EC)*E formulas.

1. Introduction. Multistep methods are the oldest and the most important class of numerical methods for solving systems of ordinary differential equations. Implementations of these methods have become increasingly sophisticated in the last two decades. One paper having a considerable impact on the development of these algorithms is that of Nordsieck [1962], which introduces a class of formulas closely related to multistep formulas. It is the purpose of the present paper to explore thoroughly the relationship between multistep and Nordsieck formulas on a uniform mesh. The results apply to a limited extent to variable-order variable-stepsize algorithms which discourage order and stepsize changes, because such algorithms often produce sequences of values computed with the same formula and stepsize.

Consider the system $y' = f(t, y)$ with exact solution $y(t)$. Given starting values $y_j, y'_j, j = 0(1)k - 1$, a linear k -step formula on a mesh $t_n := t_0 + nh, n = 1(1)N$, determines approximations $\{y_n\}$ to the values $\{y(t_n)\}$ by means of $\rho(E)y_{n-k} = h\sigma(E)y'_{n-k}$ and $y'_n = f(t_n, y_n)$ where

$$\rho(\xi) := \sum_{j=0}^k \alpha_j \xi^{k-j}, \quad \sigma(\xi) := \sum_{j=0}^k \beta_j \xi^{k-j},$$

and E is the shift operator. It is assumed that $\alpha_0 \neq 0$ and $\alpha_0^2 + \beta_0^2 > 0$. For convenience a normalization such as $\alpha_0 = 1$ or $\sum \beta = 1$ is not imposed. If $\beta_0 \neq 0$, the formula is implicit and requires the solution of a system of nonlinear equations, which is always possible for small enough h if $f(t, y)$ is Lipschitz continuous as a function of y . In Section 2 a simple reformulation of the general k -step formula is given which

Received September 1, 1978.

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

Key words and phrases. Linear multistep formula, multistep formula, linear multistep method, multistep method, Nordsieck method, multivalued method, predictor-corrector method.

* Supported in part by the U. S. Air Force Office of Scientific Research under grant AFOSR 75-2854.

requires that only k values be saved between steps. In practice, $k + m$ values might be stored, from which it is possible to construct a predictor formula of order $k + m - 1$ except for pathological cases where $\rho(\xi)$ and $\sigma(\xi)$ have a common factor.

In Section 3 linear multistep formulas are described in terms of a local polynomial approximation $p_n(t)$ which interpolates the $k + m$ values which are retained after completion of the n th step. It is from this basis-free description that equivalent forms of multistep formulas are derived. One convenient representation of the polynomial $p_n(t)$ is in terms of the scaled derivatives $h^j p_n^{(j)}(t_n)/j!$, $j = 0(1)k + m - 1$. Multistep formulas represented in this way belong to the class of formulas introduced by Nordsieck. Another is in terms of ordinates $p_n(t_{n-j})$, $j = 0(1)k_1 - 1$, and derivatives $h p_n'(t_{n-j})$, $j = 0(1)k_2 - 1$, where $k_1 + k_2 = k + m$, which yields the modified multistep formulas of Gear [1971a, p. 150]. Backward differences of ordinates or derivatives are also possible.

General linear Nordsieck formulas compute approximations $\mathbf{a}_n =: [y_n, h y_n', \dots, h^q y_n^{(q)}/q!]^T$ to the vector of scaled derivatives

$$[y(t_n), h y'(t_n), \dots, h^q y^{(q)}(t_n)/q!]^T$$

by the recurrence

$$\mathbf{a}_n := P \mathbf{a}_{n-1} + \mathbf{l} \delta_n,$$

where δ_n is chosen so that $y_n' = f(t_n, y_n)$; that is, δ_n is implicitly defined by

$$\mathbf{e}_1^T P \mathbf{a}_{n-1} + l_1 \delta_n = h f(t_n, \mathbf{e}_0^T P \mathbf{a}_{n-1} + l_0 \delta_n).$$

Here P is the Pascal triangle matrix defined by

$$P_{ij} := \binom{j}{i}, \quad 0 \leq i \leq j \leq q,$$

\mathbf{e}_j^T is the j th unit vector, and $\mathbf{l} =: [l_0, l_1, \dots, l_q]^T$ is a vector of parameters which characterize the formula. It is assumed that $l_1 \neq 0$. Note that zero-origin indexing is being used. Nordsieck [1962] derives particular choices for \mathbf{l} which optimize stability and accuracy for $q = 2, 3, 4, 5, 6$, and he implements the formula for $q = 5$ as a variable-step computer program for the Illiac I. By direct calculation he shows that there is a connection between this formula and the sixth-order Adams-Moulton formula; specifically, it is shown that the zeroth and first component of $\mathbf{a}_n, \mathbf{a}_{n-1}, \dots, \mathbf{a}_{n-5}$ satisfy the difference equation of the Adams-Moulton formula.

Further understanding of Nordsieck formulas resulted from the work of Descloux [1963]. An actual equivalence is shown between Nordsieck's $(q + 1)$ -value formula and the $(q + 1)$ st order Adams-Moulton formula, in the sense that a linear transformation was constructed relating the Nordsieck vector to the vector values $[y_n, h y_n', \dots, h y_{n-q+1}']^T$. An equivalence was also shown for the q th-order Adams-Moulton formula and a $(q + 1)$ -value Nordsieck formula with a slightly different \mathbf{l} vector. This second family of formulas is used in the nonstiff option of DIFSUB (Gear [1971b]) and its descendants. Furthermore, Descloux showed that in both cases the predicted value $y_{n,0} := \mathbf{e}_0^T P \mathbf{a}_{n-1}$ satisfies the difference equation of the k th-order Adams-

Bashforth formula. General k -step formulas were also considered and equivalent Nordsieck formulas were constructed having the more general form

$$(1.1) \quad \mathbf{a}_n = A\mathbf{a}_{n-1} + l\delta_n,$$

where the dimensionality of the arrays might be as high as $2k$ and where A is not necessarily the Pascal triangle matrix. In the appendix of Descloux's report is a theorem stating that the values y_n and f_n computed by formulas even more general than (1.1) satisfy the difference equation of an associated linear multistep formula. The proof is not given by Descloux because it "is rather painful and long." However, in a later paper of Osborne [1966] a similar result is stated and a short and elegant proof is given.

In Section 4 a one-to-one correspondence is established between (i) the class of all $(q + 1)$ -value linear Nordsieck formulas and (ii) the class of all linear multistep formulas of formal order at least q and stepnumber at most q . This correspondence is given in closed form by

$$\xi^{m-1}\rho(\xi) = \det(\xi I - P)\mathbf{e}_1^T(\xi I - P)^{-1}\mathbf{1}$$

and

$$\xi^{m-1}\sigma(\xi) = \det(\xi I - P)\mathbf{e}_0^T(\xi I - P)^{-1}\mathbf{1}.$$

Osborne [1966] gives an expression for $\xi^{m-1}\rho(\xi)/(\xi - 1)$ similar to that given above and proves a correspondence between Nordsieck and multistep formulas of order at least $q + 1$ (although it is not clear that l_0 is correctly chosen in the proof). Presumably at that time only formulas of optimal order were considered to be of practical value. More recently Wallace and Gupta [1973] obtained expressions for $\xi^{m-1}\rho(\xi)$ and $\xi^{m-1}\sigma(\xi)$ which are equivalent to those given above.

In Section 5 it is shown that the zeroth and first components of $\mathbf{a}_n, \mathbf{a}_{n-1}, \dots, \mathbf{a}_{n-k}$ satisfy the difference equation of the corresponding multistep formulas. Also, if $\rho(\xi)$ and $\sigma(\xi)$ have no common factor, then the Nordsieck formula is shown to be equivalent in the sense of Gear [1971a, p. 143] to the corresponding multistep formula, which means that \mathbf{a}_n is a linear transformation of certain linear combinations of the values $y_{n-j}, y'_{n-j}, j = 0(1)k - 1$. It is proved that there is no equivalence to a multistep formula if $\rho(\xi)$ and $\sigma(\xi)$ have a common factor.

In Section 6 predictor-corrector Nordsieck formulas are considered. An equivalence to predictor-corrector multistep formulas is shown to be the case for P(EC)* formulas but not for P(EC)*E formulas.

In Section 7 recent applications of equivalence results are discussed.

2. Minimum Storage for Multistep Formulas. The computation of each solution value y_n and its derivative y'_n by a general linear multistep formula would seem to require saving $2k$ previously computed values: $y_{n-1}, y_{n-2}, \dots, y_{n-k}$ and $y'_{n-1}, y'_{n-2}, \dots, y'_{n-k}$. For this reason it has been considered desirable to have formulas with most of either the α or β coefficients set to zero (see, for example, the "minimum

storage methods" of Dill and Gear [1971]). However, there is an easy way to get by with just k values without reevaluating the function f . Define

$$s_n^j := \sum_{i=0}^j (-\alpha_{k-j+1} y_{n-i} + h\beta_{k-j+1} y'_{n-i})$$

for $j = 0(1)k - 1$. In terms of the partial sums s_n^j , the linear multistep formula becomes

$$\begin{aligned} \text{solve } \alpha_0 y_n &= h\beta_0 f(t_n, y_n) + s_{n-1}^{k-1} \quad \text{for } y_n, \\ \text{set } y'_n &= f(t_n, y_n), \\ s_n^{k-1} &= -\alpha_1 y_n + h\beta_1 y'_n + s_{n-1}^{k-2}, \\ &\vdots \\ &\vdots \\ s_n^1 &= -\alpha_{k-1} y_n + h\beta_{k-1} y'_n + s_{n-1}^0, \\ s_n^0 &= -\alpha_k y_n + h\beta_k y'_n. \end{aligned} \tag{2.1}$$

Thus, only the k partial sums $s_{n-1}^0, s_{n-1}^1, \dots, s_{n-1}^{k-1}$ need to be saved in order to advance the computation. Of course, this is not the only set of values that could be saved; any nonsingular linear transformation of them could be used instead.

In practice, the implicit equation (2.1) is approximately solved by some iterative process. An initial guess $y_{n,0}$ is obtained with a predictor formula

$$\alpha_0^* y_{n,0} + \sum_{j=1}^{k^*} \alpha_j^* y_{n-j} = h \sum_{j=1}^{k^*} \beta_j^* y'_{n-j},$$

and subsequent iterates $y_{n,\mu}$ are given by an iteration of the form

$$y_{n,\mu+1} := y_{n,\mu} - [\alpha_0 1 - h\beta_0 J_{n,\mu}]^{-1} [\alpha_0 y_{n,\mu} - h\beta_0 f_{n,\mu} - s_{n-1}^{k-1}],$$

$$\mu = 0(1)M(n) - 1,$$

where $f_{n,\mu} := f(t_n, y_{n,\mu})$, 1 is the identity matrix and $J_{n,\mu}$ is a matrix depending on the type of corrector iteration. For functional iteration $J_{n,\mu} := 0$, for the Newton iteration $J_{n,\mu} := f_y(t_n, y_{n,\mu})$, and for the chord iteration $J_{n,\mu} := f_y(t_n, y_{n,0})$. The number of iterations $M(n)$ might vary from step to step. After determining $y_n := y_{n,M(n)}$ a final function evaluation may or may not be performed. For a P(EC)* formula $y'_n := f_{n,M(n)-1}$ is accepted and for a P(EC)*E formula $y'_n := f_n = f_{n,M(n)}$. For our purposes the essential difference between the two types of predictor-corrector schemes is that the values y_n and y'_n satisfy the corrector formula for the P(EC)* scheme and they satisfy the differential equation for the P(EC)*E scheme. In fact, one could view the P(EC)* formula as computing values y_n, y'_n that satisfy a slightly different differential equation, which implies that equivalence results for linear formulas apply equally well to P(EC)* formulas.

How many values must be saved between steps for these predictor-corrector formulas? The answer would seem to be the row rank of the following matrix:

$$\left[\begin{array}{cccc}
 \alpha_K & & & \beta_K \\
 \alpha_{K-1} & \alpha_K & \circ & \beta_{K-1} \\
 \vdots & \vdots & \diagdown & \vdots \\
 \vdots & \vdots & & \vdots \\
 \alpha_1 & \alpha_2 & \dots & \alpha_K & \beta_1 & \beta_2 & \dots & \beta_K \\
 \alpha_K^* & & & \beta_K^* \\
 \alpha_{K-1}^* & \alpha_K^* & \circ & \beta_{K-1}^* \\
 \vdots & \vdots & \diagdown & \vdots \\
 \vdots & \vdots & & \vdots \\
 \alpha_1^* & \alpha_2^* & \dots & \alpha_K^* & \beta_1^* & \beta_2^* & \dots & \beta_K^*
 \end{array} \right],$$

where $K := \max\{k, k^*\}$. The row rank of this matrix is at least K and could be as high as $k + k^*$. For reasons of storage economy the predictor should be chosen so that the rank of the matrix is low. The most economical choice would be to use a predictor of the form

$$y_{n,0} = \gamma_0 s_{n-1}^0 + \gamma_1 s_{n-1}^1 + \dots + \gamma_{k-1} s_{n-1}^{k-1},$$

where the coefficients γ_j are chosen to make the formula of order $k - 1$ or greater.

A predictor of order $k - 1$ may not be accurate enough for a corrector of order k or greater. For this reason it may be preferable to store more than k values between steps. For a predictor of order at least q where $q \geq k - 1$, one would expect to need an additional $m = q - (k - 1)$ values. This can be done by including information from preceding meshpoints. Since each step of the computation introduces exactly one new item of information, namely y'_n , the following set of $k + m$ values is suggested:

$$s_{n-m}^j, \quad j = 0(1)k - 1, \quad hy'_{n-j}, \quad j = 0(1)m - 1.$$

From these values one can apply the corrector formula to generate $y_{n-m+j} s_{n-m+j}^{k-1}, \dots, s_{n-m+j}^0, j = 1(1)m$. Moreover, if the corrector formula is of order q or more, then the polynomial of degree q or less which interpolates the original set of $k + m$ values also interpolates the values generated from them. An alternative set of values from which the others can be recovered (by reversing the forward step procedure) is the following:

$$s_{n-m}^j, \quad j = 0(1)k - m - 1^*, \quad y_{n-j}, \quad j = 0^*(1)m - 1, \quad hy'_{n-j}, \quad j = 0(1)m - 1,$$

where $0^* = 0, 1^* = 1$ if $m \leq k$ and $0^* = 1, 1^* = 0$ if $m = k + 1$. This set has the important advantage of requiring values from only $k + 0^*$ meshpoints instead of $k + m$ meshpoints, although it is a little more awkward to specify. The case $m = k + 1$ can only apply to the trapezoid formula, the Milne formula, and other (unstable) k -step formulas of maximal order $2k$. For this case the set of values $y_{n-j}, j = 1(1)k$, and $hy'_{n-j}, j = 0(1)k$, is used to define a polynomial of degree at most $2k$.

This polynomial also interpolates y_n because y_n is determined from the other values by a corrector formula which is exact for polynomials of degree at most $2k$.

As an example, a k th order predictor for the $(k + 1)$ st-order Adams-Moulton formula would be constructed from the values

$$h \sum_{i=0}^j \beta_{k-j+i} y'_{n-1-i} \quad j = 0(1)k - 2, y_n, \text{ and } hy'_n$$

or equivalently linear combinations of these values: y_n , and hy'_{n-j} , $j = 0(1)k - 1$. The predictor so constructed is the k th-order Adams-Bashforth formula.

It is not obvious, however, that such a predictor can always be constructed. Clearly, the question is of practical interest only for implicit formulas (ρ, σ) of order at least q . It is claimed that the construction of a suitable predictor is possible if and only if a unique polynomial $p(t)$ of degree q or less is uniquely determined from the values $L_h^j p(t_{n-m})$, $j = 0(1)k - 1$, and $p'(t_{n-j})$, $j = 0(1)m - 1$, where the operators L_h^j are defined by

$$L_h^j z(t) = \sum_{i=0}^j (-\alpha_{k-j+i} z(t - ih) + h\beta_{k-j+i} z'(t - ih)).$$

Clearly, the possibility of constructing $p(t)$ implies the existence of a suitable predictor. On the other hand, if $p(t_{n+1})$ can be determined by a predictor, then from the corrector formula we can get $p'(t_{n+1})$. This can be repeated to generate $p(t_{n+2})$, $p(t_{n+3})$, . . . until there are enough values to construct $p(t)$ by interpolation of ordinate values. Necessary and sufficient conditions for the possible construction of $p(t)$ is given by the theorem that follows. First a precise definition of order is needed.

Definition. A linear multistep formula (ρ, σ) is of *order* at least q if

$$\frac{\rho(1+z)}{\sigma(1+z)} = \log(1+z) + O(z^{q+1})$$

and is of *formal order* at least q if

$$\rho(1+z) = \log(1+z)\sigma(1+z) + O(z^{q+1})$$

(cf. Gear [1971a, p. 119]).

THEOREM 2.1. *Let (ρ, σ) be a linear k -step formula of order at least q . Values $L_h^j p(t_{n-m})$, $j = 0(1)k - 1$, and $p'(t_{n-j})$, $j = 0(1)m - 1$, uniquely determine a polynomial $p(t)$ of degree at most q if and only if $\rho(\xi)$ and $\sigma(\xi)$ have no common factors.*

Proof. If polynomials are identified with the column vector of their scaled derivatives, then the unique existence of $p(t)$ is equivalent to the linear independence of the rows of the matrix $T := \text{col}(\lambda_0^T, \lambda_1^T, \dots, \lambda_{k-1}^T, \mathbf{e}_1^T, \mathbf{e}_1^T P^{-1}, \dots, \mathbf{e}_1^T P^{1-m})$ where

$$\lambda_j^T := \sum_{i=0}^j \{-\alpha_{k-j+i} \mathbf{e}_0^T + \beta_{k-j+i} \mathbf{e}_1^T\} P^{-i-m},$$

and P is the $(q + 1) \times (q + 1)$ Pascal triangle matrix. The row vectors λ_j^T , $j =$

$0(1)k - 1$, have a simple generating function, for

$$\begin{aligned}
 &(-\rho(\xi)\mathbf{e}_0^T + \sigma(\xi)\mathbf{e}_1^T)(I - \xi P^{-1})^{-1}P^{-m} \\
 &= \sum_{j=0}^k \xi^{k-j}(-\alpha_j\mathbf{e}_0^T + \beta_j\mathbf{e}_1^T) \sum_{i=0}^{\infty} \xi^i P^{-i-m} = \sum_{l=0}^{k-1} \xi^l \lambda_l^T,
 \end{aligned}$$

where we have used the fact that $\lambda_k^T P^{k-1} = \mathbf{0}^T$. Let $\rho(\xi) = \pi(\xi)\bar{\rho}(\xi)$ and $\sigma(\xi) = \pi(\xi)\bar{\sigma}(\xi)$, where $\pi(\xi)$ is the g.c.d. of $\rho(\xi)$ and $\sigma(\xi)$. Then $(\bar{\rho}, \bar{\sigma})$ is of order q or greater, and so $(-\bar{\rho}(\xi)\mathbf{e}_0^T + \bar{\sigma}(\xi)\mathbf{e}_1^T)(I - \xi P^{-1})^{-1}P^{-m}$ is a generating function for the $\bar{\gamma}_j^T$, which are defined in the obvious way. The relation

$$(-\rho(\xi)\mathbf{e}_0^T + \sigma(\xi)\mathbf{e}_1^T)(I - \xi P^{-1})^{-1}P^{-m} = \pi(\xi)(-\bar{\rho}(\xi)\mathbf{e}_0^T + \bar{\sigma}(\xi)\mathbf{e}_1^T)(I - \xi P^{-1})^{-1}P^{-m}$$

implies that the first k rows of T are linear combinations of the row vectors $\bar{\gamma}_j^T$, $j = 0(1)\bar{k} - 1$. If $\rho(\xi)$ and $\sigma(\xi)$ have a common factor so that $\bar{k} < k$, then $p(t)$ is obviously not uniquely determined. Assume that $\rho(\xi)$ and $\sigma(\xi)$ have no common factors. Suppose that the rows of T are linearly dependent. Then, $T\mathbf{v} = \mathbf{0}$ for some $\mathbf{v} \neq \mathbf{0}$. Hence,

$$(-\rho(\xi)\mathbf{e}_0^T + \sigma(\xi)\mathbf{e}_1^T)(I - \xi P^{-1})^{-1}P^{-m}\mathbf{v} = \mathbf{0}$$

or equivalently

$$(2.2) \quad \rho(\xi)\tau_0(\xi) = \sigma(\xi)\tau_1(\xi),$$

where $\tau_i(\xi) = \det((I - \xi P^{-1})P^m)\mathbf{e}_i^T(I - \xi P^{-1})^{-1}P^{-m}\mathbf{v}$, which by Cramer's rule is simply the determinant of $(I - \xi P^{-1})P^m$ with the i th column replaced by \mathbf{v} . Thus, $\tau_i(\xi)$ is a nonzero polynomial of degree q or less. Furthermore,

$$\begin{aligned}
 \tau_1(\xi) &= -(1 - \xi)^{q+1}\xi^{-1}\mathbf{e}_1^T(I - \xi^{-1}P)^{-1}P^{1-m}\mathbf{v} \\
 &= (1 - \xi)^{q+1} \sum_{j=0}^{\infty} \xi^{-1-j}\mathbf{e}_1^T P^{j-m+1}\mathbf{v} \\
 &= (-1)^q \xi^{q-m}\mathbf{e}_1^T P\mathbf{v} + \text{lower-order terms,}
 \end{aligned}$$

and so $\tau_1(\xi)$ is of degree at most $k - 1$. Hence, from (2.2) at least one of the k roots of $\rho(\xi)$ must be a root of $\sigma(\xi)$. This is a contradiction; and therefore, the rows of T are linearly independent. \square

COROLLARY. *The values $L_h^j p(t_{n-m}), j = 0(1)k - m - 1^*, p(t_{n-j}), j = 0^*(1)m - 1$, and $p'(t_{n-j}), j = 0(1)m - 1$, uniquely determine $p(t)$ if and only if $\rho(\xi)$ and $\sigma(\xi)$ have no common factor.*

Remark 1. It is not enough to have formal order of at least q . Consider $\rho(\xi) = (\xi - 1)^2$ and $\sigma(\xi) = 0$; the values $-p(t_n)$ and $2p(t_n) - p(t_{n-1})$ uniquely determine a polynomial of degree one or less.

Remark 2. Theorem 2.1 appears to be related to a result of Dahlquist [1975] concerning the equivalence between a linear k -step formula and the corresponding one-

leg k -step formula, under the assumption that $\rho(\xi)$ and $\sigma(\xi)$ have no common factor.

Apart from the need for a predictor, there is another complication that affects the amount of storage needed. In practice, it has been found desirable to vary the step-size, and thus a variable step form of the multistep formula must be used. For any given fixed step formula there are numerous variable step formulas, some of which require less storage than others. This topic will be studied in a future paper.

3. Construction of Linear Nordsieck Formulas. In this section linear multistep formulas are described in terms of local polynomial approximations, and an equivalent Nordsieck formula is constructed by representing the polynomials in terms of their coordinates with respect to local monomial bases.

Consider a linear k -step formula of order at least q , where $q \geq k$, such that $\rho(\xi)$ and $\sigma(\xi)$ have no common factor. (The case $q = k - 1$ is treated at the end of this section.) Then as a consequence of the corollary to Theorem 2.1 a unique polynomial of degree at most q is determined by the $q + 1$ conditions

$$\begin{aligned} L_n^j p_{n-1}(t_{n-1-m}) &= s_{n-1-m}^j, & j &= 1(1)k - m - 1^*, \\ p_{n-1}(t_{n-1-j}) &= y_{n-1-j}, & j &= 0^*(1)m - 1, \\ p'_{n-1}(t_{n-1-j}) &= y'_{n-1-j}, & j' &= 0(1)m - 1. \end{aligned}$$

The polynomial $p_{n-1}(t)$ can be regarded as an approximation to the solution $y(t)$ near $t = t_{n-1}$, and it can be used to obtain a predicted value $y_{n,0} = p_{n-1}(t_n)$ as an initial approximation to y_n . In terms of the values y_{n-1}, \dots, y_{n-k} , and $y'_{n-1}, \dots, y'_{n-k}$, the relation $y_{n,0} = p_{n-1}(t_n)$ becomes an explicit linear k -step formula.

Advancing the numerical solution one step yields values which determine the polynomial $p_n(t)$. However, there is a more direct way of expressing $p_n(t)$ in terms of $p_{n-1}(t)$. First, we examine how y_n and y'_n are determined. From the multistep formula we have

$$\alpha_0 y_n = h\beta_0 y'_n + s_n^{k-1};$$

and because L_n^k annihilates $p_{n-1}(t)$, we have

$$\alpha_0 y_{n,0} = h\beta_0 y'_{n,0} + L_n^{k-1} p_{n-1}(t_{n-1}),$$

where $y'_{n,0} = p'_{n-1}(t_n)$. It follows from the defining conditions that $L_n^{k-1} p_{n-1}(t_{n-1}) = s_n^{k-1}$. Hence, $\alpha_0(y_n - y_{n,0}) = h\beta_0(y'_n - y'_{n,0})$. Also, $y'_n = f(t_n, y_n)$. Together these last two equations define y_n and y'_n . A more convenient way of expressing this is to introduce an increment δ_n which satisfies

$$hy'_{n,0} + \alpha_0 \delta_n = hf(t_n, y_{n,0} + \beta_0 \delta_n)$$

and set

$$y_n = y_{n,0} + \beta_0 \delta_n, \quad hy'_n = hy'_{n,0} + \alpha_0 \delta_n.$$

Second, we construct $d_n(t) := p_n(t) - p_{n-1}(t)$. For $j = 0(1)k - m - 1^*$,

$$\begin{aligned} L_h^j d_n(t_{n-m}) &= L_h^j p_n(t_{n-m}) - \{-\alpha_{k-j} p_{n-1}(t_{n-m}) + h\beta_{k-j} p'_{n-1}(t_{n-m}) \\ &\quad + L_h^{j-1} p_{n-1}(t_{n-1-m})\} \\ &= s_{n-m}^j - \{-\alpha_{k-j} y_{n-m} + h\beta_{k-j} y'_{n-m} + s_{n-1-m}^{j-1}\} \\ &= 0; \end{aligned}$$

for $j = 0^*(1)m - 1$,

$$d_n(t_{n-j}) = \begin{cases} \beta_0 \delta_n & \text{if } j = 0, \\ 0 & \text{otherwise;} \end{cases}$$

and for $j = 0(1)m - 1$,

$$hd'_n(t_{n-j}) = \begin{cases} \alpha_0 \delta_n & \text{if } j = 0, \\ 0 & \text{otherwise.} \end{cases}$$

Hence,

$$d_n(t) = \Lambda\left(\frac{t - t_n}{h}\right) \delta_n,$$

where $\Lambda(x)$ is the unique polynomial of degree at most q which satisfies

$$L_1^j \Lambda(-m) = 0, \quad j = 0(1)k - m - 1^*,$$

$$\Lambda(-j) = \begin{cases} \beta_0, & j = 0^*(1)0, \\ 0, & j = 1(1)m - 1, \end{cases}$$

$$\Lambda'(-j) = \begin{cases} \alpha_0, & j = 0, \\ 0, & j = 1(1)m - 1. \end{cases}$$

Briefly, we have shown that the numerical solution is advanced one step by setting

$$(3.1) \quad p_n(t) = p_{n-1}(t) + \Lambda\left(\frac{t - t_n}{h}\right) \delta_n,$$

where δ_n is chosen so that $p_n(t)$ satisfies the differential equation and $\Lambda(x)$ is characteristic of the multistep formula.

The polynomial formulation of the k th and $(k + 1)$ st order Adams-Moulton formulas was discovered by Descloux [1963]. Schemes based on general choices of $\Lambda(x)$ are discussed by Skeel [1973] and by Wallace and Gupta [1973], who give an interesting interpretation of polynomial schemes in terms of polynomial predictive filters. They derive new formulas for stiff problems by choosing $\Lambda(x)$ to be a monic polynomial which best approximates zero for $x < 0$. Different types of approximations yield different formulas. Still more formulas are presented in Gupta and Wallace [1975] and Gupta [1976]. The 1975 paper uses the term *modifier polynomial* for $\Lambda(x)$. (In the 1973 paper this term is applied to $\Lambda(x/h)$.)

A very simple identification of these polynomial schemes with Nordsieck formulas becomes apparent if the polynomials are represented by vectors of scaled derivatives. Let $\mathbf{a}_n = [p_n(t_n), hp'_n(t_n), \dots, h^q p_n^{(q)}(t_n)/q!]^T$ and $\mathbf{1} = [\Lambda(0), \Lambda'(0), \dots, \Lambda^{(q)}(0)/q!]^T$. Then from (3.1) we have

$$p_n^{(j)}(t_n) = p_{n-1}^{(j)}(t_n) + h^{-j} \Lambda^{(j)}(0) \delta_n,$$

from which it easily follows that

$$\mathbf{a}_n = P\mathbf{a}_{n-1} + \mathbf{1}\delta_n$$

with δ_n chosen so that $h^{-1} \mathbf{e}_1^T \mathbf{a}_n = f(t_n, \mathbf{e}_0^T \mathbf{a}_n)$.

For the k th-order backward differentiation formula

$$\sum_{j=0}^k \alpha_j y_{n-j} = h y'_n,$$

the modifier polynomial $\Lambda(x)$ must satisfy

$$\Lambda(-j) = 0, \quad j = 1(1)k-1, \quad \Lambda(0) = 1, \quad \Lambda'(0) = \alpha_0,$$

whence

$$\begin{aligned} \Lambda(-k) &= \left\{ \Lambda'(0) - \sum_{j=0}^{k-1} \alpha_j \Lambda(-j) \right\} / \alpha_k \\ &= 0. \end{aligned}$$

Therefore,

$$\Lambda(x) = \binom{x+k}{k} = \sum_{j=0}^k \frac{1}{k!} \left[\begin{matrix} k+1 \\ j+1 \end{matrix} \right] x^j,$$

where the (square) brackets denote Stirling numbers of the first kind (see, for example Knuth [1968, p. 66]). The fact that $\Lambda(x)$ vanishes for $x = -1, -2, \dots, -k$ means that the values $y_{n-1}, y_{n-2}, \dots, y_{n-k}$ are "remembered" after advancing from t_{n-1} to t_n .

For the $(k+1)$ st-order Adams-Moulton formula

$$y_n - y_{n-1} = h \sum_{j=0}^k \beta_j y'_{n-j}$$

the modifier polynomial of the $(k+2)$ -value form must satisfy

$$\Lambda'(-j) = 0, \quad j = 2(1)k-1,$$

$$\Lambda(0) = \beta_0, \quad \Lambda'(0) = 1, \quad \Lambda(-1) = \Lambda'(-1) = 0,$$

whence

$$\begin{aligned} \Lambda'(-k) &= \left\{ \Lambda(0) - \Lambda(-1) - \sum_{j=0}^{k-1} \beta_j \Lambda'(-j) \right\} / \beta_k \\ &= 0. \end{aligned}$$

Therefore $\Lambda'(x) = \binom{x+k}{k}$, and so

$$\Lambda(x) = \int_{-1}^x \binom{y+k}{k} dy = - \int_0^{-1} \binom{y+k}{k} dy + \int_0^x \binom{y+k}{k} dy.$$

In terms of Stirling numbers

$$\int_0^x \binom{y+k}{k} dy = \sum_{j=1}^{k+1} \frac{1}{jk!} \left[\begin{matrix} k+1 \\ j \end{matrix} \right] x^j.$$

These are the formulas used by the nonstiff options of the codes DIFSUB of Gear [1971], GEAR Rev. 3 of Hindmarsh [1974], and EPISODE of Byrne and Hindmarsh [1975] when the stepsize does not vary.

The $(k+1)$ st-order Adams-Moulton formula could also be written as a $(k+1)$ -value Nordsieck formula. The modifier polynomial for such a formula can be determined by applying the theorem that follows.

THEOREM 3.1. *Let $\Lambda(x)$ be the modifier polynomial of the $(q+1)$ -value form of a linear k -step formula of order at least q , where $q \geq k+1$ and $\rho(\xi)$ and $\sigma(\xi)$ have no common factors. Then $\hat{\Lambda}(x) := \Lambda(x) - \Lambda(x-1)$ is the modifier polynomial for the q -value form of the multistep formula.*

Proof. For $m \leq k$

$$\begin{aligned} L^{k-m} \Lambda(-m) &= L^k \Lambda(-m) - \sum_{i=0}^{m-1} \{-\alpha_i \Lambda(-i) + \beta_i \Lambda'(-i)\} \\ &= 0, \end{aligned}$$

and for $m = k+1$

$$\begin{aligned} \Lambda(0) &= \frac{\beta_0}{\alpha_0} \Lambda'(0) + \frac{1}{\alpha_0} \sum_{i=1}^k \{-\alpha_i \Lambda(-i) + \beta_i \Lambda'(-i)\} \\ &= \beta_0. \end{aligned}$$

Hence

$$L_1^j \Lambda(-m) = 0, \quad j = 0(1)k-m,$$

$$\Lambda(0) = \beta_0, \quad \Lambda'(0) = \alpha_0, \quad \Lambda(-j) = \Lambda'(-j) = 0, \quad j = 1(1)m-1,$$

from which the theorem follows. \square

Therefore, the $(k+1)$ -value form of the $(k+1)$ st-order Adams-Moulton formula has

$$\begin{aligned} \Lambda(x) &= \int_{-1}^x \binom{y+k}{k} dy - \int_{-1}^{x-1} \binom{y+k}{k} dy \\ &= \int_{-1}^0 \binom{y+k}{k} dy + \int_0^x \left\{ \binom{y+k}{k} - \binom{y+k-1}{k} \right\} dy \\ &= - \int_0^{-1} \binom{y+k}{k} dy + \int_0^x \binom{y+k-1}{k-1} dy. \end{aligned}$$

These are the original formulas of Nordsieck.

Let us consider now the case $m = 0$. The values δ_n , y_n , and y'_n are determined as before. However, for $j = 0(1)k - 1$

$$\begin{aligned} L_h^j d_n(t_n) &= -\alpha_{k-j}(y_n - y_{n,0}) + h\beta_{k-j}(y'_n - y'_{n,0}) \\ &= (-\alpha_{k-j}\beta_0 + \beta_{k-j}\alpha_0)\delta_n. \end{aligned}$$

Therefore, the solution is advanced one step as given by (3.1) with $\Lambda(x)$ determined by

$$L_1^j \Lambda(0) = -\alpha_{k-j}\beta_0 + \beta_{k-j}\alpha_0, \quad j = 0(1)k - 1.$$

However, $p_n(t)$ does not necessarily interpolate y_n and y'_n nor does it generally satisfy the differential equation at $t = t_n$. Therefore, in the scaled derivative form this scheme is a generalized Nordsieck formula in the sense that δ_n is determined by a condition other than the satisfaction of the differential equation at $t = t_n$. Such formulas are potentially useful because of their minimum storage property. Nordsieck [1962, p. 27] considers the possibility of having $l_0 = \beta_0$ but $l_1 \neq \alpha_0$, so that $p_n(t)$ interpolates y_n but not y'_n ; however, the results of his experiments were not promising. Also, Wallace and Gupta [1973] mention that "Other choices of δ_n are rationally possible, and we hope to explore some other choices later." Finally, it is worth noting that Theorem 3.1 extends to the case $q = k$.

4. The Correspondence Between Multistep and Nordsieck Formulas. In the preceding section it was shown how to construct the modifier polynomial $\Lambda(x)$ of a Nordsieck formula from the polynomials $\rho(\xi)$ and $\sigma(\xi)$ of a linear multistep formula provided that $\rho(\xi)$ and $\sigma(\xi)$ have no common divisors. In this section we show how to obtain $\rho(\xi)$ and $\sigma(\xi)$ from $\Lambda(x)$, and we establish a one-to-one correspondence between (i) the class of all $(q + 1)$ -value linear Nordsieck formulas and (ii) the class of all linear multistep formulas of formal order at least q and of stepnumber at most q (including those for which $\rho(\xi)$ and $\sigma(\xi)$ have common factors).

For each linear Nordsieck formula we define a *corresponding* linear multistep formula $(\tilde{\rho}, \tilde{\sigma})$ by

$$(4.1) \quad \tilde{\rho}(\xi) := \det(\xi I - P) \mathbf{e}_1^T (\xi I - P)^{-1} \mathbf{1}, \quad \tilde{\sigma}(\xi) := \det(\xi I - P) \mathbf{e}_0^T (\xi I - P)^{-1} \mathbf{1}.$$

Applying Cramer's rule to $\mathbf{e}_j^T (\xi I - P)^{-1} \mathbf{1}$ for $j = 1$ and $j = 0$ yields

$$\tilde{\rho}(\xi) = \det \begin{bmatrix} \xi - 1 & l_0 & -1 & \cdots & -1 \\ 0 & l_1 & -2 & \cdots & -q \\ 0 & l_2 & \xi - 1 & & -\binom{q}{2} \\ \vdots & \vdots & & & \\ 0 & l_q & 0 & & \xi - 1 \end{bmatrix}$$

and

$$\tilde{\sigma}(\xi) = \det \begin{bmatrix} l_0 & -1 & -1 & \cdots & -1 \\ l_1 & \xi - 1 & -2 & \cdots & -2 \\ l_2 & 0 & \xi - 1 & & -\binom{q}{2} \\ \vdots & \vdots & & & \\ l_q & 0 & 0 & & \xi - 1 \end{bmatrix}.$$

Hence $\tilde{\sigma}(\xi)$ is a polynomial of degree q or less, and $\tilde{\rho}(\xi)$ is of degree q since $l_1 \neq 0$.

Let

$$\tilde{\rho}(\xi) =: \sum_{i=0}^q \alpha_i \xi^{q-i} \quad \text{and} \quad \tilde{\sigma}(\xi) =: \sum_{i=0}^q \beta_i \xi^{q-i}.$$

It might happen that $\alpha_q = \beta_q = 0$, and hence express

$$(4.2) \quad \tilde{\rho}(\xi) =: \xi^{m-1} \rho(\xi) \quad \text{and} \quad \tilde{\sigma}(\xi) =: \xi^{m-1} \sigma(\xi),$$

where $\alpha_k^2 + \beta_k^2 > 0$ and $m - 1 = q - k$.

The following theorem shows that the $\rho(\xi)$ and $\sigma(\xi)$ corresponding to $\Lambda(x)$ can be used to reconstruct $\Lambda(x)$ by the process of Section 3 if it is applicable.

THEOREM 4.1. *With $\rho(\xi)$ and $\sigma(\xi)$ derived from $\Lambda(x)$ by (4.1) and (4.2) we have*

$$\begin{aligned} L_1^j \Lambda(-m) &= 0, & j &= 0(1)k - 1, \\ l_0 = \Lambda(0) &= \beta_0, & l_1 = \Lambda'(0) &= \alpha_0, \\ \Lambda(-j) = \Lambda'(-j) &= 0, & j &= 1(1)m - 1, \\ \Lambda(-m) &= (-1)^q \beta_k, & \Lambda'(-m) &= (-1)^q \alpha_k. \end{aligned}$$

Proof. The relations $l_0 = \beta_0$ and $l_1 = \alpha_0$ follow from the expressions for $\rho(\xi)$ and $\sigma(\xi)$ as determinants. From (4.1) we have

$$\begin{aligned} \tilde{\rho}(\xi) &= -(\xi - 1)^{q+1} \mathbf{e}_1^T (I - \xi P^{-1})^{-1} P^{-1} \mathbf{1} \\ &= -(\xi - 1)^{q+1} \sum_{j=0}^{\infty} \mathbf{e}_1^T \xi^{-j} P^{-j-1} \mathbf{1} = -(\xi - 1)^{q+1} \sum_{j=0}^{\infty} \Lambda'(-j-1) \xi^j. \end{aligned}$$

Clearly, $\Lambda'(-j) = 0$, $j = 1(1)m - 1$, $\alpha_k = (-1)^q \Lambda'(-m)$, and

$$\rho(\xi) = -(\xi - 1)^{q+1} \sum_{j=0}^{\infty} \Lambda'(-j-m) \xi^j.$$

Similarly, $\Lambda(-j) = 0$, $j = 1(1)m - 1$, $\beta_k = (-1)^q \Lambda(-m)$, and

$$\sigma(\xi) = -(\xi - 1)^{q+1} \sum_{j=0}^{\infty} \Lambda(-j-m) \xi^j.$$

From these expressions for $\rho(\xi)$ and $\sigma(\xi)$ we get

$$-\rho(\xi) \sum_{j=0}^{\infty} \Lambda(-j-m)\xi^j + \sigma(\xi) \sum_{j=0}^{\infty} \Lambda'(-j-m)\xi^j = 0.$$

Equating coefficients of the powers of ξ completes the proof. \square

The next theorem shows that the multistep formula $(\tilde{\rho}, \tilde{\sigma})$ corresponding to $\Lambda(x)$ is of order at least q , and therefore, it can be used to reconstruct $\Lambda(x)$ by the process of Section 3, if $\rho(\xi)$ and $\sigma(\xi)$ have no common factors.

THEOREM 4.2. *The linear multistep formula corresponding to a $(q + 1)$ -value linear Nordsieck formula has formal order at least q . Moreover, it is of order at least $q + 1$ if and only if $\mathbf{b}^T \mathbf{1} = 0$, where the components of \mathbf{b} are the Bernoulli numbers defined by*

$$\frac{z}{e^z - 1} = \sum_{j=0}^{\infty} \frac{b_j z^j}{j!}.$$

Proof. According to the definition of formal order and Eq. (4.1), it must be shown that

$$z^{q+1} \mathbf{e}_1^T ((1+z)I - P)^{-1} \mathbf{1} = \log(1+z) z^{q+1} \mathbf{e}_0^T ((1+z)I - P)^{-1} \mathbf{1} + O(z^{q+1});$$

and so it is enough to show that

$$\mathbf{e}_1^T (I - z^{-1}F)^{-1} = \log(1+z) \mathbf{e}_0^T (I - z^{-1}F)^{-1} + O(z),$$

where $F := P - I$. Because $F^{q+1} = 0$, it is not difficult to verify that

$$\log(1+z)(I - z^{-1}F)^{-1} = \log(I + F)(I - z^{-1}F)^{-1} + O(z);$$

and so it suffices to show that $\mathbf{e}_0^T \log(I + F) = \mathbf{e}_1^T$. This holds because $I + F = P = \exp(Q)$, where

$$Q = \begin{bmatrix} 0 & 1 & & & \\ & 0 & & \circ & \\ & & 2 & & \\ \circ & & & 0 & \\ & & & & q \\ & & & & & 0 \end{bmatrix}.$$

To prove the second assertion, retrace the first two steps of the proof to get the following condition for formal order of at least $q + 1$:

$$\mathbf{e}_0^T \left\{ I - \frac{1}{2}F + \dots + \frac{1}{q+1}(-F)^q \right\} \mathbf{1} = 0.$$

The matrix in braces is

$$\frac{\log(I + F)}{F} = \frac{Q}{\exp(Q) - 1} = \sum_{j=0}^q \frac{b_j}{j!} Q^j,$$

from which the second part of the theorem follows. \square

Remark 1. The matrices P , F , and Q obviously represent the linear operators shift, forward difference, and derivative, respectively, for polynomials of degree q or less. Defining $B = I - P^{-1}$ to be the matrix representing the backward difference operator ∇ it is not difficult to show that the condition $\mathbf{b}^T \mathbf{1} = 0$ is equivalent to

$$\sum_{j=0}^q \frac{1}{j+1} \nabla^j \Lambda(-1) = 0$$

as well as

$$\Lambda(-1) - \sum_{j=0}^{q-1} \gamma_{j+1}^* \nabla^j \Lambda'(-1) = 0,$$

where the γ_{j+1}^* are the coefficients for the backward difference form of the Adams-Moulton formulas. These conditions on the polynomial $\Lambda(x)$ arise in another situation in Henrici [1962, p. 342].

Remark 2. The condition $\mathbf{b}^T \mathbf{1} = 0$ is obtained by Wallace and Gupta [1973], although the coefficients b_j are not identified as the Bernoulli numbers. Instead, a recursive definition is given for the b_j , which should read

$$b_0 = 1, \quad \sum_{i=0}^j \binom{j+1}{i} b_i = 0.$$

Remark 3. It is shown by Gear [1966] that a stable linear Nordsieck method is convergent of order q in all components of \mathbf{a}_n , and it is shown by Skeel and Jackson [1977] that it is convergent of order $q+1$ in all components if and only if $\mathbf{b}^T \mathbf{1} = 0$.

It has thus been shown that to each linear Nordsieck formula, which is uniquely specified in terms of $\mathbf{1}$, there corresponds a linear multistep formula of formal order at least q and of stepnumber at most q . We now establish a correspondence in the opposite direction. We have from (4.1) that

$$\tilde{\sigma}(1+z) = z^{q+1} \mathbf{e}_0^T (zI - F)^{-1} \mathbf{1} = z^q \mathbf{e}_0^T \sum_{j=0}^q z^{-j} F^j \mathbf{1},$$

where $F := P - I$ is the matrix of the forward difference operator. Equating coefficients of powers of z gives $\tilde{\sigma}^{(j)}(1)/j! = \Delta^{q-j} \Lambda(0)$, and hence $\Lambda(x)$ is determined from its corresponding $\tilde{\sigma}(\xi)$ by Newton's forward difference formula

$$(4.3) \quad \Lambda(x) = \sum_{j=0}^q \frac{\tilde{\sigma}^{(j)}(1)}{j!} \binom{x}{q-j}.$$

Since the linear multistep formula is uniquely specified by $\tilde{\sigma}(\xi)$, the one-to-one correspondence is established. The inverse transformation is conveniently expressed as

$$\tilde{\sigma}(\xi) = \sum_{j=0}^q \Delta^j \Lambda(0) (\xi - 1)^{q-j} \quad \text{and} \quad \tilde{\rho}(\xi) = \sum_{j=0}^{q-1} \Delta^j \Lambda'(0) (\xi - 1)^{q-j}.$$

Note from (4.3) that the popular normalization $\sigma(1) = 1$ corresponds to $l_q = 1/q!$.

Remark. Theorem 3.1 can be generalized to any modifier polynomial $\Lambda(x)$ for which $m \geq 2$. Express $\tilde{\sigma}(\xi) =: \xi \hat{\sigma}(\xi)$. Then

$$\tilde{\sigma}^{(j)}(1) = \hat{\sigma}^{(j)}(1) + j\hat{\sigma}^{(j-1)}(1),$$

from which it can be shown that

$$\Lambda(x) = \sum_{j=0}^{q-1} \frac{\hat{\sigma}^{(j)}(1)}{j!} \binom{x+1}{q-j}$$

and

$$\hat{\Lambda}(x) := \Lambda(x) - \Lambda(x-1) = \sum_{j=0}^{q-1} \frac{\hat{\sigma}^{(j)}(1)}{j!} \binom{x}{q-1-j}.$$

Clearly, $\hat{\Lambda}(x)$ is the modifier polynomial for the q -value formula.

5. Equivalence of Linear Nordsieck Formulas to Linear Multistep Formulas.

Recall that a linear Nordsieck formula determines successive values by the system of difference equations

$$(5.1) \quad \mathbf{a}_n = P\mathbf{a}_{n-1} + \mathbf{l}\delta_n,$$

where δ_n is chosen so that $y'_n = f(t_n, y_n)$. We do not consider the more general formula $\mathbf{a}_n = A\mathbf{a}_{n-1} + \mathbf{l}\delta_n$ because formulas with $A \neq P$ are of dubious value, and in any case, most of the results for $A = P$ generalize if minor restrictions are imposed on A .

For theoretical purposes a rewriting of (5.1) is often useful. Premultiplying (5.1) by \mathbf{e}_1^T yields $\delta_n = l_1^{-1}(hf_n - \mathbf{e}_1^T P\mathbf{a}_{n-1})$. Let $\tilde{\mathbf{I}} := l_1^{-1}\mathbf{I}$ and $S := (I - \tilde{\mathbf{I}}\mathbf{e}_1^T)P$; and we get

$$(5.2) \quad \mathbf{a}_n = S\mathbf{a}_{n-1} + h\tilde{\mathbf{l}}f_n,$$

where f_n is chosen so that $f_n = f(t_n, \mathbf{e}_1^T S\mathbf{a}_{n-1} + h\tilde{\mathbf{l}}_0 f_n)$. Expressions for $\tilde{\rho}(\xi)$ and $\tilde{\sigma}(\xi)$ in terms of S are given by the following theorem, whose proof uses an idea from Osborne [1966, Eq. (4.3)].

THEOREM 5.1. *The polynomials $\tilde{\rho}(\xi)$ and $\tilde{\sigma}(\xi)$ defined by (4.1) satisfy*

$$\tilde{\rho}(\xi) = \det(\xi I - S)\mathbf{e}_1^T(\xi I - S)^{-1}\mathbf{1}$$

and

$$\tilde{\sigma}(\xi) = \det(\xi I - S)\mathbf{e}_0^T(\xi I - S)^{-1}\mathbf{1}.$$

Proof. We have

$$\xi I - S = \xi I - P + \tilde{\mathbf{I}}\mathbf{e}_1^T P = (\xi I - P)(I + (\xi I - P)^{-1}\tilde{\mathbf{I}}\mathbf{e}_1^T P),$$

and so

$$\begin{aligned} \det(\xi I - S) \cdot (\xi I - S)^{-1}\tilde{\mathbf{I}} &= \det(\xi I - P) \det(I + (\xi I - P)^{-1}\tilde{\mathbf{I}}\mathbf{e}_1^T P) \\ &\quad \cdot (I + (\xi I - P)^{-1}\tilde{\mathbf{I}}\mathbf{e}_1^T P)^{-1}(\xi I - P)^{-1}\tilde{\mathbf{I}}. \end{aligned}$$

This can be simplified by means of the identity

$$\det(I + \mathbf{xy}^T) \cdot (I + \mathbf{xy}^T)^{-1}\mathbf{x} = \mathbf{x},$$

which follows from Cramer's rule. Thus, we get

$$\det(\xi I - S) \cdot (\xi I - S)^{-1} \tilde{\Gamma} = \det(\xi I - P) \cdot (\xi I - P)^{-1} \tilde{\Gamma},$$

from which the lemma immediately follows. \square

Note that $\mathbf{e}_1^T S = \mathbf{0}^T$; and so $\tilde{\rho}(\xi) = \det(\xi I - S) \xi^{-1} l_1$, which gives the characteristic polynomial of S as $\det(\xi I - S) = l_1^{-1} \xi^m \rho(\xi)$. Thus, the strict root condition is satisfied by the linear Nordsieck formula (cf. Skeel and Jackson [1977]) if and only if it is satisfied by the corresponding linear multistep formula.

The theorem that follows shows that in the case of *all* Nordsieck formulas the zeroth and first components of the vectors \mathbf{a}_n satisfy the difference equation of the corresponding linear multistep formula (ρ, σ) . Hence, all the limitations on the accuracy of multistep formulas (Dahlquist [1956], [1963]) apply also to Nordsieck formulas.

THEOREM 5.2. *The values $y_n, y'_n, n = 0(1)N$, computed from (5.1), satisfy the linear multistep formula*

$$\rho(E)y_{n-k} = h\sigma(E)y'_{n-k}$$

defined by (4.1) and (4.2) for $n = k(1)N$.

Proof. From (5.2) we have

$$S^{k-j} \mathbf{a}_{n-k} = \mathbf{a}_{n-j} - \sum_{i=j}^{k-1} S^{i-j} \tilde{\Gamma} h f_{n-i}$$

and

$$\rho(S) \mathbf{a}_{n-k} = \sum_{j=0}^k \alpha_j \mathbf{a}_{n-j} - \sum_{j=0}^{k-1} \alpha_j \sum_{i=j}^{k-1} S^{i-j} \tilde{\Gamma} h f_{n-i}.$$

Premultiplying by \mathbf{e}_0^T and rearranging gives

$$\sum_{j=0}^k \alpha_j y_{n-j} = \sum_{i=0}^{k-1} \mathbf{e}_0^T \left\{ \sum_{j=0}^i \alpha_j S^{i-j} \right\} \tilde{\Gamma} h f_{n-i} + \mathbf{e}_0^T \rho(S) \mathbf{a}_{n-k}.$$

It needs to be shown first that

$$(5.3) \quad \tilde{\sigma}(\xi) = \sum_{i=0}^q \mathbf{e}_0^T \left\{ \sum_{j=0}^i \alpha_j S^{i-j} \right\} \tilde{\Gamma} \xi^{q-i}.$$

We have that

$$\begin{aligned} \tilde{\sigma}(\xi) &= l_1^{-1} \xi \rho(\xi) \mathbf{e}_0^T (\xi I - S)^{-1} \mathbf{1} = \rho(\xi) \mathbf{e}_0^T (I - \xi^{-1} S)^{-1} \tilde{\Gamma} \\ &= \sum_{j=0}^q \alpha_j \xi^{q-j} \sum_{i=0}^{q-j} \xi^{-i} \mathbf{e}_0^T S^i \tilde{\Gamma}, \end{aligned}$$

which leads to (5.3). It remains to be shown that

$$\mathbf{e}_0^T \rho(S) = \beta_k \mathbf{e}_1^T.$$

Because S satisfies its characteristic polynomial, $\mathbf{e}_0^T \rho(S) S^m = \mathbf{0}^T$. Also, because S is of

rank q , it has only one linearly independent left eigenvector associated with the eigenvalue 0; and so $\mathbf{e}_0^T \rho(S)S^{m-1}$ is a multiple of \mathbf{e}_1^T . This implies $\mathbf{e}_0^T \rho(S)S^{m-1}(I - \tilde{\mathbf{1}}\mathbf{e}_1^T) = 0$, and so by (5.3)

$$\mathbf{e}_0^T \rho(S)S^{m-1} = \beta_{k+m-1} \mathbf{e}_1^T.$$

If $m > 1$, then $\mathbf{e}_0^T \rho(S)S^{m-1} = \mathbf{0}^T$, and this argument can be repeated until the assertion is established. \square

Remark 1. We have $y'_n = f_n$ except possibly for $n = 0$.

Remark 2. This theorem improves the result of Descloux [1963] and Osborne [1966] in two respects: first, $\rho(\xi)$ and $\sigma(\xi)$ are given in closed form, and second, the result is shown for $n \geq k$ rather than $n \geq q$.

The next theorem proves that Nordsieck and multistep methods are equivalent in the sense of Gear, for it is shown that there is a nonsingular matrix that relates the scaled derivative approximations \mathbf{a}_n to certain linear combinations of computed y and y' values. As a consequence, in practically all cases linear Nordsieck methods cannot be regarded as generalizations of linear multistep formulas (cf. Gear [1971a, pp. 102, 136]).

THEOREM 5.3. *Assume the polynomials $\rho(\xi)$ and $\sigma(\xi)$ corresponding to a given $(q + 1)$ -value Nordsieck formula have no common factors. Then there exists a unique nonsingular $(q + 1) \times (q + 1)$ matrix T depending only on l such that $\mathbf{a}_n = T^{-1}\mathbf{y}_n$ for $n = k - 1^*(1)N$, where*

$$\mathbf{y}_n := [s_{n-m}^0, \dots, s_{n-m}^{k-m-1^*}, y_{n-0^*}, \dots, y_{n-m+1}, hy'_n, \dots, hy'_{n-m+1}]^T.$$

Proof. By Theorem 4.2 the formula (ρ, σ) is of order at least q , and hence by the Corollary to Theorem 2.1 there exists a unique nonsingular matrix T such that

$$\begin{bmatrix} p(t_n) \\ hp'(t_n) \\ \vdots \\ h^q p^{(q)}(t_n)/q! \end{bmatrix} = T^{-1} \begin{bmatrix} L_h^0 p(t_{n-m}) \\ L_h^1 p(t_{n-m}) \\ \vdots \\ hp'(t_{n-m+1}) \end{bmatrix}$$

for any polynomial of degree q or less. Define values y_j and y'_j for $j = 1^* - k(1) - 1$ such that

$$\mathbf{y}_0 = T\mathbf{a}_0.$$

The process of Section 3 may be applied to construct a $(q + 1)$ -value linear Nordsieck formula which would compute vectors $T^{-1}\mathbf{y}_n$, $n = 1(1)N$. The results of Section 4 imply that this formula would be identical to the given Nordsieck formula, and hence $\mathbf{a}_n = T^{-1}\mathbf{y}_n$. The uniqueness of T follows from the fact that the method is exact for all polynomials of degree at most q if the starting values are exact. \square

Remark. For implicit formulas Wallace and Gupta [1973] give an informal argument suggesting that the quantities \mathbf{a}_{n-q} , and δ_{n-j} , $j = 1(1)q - 1$, can be ex-

pressed as linear combinations of $y_{n-j} f_{n-j}$ $j = 1(1)q$. Their conjecture is correct as long as $\xi^{m-1}\rho(\xi)$ and $\xi^{m-1}\sigma(\xi)$ have no common factors. In that case it is an immediate consequence that the components of \mathbf{a}_{n-1} are linear combinations of $y_{n-j} f_{n-j}$ $j = 1(1)q$.

Theorem 5.3 does not extend to Nordsieck formulas for which the corresponding ρ and σ polynomials have common factors.

THEOREM 5.4. *Assume that $\rho(\xi)$ and $\sigma(\xi)$ have a common factor and that the formula (ρ, σ) is of order at least q . Then \mathbf{a}_n cannot be expressed only in terms of $y_{n-j} hy'_{n-j}$ $j = 0(1)n$.*

Proof. We begin by showing that there exists an eigenvector \mathbf{v} associated with a nonzero eigenvalue ξ_0 of S such that $\mathbf{e}_0^T \mathbf{v} = 0$. From (4.3) it follows that $\rho(\xi)$ and $\sigma(\xi)$ have the common factor 1 if and only if $l_q = 0$. First, suppose that $l_q = 0$. Then $\xi = 1$ is a common root, and so the formula must be of formal order at least $k + 1$. Hence $\mathbf{b}^T \mathbf{1} = 0$ where

$$\mathbf{b}^T = \mathbf{e}_0^T \frac{\log(P)}{P - I},$$

and so

$$(\mathbf{b}^T + \mathbf{e}_1^T)S = \mathbf{b}^T P = \mathbf{b}^T + \mathbf{e}_0^T \log P = \mathbf{b}^T + \mathbf{e}_1^T.$$

Also, $l_q = 0$ implies $\mathbf{e}_q^T S = \mathbf{e}_q^T$. Therefore, the null space of $S - I$ is at least of dimension two; and so we can choose $\mathbf{v} \neq \mathbf{0}$ such that $S\mathbf{v} = \mathbf{v}$ and $\mathbf{e}_0^T \mathbf{v} = 0$. Second, suppose that $l_q \neq 0$. Let

$$\mathbf{v}(\xi) = \det(\xi I - S) \cdot (\xi I - S)^{-1} \mathbf{1},$$

which is a vector of polynomials in ξ . Let ξ_0 be a common root of $\rho(\xi)$ and $\sigma(\xi)$. Then $\mathbf{v}(\xi_0) \neq \mathbf{0}$, since $\mathbf{e}_q^T \mathbf{v}(\xi) = (\xi - 1)^q l_q$. Also

$$(\xi_0 I - S)\mathbf{v}(\xi_0) = \det(\xi_0 I - S) \cdot \mathbf{1} = \mathbf{0},$$

and $\mathbf{e}_0^T \mathbf{v}(\xi_0) = \sigma(\xi_0) = 0$. Therefore, in either case there exists $\xi_0 \neq 0$ and $\mathbf{v} \neq \mathbf{0}$ such that $S\mathbf{v} = \xi_0 \mathbf{v}$ and $\mathbf{e}_0^T \mathbf{v} = 0$. Moreover, $\mathbf{e}_1^T \mathbf{v} = 0$, because \mathbf{e}_1^T is a left eigenvector for the eigenvalue 0. This means that if \mathbf{a}_0 were changed to $\mathbf{a}_0 + \mathbf{v}$, the values of y_n and hy'_n would remain unchanged for all n , and yet \mathbf{a}_n would become $\mathbf{a}_n + \xi_0^n \mathbf{v}$. This proves the nonexistence of an expression for \mathbf{a}_n in terms of the values y_n and y'_n . \square

Theorems 5.3 and 5.4 do not cover the case where the order is less than q due to a common factor of $\xi - 1$, but extending the results to such formulas does not seem worthwhile.

6. Equivalence of Predictor-Corrector Formulas. The use of a linear Nordsieck formula requires the solution for δ_n of the equation

$$hy'_{n,0} + l_1 \delta_n = hf(t_n, y_{n,0} + l_0 \delta_n),$$

where $y_{n,0} = \mathbf{e}_0^T P \mathbf{a}_{n-1}$ and $hy'_{n,0} = \mathbf{e}_1^T P \mathbf{a}_{n-1}$. In practice, this must be approximately

solved by some iterative process:

$$\begin{aligned}\delta_{n,0} &:= 0, \\ \delta_{n,\mu+1} &:= \delta_{n,\mu} - [l_1 1 - l_0 h J_{n,\mu}]^{-1} [h y'_{n,0} + l_1 \delta_{n,\mu} - h f_{n,\mu}], \\ &\mu = 0, 1, \dots, M(n) - 1,\end{aligned}$$

where $f_{n,\mu}$, 1 , and $J_{n,\mu-1}$ are defined as in Section 2 with $y_{n,\mu} = y_{n,0} + l_0 \delta_{n,\mu}$. After determining $\delta_n := \delta_{n,M(n)}$ and adding the increment $l\delta_n$ a final function evaluation may or may not be performed. For a P(EC)* Nordsieck formula

$$\mathbf{a}_n := P\mathbf{a}_{n-1} + l\delta_n$$

and for a P(EC)*E Nordsieck formula

$$\mathbf{a}_n := P\mathbf{a}_{n-1} + l\delta_n + \mathbf{e}_1 h(f'_n - \bar{y}'_n),$$

where $h\bar{y}'_n := h y'_{n,0} + l_1 \delta_{n,M(n)-1}$.

Remark. The predictor-corrector Nordsieck formulas of Gear [1971a] are introduced independently of linear ("corrector only") Nordsieck formulas. For this reason these predictor-corrector formulas use $l_1 1 - l_0 h J_{n,\mu} := 1$.

For the P(EC)* formulas we have that $h y'_n = \mathbf{e}_1^T P\mathbf{a}_{n-1} + l_1 \delta_n$, and so the recurrence can also be expressed as

$$\mathbf{a}_n = S\mathbf{a}_{n-1} + \tilde{1} h y'_n.$$

Therefore, the equivalence results apply to P(EC)* formulas as much as they do to linear formulas. The underlying multistep predictor formula can be obtained from $y_{n,0} = \mathbf{e}_0^T P\mathbf{a}_{n-1}$ by expressing \mathbf{a}_{n-1} as linear combinations of the components of \mathbf{y}_{n-1} ; thus, it has stepnumber of at most $k + 0^*$.

The situation is quite different for P(EC)*E formulas. To determine the equivalent multistep formula, one begins with the recurrence

$$\mathbf{a}_n = S\mathbf{a}_{n-1} + (\tilde{1} - \mathbf{e}_1) h \bar{y}'_n + \mathbf{e}_1 h f'_n.$$

Proceeding as in the proof of Theorem 5.2, one obtains a difference equation involving y_{n-j} , $j = 0(1)k$, \bar{y}'_{n-j} , $j = 0(1)k - 1$, and y'_{n-j} , $j = 1(1)k$, which is not a true P(EC)*E multistep formula.

For example, consider the three-value Nordsieck formula. By expressing y_n and y_{n-1} as functions of \bar{y}'_n , \bar{y}'_{n-1} , f_{n-1} , and \mathbf{a}_{n-2} and then eliminating $h^2 y''_{n-2}/2$, one obtains the difference equation

$$\begin{aligned}y_n + (2l_2 - 2)y_{n-1} + (1 - 2l_2)y_{n-2} \\ = l_0 h \bar{y}'_n + [(1 - l_0) h y'_{n-1} + (l_2 - l_0) h \bar{y}'_{n-1}] + (l_0 + l_2 - 1) h y'_{n-2},\end{aligned}$$

which is not a P(EC)*E formula unless $l_0 = l_2$. For the third-order Adams-Moulton formula this is

$$y_n - y_{n-1} = \frac{5}{12} h \bar{y}'_n + \left[\frac{7}{12} h y'_{n-1} + \frac{1}{12} h \bar{y}'_{n-1} \right] + \frac{1}{12} h y'_{n-2}.$$

7. **Applications.** An important practical consequence of the equivalence theory is that all multistep formulas are minimum storage formulas. This idea is certainly implicit in the investigations by Gupta and Wallace of new multistep formulas. It is also the rationale for the computer search of Kong [1977] for k -step formulas of order k having the smallest error coefficient for a given angle α of $A(\alpha)$ -stability. Previous searches by Dill and Gear [1971] (the error coefficient for their formula should be 14.0, rather than 0.0108) and by Jain and Srivastava [1970] concentrated on formulas for which most of the trailing coefficients of $\sigma(\xi)$ were set to zero. The computer results of Kong suggested formulas which lead to a proof of the following result: for any $\alpha < \pi/2$ there exists an $A(\alpha)$ -stable k -step formula of order k . This is also proved by Grigorieff and Schroll [1977].

The paper of Nordsieck mentions that "the potential advantage of a more elaborate procedure in which the matrix hf_y is numerically computed at every step and \mathbf{l} is made a chosen function of hf_y , implying a nonlinear process tailored to the subject differential equation system, is an interesting topic for future investigation," where we have substituted our notation for his. The idea is developed further in a report of Gear [1966], which proposes a formula for scalar implicit differential equations of any order. For the equation $y' - f(t, y) = 0$ the formula becomes

$$\mathbf{l} := \mathbf{l}^{\text{AM}} - \mathbf{l}^{\text{BD}} hf_y$$

and

$$l_1 \mathbf{l} - l_0 h J_{n,\mu} := \mathbf{l} + h^2 f_y^2,$$

where AM and BD refer to the $(q + 1)$ st Adams-Moulton and q th order backward differentiation formulas, respectively. (The subscripts $y^{(q)}$ in the equation on p. 21 of Gear's report should read a_q .) This idea is extended to systems of differential equations by Skeel and Kong [1977], who suggest

$$\mathbf{l} := \mathbf{l}^{\text{AM}} - \gamma \mathbf{l}^{\text{BD}} hf_y$$

and

$$l_1 \mathbf{l} - l_0 h J_{n,\mu} := (\mathbf{l} - c hf_y)^2,$$

where γ and c are free parameters. In Section 4 it was shown that $\rho(\xi)$ and $\sigma(\xi)$ are linear transformations of \mathbf{l} , and so a linear combination of \mathbf{l} vectors corresponds to that same linear combination of the corresponding linear multistep formulas. Thus, under the assumption of constant hf_y the "blended" Nordsieck formula is equivalent to the same blend of linear multistep formulas.

Department of Computer Science
222 Digital Computer Laboratory
University of Illinois at Urbana-Champaign
Urbana, Illinois 61801

G. D. BYRNE & A. C. HINDMARSH (1975), "A polyalgorithm for the numerical solution of ordinary differential equations," *ACM Trans. Math. Software*, v. 1, pp. 71-96.

G. DAHLQUIST (1956), "Numerical integration of ordinary differential equations," *Math. Scand.*, v. 4, pp. 33-50.

- G. DAHLQUIST (1963), "A special stability problem for linear multistep methods," *BIT*, v. 3, pp. 27-43.
- G. DAHLQUIST (1975), *On Stability and Error Analysis for Stiff Non-linear Problems, Part I*, Report TRITA-NA-7508, Dept. of Computer Sci., Roy. Inst. of Technology, Stockholm.
- J. DESCLOUX (1963), *A Note on a Paper by A. Nordsieck*, Report #131, Dept. of Computer Sci., Univ. of Illinois at Urbana-Champaign.
- C. DILL & C. W. GEAR (1971), "A graphical search for stiffly stable methods for ordinary differential equations," *J. Assoc. Comput. Mach.*, v. 18, pp. 75-79.
- C. W. GEAR (1966), *The Numerical Integration of Ordinary Differential Equations of Various Orders*, Report #ANL-7126, Argonne National Laboratory, Argonne, Illinois.
- C. W. GEAR (1971a), *Numerical Initial Value Problems in Ordinary Differential Equations*, Prentice-Hall, Englewood Cliffs, N. J.
- C. W. GEAR (1971b), "Algorithm 407: DIFSUB for solution of ordinary differential equations," *Comm. ACM*, v. 14, pp. 185-190.
- R. D. GRIGORIEFF & J. SCHROLL (1977), *Über $A(\alpha)$ -stabile Verfahren hoher Konsistenzordnung*, Nr. 34, Fachbereich Mathematik (3), Technische Universität Berlin.
- G. K. GUPTA (1976), "Some new high-order multistep formulae for solving stiff equations," *Math. Comp.*, v. 30, pp. 417-432.
- G. K. GUPTA & C. S. WALLACE (1975), "Some new multistep methods for solving ordinary differential equations," *Math. Comp.*, v. 29, pp. 489-500.
- P. HENRICI (1962), *Discrete Variable Methods in Ordinary Differential Equations*, Wiley, New York.
- A. C. HINDMARSH (1974), *GEAR: Ordinary Differential Equation Solver*, UCID-3001, Rev. 3, Lawrence Livermore Laboratory, Univ. of California, Livermore, California.
- M. K. JAIN & V. K. SRIVASTAVA (1970), *High Order Stiffly Stable Methods for Ordinary Differential Equations*, Report #394, Dept. of Computer Sci., Univ. of Illinois at Urbana-Champaign.
- D. E. KNUTH (1968), *The Art of Computer Programming*, Vol. 1: *Fundamental Algorithms*, Addison-Wesley, Reading, Mass.
- A. K. KONG (1977), *A Search for Better Linear Multistep Methods for Stiff Problems*, Report R-77-899, Dept. of Computer Sci., Univ. of Illinois at Urbana-Champaign.
- A. NORDSIECK (1962), "On numerical integration of ordinary differential equations," *Math. Comp.*, v. 16, pp. 22-49.
- M. R. OSBORNE (1966), "On Nordsieck's method for the numerical solution of ordinary differential equations," *BIT*, v. 6, pp. 51-57.
- R. D. SKEEL (1973), *Convergence of Multivalued Methods for Solving Ordinary Differential Equations*, Report TR73-16, Dept. of Computing Sci., Univ. of Alberta, Edmonton.
- R. D. SKEEL & L. W. JACKSON (1977), "Consistency of Nordsieck methods," *SIAM J. Numer. Anal.*, v. 14, pp. 910-924.
- R. D. SKEEL & A. K. KONG (1977), "Blended linear multistep methods," *ACM Trans. Math. Software*, v. 3, pp. 326-345.
- C. S. WALLACE & G. K. GUPTA (1973), "General linear multistep methods to solve ordinary differential equations," *Austral. Comput. J.*, v. 5, pp. 62-69.