

Chapter 4

Linear Multistep Methods

4.1 Introduction

Runge-Kutta methods are a one-step method because they require only the value at one mesh point x_{n-1} to compute the value of the approximate solution at the next x_n . Notice, however, that the Runge-Kutta method does invoke a series of intermediate values during the computation. The idea of a multistep method is to use previously calculated values at a number of mesh points to aid the computation for later points much as those intermediate values used in a Runge-Kutta method.

Consider the initial value problem

$$\frac{dy}{dx} = f(x, y), \quad y(a) = y_0 \quad (4.1)$$

where $x, y, f \in R$. A method that makes use of the values of the dependent variable $y(x)$ and its derivative $f(x, y)$ at k different mesh points $x_{n-1}, x_{n-2}, \dots, x_{n-k}$ is called a *multistep* or a *k-step* method. More precisely, after approximations at x_{n-k}, \dots, x_{n-1} have been determined, we have values y_{n-k}, \dots, y_{n-1} and $hy'_{n-k}, \dots, hy'_{n-1}$ available. We want to use this information to determine y_n and hy'_n at x_n . We define

$$Y_n := [y_n, y_{n-1}, \dots, y_{n-k+1}; hy'_n, y'_{n-1}, \dots, hy'_{n-k+1}]^T. \quad (4.2)$$

The objective of a multistep method is to find a numerical approximation for Y_n from Y_{n-1} and to repeat this process.

Remark. Apparently the vector Y_{k-1} , whose components are called the *starting values*, must be known prior to the iteration. A common technique to generate the starting values is to use the Runge-Kutta method to compute y_1, \dots, y_{k-1} first and then set $hy'_i := hf_i$ for $i = 1, \dots, k-1$.

Suppose the starting value Y_{k-1} is known somehow, an iteration process may be as follows:

$$Y_{n,(0)} := BY_{n-1} \quad (4.3)$$

where

$$Y_{n,(0)} := [y_{n,(0)}, y_{n-1}, \dots, y_{n-k+1}; hy'_{n,(0)}, y'_{n-1}, \dots, hy'_{n-k+1}]^T, \quad (4.4)$$

$$B := \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_k & \beta_1 & \beta_2 & \dots & \beta_k \\ 1 & 0 & \dots & 0 & 0 & & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & & & \\ \vdots & & \ddots & & & & & \\ 0 & 0 & & 1 & 0 & & & \\ \gamma_1 & \gamma_2 & \dots & \gamma_k & \delta_1 & \delta_2 & \dots & \delta_k \\ 0 & 0 & & & 1 & 0 & & \\ & & & & 0 & 1 & & \\ \vdots & & & & & & \ddots & \\ 0 & 0 & & & 0 & & & 1 & 0 \end{bmatrix} \quad (4.5)$$

and $\alpha_i, \beta_i, \gamma_i, \delta_i$ are constants to be determined so that

$$y(x_n) \approx y_{n,(0)} := \sum_{i=1}^k (\alpha_i y_{n-i} + \beta_i hy'_{n-i}), \quad (4.6)$$

$$hy'(x_n) \approx hy'_{n,(0)} := \sum_{i=1}^k (\gamma_i y_{n-i} + \delta_i hy'_{n-i}). \quad (4.7)$$

Note that because $y_{n,(0)}$ is not exact,

$$G(Y_{n,(0)}) := -(Y_{n,(0)})_{k+1} + hf(x_n, (Y_{n,(0)})_1) \neq 0. \quad (4.8)$$

One may correct this error by a process like

$$Y_{n,(1)} := Y_{n,(0)} + \mathbf{c}G(Y_{n,(0)})$$

where \mathbf{c} is a suitable constant vector and may even repeat this process by

$$Y_{n,(m+1)} := Y_{n,(m)} + \mathbf{c}G(Y_{n,(m)}) \quad (4.9)$$

for $m = 1, 2, \dots$. The process may be repeated for a fix number of iterations or until $|G| \ll 1$. The scheme defined in (4.3) usually is known as the *predictor* and the scheme in (4.9) is known as the *corrector*.

Definition 4.1.1 Suppose that $\gamma_i = \delta_i \equiv 0$ for $1 \leq i \leq k$ in (4.3) and that $c_i = 0$ in (4.9) expect the $(k+1)$ -th position which is 1. The resulting scheme

$$y_n := \sum_{i=1}^k (\alpha_i y_{n-i} + \beta_i hy'_{n-i}) \quad (4.10)$$

$$hy'_n := hf_n \quad (4.11)$$

is called an *explicit k-step method*.

Remark. Note that the additional iteration in the corrector process has no effect.

Remark. Suppose $\mathbf{c} = [\beta_0^*, 0, \dots, 0; 1, 0, \dots]^T$ in (4.9). Then

$$\begin{aligned} y_{n,(1)} &= y_{n,(0)} + \beta_0^*(hf_{n,(0)} - hy'_{n,(0)}) \\ &= \sum_{i=1}^k [(\alpha_i - \beta_0^*)\gamma_i y_{n-i} + (\beta_i - \beta_0^*\delta_i)hy'_{n-i}] + \beta_0^*hf_{n,(0)}. \end{aligned}$$

If we define

$$\begin{aligned} \alpha_i^* &:= \alpha_i - \beta_0^*\gamma_i, \\ \beta_i^* &:= \beta_i - \beta_0^*\delta_i, \end{aligned}$$

then we have a new scheme

$$y_{n,(m+1)} = \sum_{i=1}^k (\alpha_i^* y_{n-i} + \beta_i^* hy'_{n-i}) + \beta_0^* hf_{n,(m)}, \quad (4.12)$$

$$hy'_{n,(m+1)} = hf_{n,(m)}. \quad (4.13)$$

Definition 4.1.2 *In the sense of convergence, we may rewrite the scheme as*

$$y_n = \sum_{i=1}^k (\alpha_i^* y_{n-i} + \beta_i^* hy'_{n-i}) + \beta_0^* hf(x_n, y_n), \quad (4.14)$$

$$hy'_n = hf_n \quad (4.15)$$

which is called an implicit k -step method.

Remark. Note that (4.14) is a nonlinear equation that must be solved for y_n .

4.2 Derivation of Linear Multistep Methods

There are three ways to derive linear multistep methods. We shall pay our attention to the derivation of (4.10) and (4.12) in particular.

Through the Taylor expansion: We have already seen how the Euler method can be derived. Consider now the expansion:

$$y(x_n + h) = y(x_n) + hy^{(1)}(x_n) + \frac{h^2}{2}y^{(2)}(x_n) + O(h^3),$$

$$y(x_n - h) = y(x_n) - hy^{(1)}(x_n) + \frac{h^2}{2}y^{(2)}(x_n) + O(h^3).$$

Clearly we have

$$y(x_n + h) - y(x_n - h) = 2hy^{(1)}(x_n) + O(h^3) = 2hf(x_n, y(x_n)) + O(h^3). \quad (4.16)$$

Suppose $y_{n+1} \approx y(x_n + h)$, $y_{n-1} \approx y(x_{n-1})$ and $y_n \approx y(x_n)$. Then we have arrived at an explicit 2-step method,

$$y_{n+1} = y_{n-1} + 2hf_n, \quad (4.17)$$

that is also known as the *midpoint rule* and has local truncation error $O(h^3)$.

Suppose we are looking for a general linear 1-step method, i.e.,

$$y_n = \alpha_1 y_{n-1} + h\beta_0 f_n + h\beta_1 f_{n-1} \quad (4.18)$$

where $\alpha_1, \beta_0, \beta_1$ are to be determined. We consider the equation

$$y(x_{n-1} + h) = \alpha_1 y(x_{n-1}) + h\beta_0 y^{(1)}(x_{n-1} + h) + h\beta_1 y^{(1)}(x_{n-1}). \quad (4.19)$$

Upon expanding (4.19) about x_{n-1} and collecting like powers of h , we obtain

$$(1 - \alpha_1)y(x_{n-1}) + h(1 - \beta_0 - \beta_1)y^{(1)}(x_{n-1}) + h^2\left(\frac{1}{2} - \beta_0\right)y^{(2)}(x_{n-1}) + O(h^3) = 0. \quad (4.20)$$

Setting the coefficients zero, we arrive at a 1-step scheme

$$y_n = y_{n-1} + \frac{h}{2}(f_n + f_{n-1}) \quad (4.21)$$

that, also known as the Trapezoidal rule, is implicit and has local truncation error $O(h^3)$.

Example. It can be shown that the most accurate linear 2-step method is the so called Simpson's rule:

$$y_n = y_{n-2} + \frac{h}{3}(f_n + 4f_{n-1} + f_{n-2}) \quad (4.22)$$

that has local truncation error $O(h^5)$. (This is a homework problem.)

Through numerical integration: Clearly we should have

$$y(x_n) - y(x_{n-1}) = \int_{x_{n-1}}^{x_n} y'(x)dx = \int_{x_{n-1}}^{x_n} f(x, y(x))dx. \quad (4.23)$$

We approximate the integrand by a polynomial of degree zero, say by the constant $f(x_{n-1}, y(x_{n-1}))$. This leads to

$$y(x_n) - y(x_{n-1}) \approx hf(x_{n-1}, y(x_{n-1}))$$

and hence suggests the scheme

$$y_n = y_{n-1} + hf(x_n, y_n)$$

that is the Euler method. If the integrand is approximated by a polynomial of degree one, say by the segment connecting the two endpoints, then we obtain the Trapezoidal rule (4.21).

Similarly, consider the integral equation

$$y(x_n) - y(x_{n-2}) = \int_{x_{n-2}}^{x_n} f(x, y(x)) dx. \quad (4.24)$$

If we assume $f \approx f(x_{n-1}, y(x_{n-1}))$, then we have the midpoint rule (4.17). If we use a quadratic polynomial to approximate the integrand, then we obtain the Simpson's rule (4.22).

In general, there is a correspondence between linear multistep methods and the so called Newton-Cotes integration formulas. Specifically, if one approximates $f(x, y(x))$ in (4.23) by an interpolation polynomial with data (x_{n-i}, f_{n-i}) for $i = 1, \dots, k$, the resulting numerical scheme is called the *Adams-Bashforth (AB)* method. If one interpolates f with one additional data (x_n, f_n) , the resulting implicit scheme is called the *Adams-Molton (AS)* method.

Example. Suppose the function f in (4.23) is approximated by a quadratic polynomial that interpolates f at x_{n-2}, x_{n-1} and x_n . It can be shown that the following scheme is arrived:

$$y_n = y_{n-1} + \frac{h}{12}(5f_n + 8f_{n-1} - f_{n-2}).$$

Remark. Derivation through numerical integration can only lead to methods of the form

$$y_n = y_{n-i} + h \sum_{j=0}^k \beta_j f_{n-j}. \quad (4.25)$$

The subclass of methods when $i = 1$ are usually referred to as *Adams methods*. The methods when $i = 2$ are referred to as *Nyström methods*.

Through interpolation: We only sketch the idea here. Let the exact solution $y(x)$ be interpolated by a Hermite polynomial over the interval $[x_{n-1}, x_n]$, i.e.,

$$\begin{aligned} p(x_n) &= y_n, p(x_{n-1}) &= y_{n-1}; \\ p'(x_n) &= f_n, p'(x_{n-1}) &= f_{n-1}. \end{aligned}$$

Generally, $p(x)$ should be a cubic polynomial. If, instead, we insist that $p(x) = ax^2 + bx + c$ so that there are only three free parameters. Then to satisfy the four interpolating conditions, there must be a relationship among the four quantities y_n, y_{n-1}, f_n and f_{n-1} . Indeed, we must have

$$y_n = y_{n-1} + \frac{h}{2}(f_{n-1} + f_n)$$

which is the Trapezoidal rule.

4.3 Convergence, Consistency and Stability

Corresponding to a given a linear multistep method

$$\sum_{i=0}^k (\alpha_i y_{n-i} + h\beta_i f_{n-i}) = 0, \quad (4.26)$$

we define a linear difference operator

$$\mathcal{L}_h(y(x)) := \sum_{i=0}^k (\alpha_i y(x - ih) + h\beta_i y'(x - ih)) \quad (4.27)$$

for any C^1 function $y(x)$. Upon expanding the right-hand side of (4.27) about x , we have

$$\mathcal{L}_h(y(x)) = \sum_{q=0}^{r+1} c_q h^q y^{(q)} + O(h^{r+2})$$

where

$$c_q := \begin{cases} \sum_{i=0}^k \alpha_i & \text{if } q = 0, \\ \sum_{i=0}^k \left[\frac{(-i)^q}{q!} \alpha_i + \frac{(-i)^{q-1}}{(q-1)!} \beta_i \right] & \text{if } q > 0. \end{cases} \quad (4.28)$$

Definition 4.3.1 *The linear difference operator \mathcal{L} and the associated linear multistep method is said to be of order r if $c_q = 0$ for all $q \leq r$ and $c_{r+1} \neq 0$.*

Remark. The equations $c_q = 0$ for all $q \leq r$ determine the coefficients for an r -th order method.

Definition 4.3.2 *The local truncation error at x_n of a linear k -step method is defined to be the quantity $\mathcal{L}_h(y(x_n))$ where $y(x)$ is the exact solution.*

Remark. The above definition is motivated by what follows: By (4.27), we have

$$\sum_{i=0}^k \alpha_i y(x_n - ih) = \mathcal{L}_h(y(x_n)) - h \sum_{i=0}^k \beta_i y'(x_n - ih). \quad (4.29)$$

On the other hand, by (4.26), we have

$$\sum_{i=1}^k \alpha_i y_{n-i} = -h \sum_{i=0}^k \beta_i f_{n-i}. \quad (4.30)$$

Suppose $y_{n-i} = y(x_n - ih)$ for $i = 1, \dots, k$ (This is called the *local assumption*). Subtracting (4.30) from (4.29) yields

$$\alpha_0(y(x_n) - y_n) = \mathcal{L}_h(y(x_n)) + h\beta_0(f(x_n, y_n) - f(x_n, y(x_n))). \quad (4.31)$$

If the method (4.26) is explicit, i.e., $\beta_0 = 0$, then $\mathcal{L}_h(y(x_n))$ does measure the local truncation error (Most schemes have $\alpha_0 = 1$). If the method (4.26) is implicit, then the mean value theorem implies that

$$\alpha_0(y(x_n) - y_n) = \mathcal{L}_h(y(x_n)) + h\beta_0 \frac{\partial f(x_n, \xi_n)}{\partial y} (y_n - y(x_n))$$

for some ξ_n lying between y_n and $y(x_n)$. It follows that

$$\mathcal{L}_h(y(x_n)) = (\alpha_0 + h\beta_0 \frac{\partial f(x_n, \xi_n)}{\partial y}) (y(x_n) - y_n). \quad (4.32)$$

Thus $\mathcal{L}_h(y(x_n))$ still measures the local truncation error.

Definition 4.3.3 Given a linear multistep method (4.26), the polynomial

$$\rho(\xi) := \sum_{i=0}^k \alpha_i \xi^{k-i} \quad (4.33)$$

$$\sigma(\xi) := \sum_{i=0}^k \beta_i \xi^{k-i} \quad (4.34)$$

are called the first and the second characteristic polynomials, respectively.

Definition 4.3.4 A linear k -step method is said to be consistent if and only if it has order $|r| \geq 1$, i.e., $c_0 = c_1 = 0$.

Remark. The linear multistep method (4.26) is consistent if and only if $\rho(1) = 0$ and $\rho'(1) + \sigma(1) = 0$.

Definition 4.3.5 The linear multistep method (4.26) is said to be zero stable if all solutions of the homogeneous linear difference equation

$$\sum_{i=1}^k \alpha_i y_{n-i} = 0 \quad (4.35)$$

are bounded for all n .

Consider any linear difference equation

$$\sum_{i=0}^k \alpha_i y_{n-i} = \phi_n, \quad n = n_0, \dots \quad (4.36)$$

where $\alpha_0, \dots, \alpha_k$ are constants independent of n , $\alpha_0 \neq 0$, $\alpha_k \neq 0$, and $\{\phi_n\}, n = n_0, \dots$, is a known sequence. A solution of this difference equation is a sequence of values $\{y_n\}, n = n_0, \dots$, that satisfies (4.36) identically for all $n = n_0, \dots$. The general solution (4.36) can be written as

$$y_n = \hat{y}_n + \psi_n \quad (4.37)$$

where \hat{y}_n is the general solution of the associated homogeneous equation (4.35) and ψ_n is a particular solution of (4.36). To determine \hat{y}_n , we try

$$\hat{y}_n = r^n. \quad (4.38)$$

Substituting (4.38) into (4.35) yields

$$\sum_{i=0}^k \alpha_i r^{n-i} = 0.$$

Dividing by r^{n-k} yields

$$\rho(r) = \sum_{i=0}^k \alpha_i r^{k-i} = 0. \quad (4.39)$$

Let the k roots of $\rho(r)$ be denoted by r_1, \dots, r_k . We calculated \hat{y}_n as follows:

Case 1: Suppose all roots are distinct. Then all solutions $r_j^n, j = 1, \dots, k, n = n_0, \dots$ are linearly independent in the sense that $\sum_{j=1}^k a_j r_j^n = 0$ for all n only if $a_j = 0$ for all j . Hence the general solution \hat{y}_n is given by

$$\hat{y}_n = \sum_{j=1}^k d_j r_j^n, \quad n = n_0, \dots \quad (4.40)$$

where d_j are arbitrary constants.

Case 2: If some of all roots are repeated, e.g., suppose $r_1 = r_2 = r_3$ and the remaining roots are distinct, then the solutions $r_1^n, nr_1^n, n^2r_1^n, r_4^n, \dots, r_k^n$ are linearly independent. The general solution \hat{y}_n is given by

$$\hat{y}_n = d_1 r_1^n + d_2 n r_1^n + d_3 n^2 r_1^n + d_4 r_4^n + \dots + d_k r_k^n. \quad (4.41)$$

Remark. Observe that $\{\hat{y}_n\}$ is bounded if and only if all roots of $\rho(r) = 0$ have modules $|r_j| \leq 1$ and those with modules 1 are simple.

Definition 4.3.6 *The linear multistep method (4.26) is said to converge if for all initial value problems*

$$\begin{aligned} y' &= f(x, y) \\ y(a) &= \eta, \end{aligned}$$

we have

$$\lim_{h \rightarrow 0, nh = x - a} y_n = y(x_n) = y(x)$$

for all $x \in [a, b]$ and for all solution $\{y_n\}$ of (4.26) with starting value $y_\mu = \eta_\mu(h)$, $\mu = 0, 1, \dots, k - 1$, satisfying $\lim_{h \rightarrow 0} \eta_\mu(h) = \eta$ for all μ .

Theorem 4.3.1 *The necessary and sufficient conditions for a linear multistep method to converge are that it be consistent and zero-stability.*

Proof. Refer to Peter Henrici, Discrete Variable Methods in Ordinary Differential Equations, John-Wiley, 1962. \square

Example. The midpoint rule (4.17) is consistent and zero stable. Applying it to the problem

$$\begin{aligned} y' &= -y \\ y(0) &= 1 \end{aligned}$$

gives rise to the difference equation

$$y_n + 2hy_{n-1} - y_{n-2} = 0. \quad (4.42)$$

The solution is of the form

$$y_n = c_1 r_1^n + c_2 r_2^n$$

where $r_1 = -h + \sqrt{1+h^2}$ and $r_2 = -h - \sqrt{1+h^2}$. Initial condition implies $c_1 + c_2 = 1$. We may express c_1 and c_2 in terms of the second starting value y_1 , i.e.,

$$\begin{aligned} c_1 &= \frac{r_2 - y_1}{r_2 - r_1} = \frac{\sqrt{1+h^2} + h + y_1}{2\sqrt{1+h^2}} \\ c_2 &= \frac{r_1 - y_1}{r_1 - r_2} = \frac{\sqrt{1+h^2} - h - y_1}{2\sqrt{1+h^2}}. \end{aligned}$$

Note that $|r_1| < 1$ and $|r_2| > 1$. Thus $|y_n|$ is unbounded as $n \rightarrow \infty$ unless y_1 is specially selected. This phenomenon is called *numerical instability* or *weak instability*. Note that the observation is not a contradiction with the convergence result because the step size h is fixed in this example.

We now analyze how the global errors are propagated. Recall that

$$\sum_{i=0}^k \alpha_i y_{n-i} = -h \sum_{i=0}^k \beta_i f_{n-i} \quad (4.43)$$

$$\sum_{i=0}^k \alpha_i y(x_{n-i}) = -h \sum_{i=0}^k \beta_i f(x_{n-i}, y(x_{n-i})) + \mathcal{L}_h(y(x_n)). \quad (4.44)$$

Now suppose that due to round-off errors and floating-point arithmetics, we have

$$\sum_{i=0}^k \alpha_i \hat{y}_{n-i} = -h \sum_{i=0}^k \beta_i f(x_{n-i}, \hat{y}_{n-i}) + \mathcal{R}_n. \quad (4.45)$$

Define

$$e_n := y(x_n) - \hat{y}_n.$$

Then we have

$$\sum_{i=0}^k \alpha_i e_{n-i} = -h \sum_{i=0}^k \beta_i (f(x_{n-i}, y(x_{n-i})) - f(x_{n-i}, \hat{y}_{n-i})) + \Phi_n \quad (4.46)$$

with

$$\Phi_n := \mathcal{L}_h(y(x_n)) - \mathcal{R}_n.$$

As usual, we apply (4.46) to the model problem $y' = \lambda y$. We also assume that $\Phi_n \equiv \Phi$ for all n (This is reasonable since Φ_n is supposed to be small anyway.). Then e_n is given by

$$e_n = \sum_{s=1}^k d_s r_s^n + \frac{\Phi}{\lambda h \sum_{i=0}^k \beta_i} \quad (4.47)$$

where d_s are constants depending upon the starting values and r_s are roots of the polynomial

$$\Pi(r, \bar{h}) := \sum_{i=0}^k (\alpha_i + \bar{h} \beta_i) r^{k-i} \quad (4.48)$$

with $\bar{h} := \lambda h$. In (4.47) we have assumed that $\Pi(r, \bar{h})$ has distinct roots. The case of multiple roots can be modified easily.

Definition 4.3.7 *The polynomial $\Pi(r, \bar{h}) = \rho(r) + \bar{h}\sigma(r)$ is called the stability polynomial.*

Definition 4.3.8 *A linear multistep method is said to be absolutely stable for a given \bar{h} if all roots r_s of $\Pi(r, \bar{h})$ satisfy $|r_s| < 1$, and to be absolutely unstable otherwise. The region absolute stability consists of all \bar{h} in the complex plan for which the method is absolutely stable.*

Theorem 4.3.2 *Every consistent zero-stable linear multistep method is absolutely unstable for small positive \bar{h} .*

Proof. Observe first that $\Pi(r, 0) = \rho(r)$ which, by zero-stability, has $r = 1$ as a simple root. Let r_1 be the root of $\Pi(r, \bar{h})$ that tends to 1 as $\bar{h} \rightarrow 0$. Suppose the method being considered is of order r so that $\mathcal{L}_h(y(x)) = O(h^{r+1})$. Now observe

$$\begin{aligned} \mathcal{L}(e^{\lambda x}) &= \sum_{i=0}^k \alpha_i e^{\lambda(x-ih)} + h\lambda \sum_{i=0}^k \beta_i e^{\lambda(x-ih)} \\ &= e^{\lambda(x-kh)} \left\{ \sum_{i=0}^k \alpha_i e^{(k-j)\bar{h}} + \bar{h} \sum_{i=0}^k \beta_i e^{(k-j)\bar{h}} \right\} \\ &= e^{\lambda(x-kh)} \Pi(e^{\bar{h}}, \bar{h}). \end{aligned}$$

Suppose the roots of $\Pi(r, \bar{h})$ are r_1, \dots, r_k so that

$$\Pi(r, \bar{h}) = (\alpha_0 + \bar{h}\beta_0)(r - r_1) \cdots (r - r_k).$$

It follows that

$$(\alpha_0 + \bar{h}\beta_0)(e^{\bar{h}} - r_1) \cdots (e^{\bar{h}} - r_k) = O(\bar{h}^{r+1}).$$

As $\bar{h} \rightarrow 0$, $e^{\bar{h}} \rightarrow 1$. Thus r_s must be significantly different from $e^{\bar{h}}$. Therefore,

$$e^{\bar{h}} - r_1 = O(\bar{h}^{r+1}).$$

That is,

$$r_1 = 1 + \bar{h} + O(\bar{h}^2). \quad (4.49)$$

This shows the unstability. \square

Remark. Motivated by the above proof, we observe that

$$r_1^n = e^{n\bar{h}} + O(\bar{h}^{r+1}) \approx e^{\lambda(x_n - a)}.$$

Thus if the error grows at a rate similar to the rate at which the solution grows, such a state of affairs generally is acceptable. We therefore have the following definition.

Definition 4.3.9 A linear multistep method is said to be relatively stable for a given \bar{h} if the roots r_s of the polynomial $\Pi(r, \bar{h})$ satisfy $|r_s| < |r_1|$ for all $s = 2, \dots, k$.

Example. Consider the Simpson's rule:

$$y_n = y_{n-2} + \frac{h}{3}(f_n + 4f_{n-1} + f_{n-2}).$$

It follows that

$$\begin{aligned}\rho(r) &= r^2 - 1, \\ \sigma(r) &= -\frac{1}{3}(r^2 + 4r + 1), \\ \Pi(r, \bar{h}) &= \left(1 - \frac{\bar{h}}{3}\right)r^2 - \frac{4\bar{h}}{3}r - \left(1 + \frac{\bar{h}}{3}\right).\end{aligned}$$

By (4.49), we know one root already, i.e.,

$$r_1 = 1 + \bar{h} + O(\bar{h}^2).$$

Assume the other root is of the form

$$r_2 = -1 + \gamma\bar{h} + O(\bar{h}^2).$$

Then we must have

$$\begin{aligned}\left(1 - \frac{\bar{h}}{3}\right)(-1 + \gamma\bar{h})^2 - \frac{4\bar{h}}{3}(-1 + \gamma\bar{h}) - \left(1 + \frac{\bar{h}}{3}\right) &= O(\bar{h}^2) \\ \left(1 - \frac{\bar{h}}{3}\right)(1 - 2\gamma\bar{h}) + \frac{4\bar{h}}{3}\bar{h} - \left(1 + \frac{\bar{h}}{3}\right) &= O(\bar{h}^2).\end{aligned}$$

It follows that

$$1 - 2\gamma\bar{h} \approx \frac{1 - \bar{h}}{1 - \frac{\bar{h}}{3}} \approx (1 - \bar{h})\left(1 + \frac{\bar{h}}{3} + O(\bar{h}^2)\right).$$

By comparing coefficients, we conclude that

$$r_2 = -1 + \frac{1}{3}\bar{h} + O(\bar{h}^2).$$

Note that if $\bar{h} > 0$, then $r_1 > 1$ and $r_2 > -1$. Note also that if $\bar{h} < 0$, then $r_1 < 1$ and $r_2 < -1$. In other words, Simpson's rule is nowhere absolutely stable. On the other hand, by solving $\Pi(r, \bar{h}) = 0$ exactly, it can be shown that the interval of relative stability for Simpson's rule is $(0, \infty)$. Hence if $\frac{\partial f}{\partial y} < 0$, it is not advisable to use this method either.

Remark. It is important not to draw sharp conclusions from the above weak stability theory for general nonlinear problems. The theory only gives a

rough guide on the maximal allowable step size h . In practice, if h is taken to be too close to that maximal step size, the method can perform poorly.

Remark. Concerning the region of absolute stability for Adams methods, it is observed that the region for the implicit Adams-Moulton method is larger than that of the explicit Adams-Bashforth method of the same order.

A k -step method has coefficients $\alpha_0 = 1, \alpha_1, \dots, \alpha_k, \beta_0, \beta_1, \dots, \beta_k$. Totally there are $2k + 1$ dimensions of freedom to choose these coefficients. It appears that we may select these numbers to satisfy the $2k + 1$ equations $c_0 = c_1 = \dots = c_{2k+1} = 0$ defined in (4.28). This indeed is possible. Thus the maximum attainable order of a k -step method is $2k$ for an implicit method and $2k - 1$ for an explicit method. This kind of choice, however, is useless because of the following theorems:

Theorem 4.3.3 *Any k -step linear method of order $r \geq k + 2$ is not zero-stable. More precisely, the highest zero-stable attainable order $p^*(k)$ is*

$$p^*(k) = \begin{cases} k + 2 & \text{if } k \text{ is even;} \\ k + 1 & \text{if } k \text{ is odd.} \end{cases} \quad (4.50)$$

Proof. See Peter Henrici's book for the proof. \square

Remark. Simpson's rule is the only linear multistep method that attains its optimal order. As we have seen, it suffers from the following property.

Theorem 4.3.4 *Any optimal linear multistep method has no interval of absolute stability.*

4.4 Predictor-Corrector Methods

We first look at some facts about the Adams-Bashforth (AB) methods and the Adams-Moulton (AM) methods:

	AB				AM			
k	1	2	3	4	1	2	3	4
r	1	2	3	4	2	3	4	5
c_{r+1}	$\frac{1}{2}$	$\frac{5}{12}$	$\frac{3}{8}$	$\frac{251}{720}$	$-\frac{1}{12}$	$-\frac{1}{24}$	$-\frac{19}{720}$	$-\frac{3}{160}$
α	-2	-1	$-\frac{6}{11}$	$-\frac{3}{10}$	$-\infty$	-6	-3	$-\frac{90}{49}$

We immediately see several advantages of implicit methods:

1. Higher order of accuracy for a given step number k .
2. Smaller error constant c_{r+1} for a given order r .
3. Larger interval $(\alpha, 0)$ of absolute stability.

On the other hand, implicit methods always involve a nonlinear system of equations that has to be solved iteratively. We are thus motivated to consider the so called *Predictor-Corrector (PC)* methods. In a PC method we first use a separate explicit method to estimate y_n . The estimate is used as the initial guess $y_n^{(0)}$ for the implicit method where iteration is taking place to find the acceptable y_n . There are two ways to proceed the PC method:

Correct to convergence: Preassign a tolerance $\epsilon > 0$, and the iteration stops only when $|y_n^{(s+1)} - y_n^{(s)}| < \epsilon$ (or $\frac{|y_n^{(s+1)} - y_n^{(s)}|}{|y_n^{(s+1)}|} < \epsilon$). In this case, we have no knowledge in advance how many function evaluations are needed. It can be shown that the local truncation error and the weak stability characteristics of such a method are precisely those of the corrector along. The properties of the predictor are not important.

$P(EC)^m$ or $P(EC)^m E$ mode: Let P stand for the predictor, E for the function evaluation of f and C for the corrector. Suppose also

$$P : \sum_{i=0}^k \alpha_i y_{n-i} + h \sum_{i=1}^k \beta_i f_{n-i} = 0, \quad (4.51)$$

$$C : \sum_{i=0}^{k^*} \alpha_i^* y_{n-i} + h \sum_{i=0}^{k^*} \beta_i^* f_{n-i} = 0. \quad (4.52)$$

Definition 4.4.1 *By a $P(EC)^m E$ mode, we mean the following scheme:*

$$\begin{aligned} & \alpha_0 y_n^{(0)} + \sum_{i=1}^k (\alpha_i y_{n-i}^{(m)} + h \beta_i f_{n-i}^{(m)}) = 0, \\ & \text{For } s = 0, 1, \dots, m-1, \text{ do:} \\ & \left. \begin{aligned} & f_n^{(s)} = f(x_n, y_n^{(s)}) \\ & \alpha_0^* y_n^{(s+1)} + \sum_{i=1}^k (\alpha_i^* y_{n-i}^{(m)} + h \beta_i^* f_{n-i}^{(m)}) + h \beta_0^* f_n^{(s)} = 0, \\ & f_n^{(m)} = f(x_n, y_n^{(m)}). \end{aligned} \right\} \quad (4.53) \end{aligned}$$

Definition 4.4.2 *By a $P(EC)^m$ mode, we mean the following scheme:*

$$\begin{aligned} & \alpha_0 y_n^{(0)} + \sum_{i=1}^k (\alpha_i y_{n-i}^{(m)} + h \beta_i f_{n-i}^{(m-1)}) = 0, \\ & \text{For } s = 0, 1, \dots, m-1, \text{ do:} \\ & \left. \begin{aligned} & f_n^{(s)} = f(x_n, y_n^{(s)}), \\ & \alpha_0^* y_n^{(s+1)} + \sum_{i=1}^k (\alpha_i^* y_{n-i}^{(m)} + h \beta_i^* f_{n-i}^{(m-1)}) + h \beta_0^* f_n^{(s)} = 0. \end{aligned} \right\} \quad (4.54) \end{aligned}$$

We now study the local truncation error and the stability theory of the *PC* methods. Let \mathcal{L} and \mathcal{L}^* represent, respectively, the linear difference operator of P and C . Suppose also

$$\mathcal{L}_h(y(x)) = c_{r+1} h^{r+1} y^{(r+1)}(x) + O(h^{r+2}), \quad (4.55)$$

$$\mathcal{L}_h^* = c_{r^*+1}^* h^{r^*+1} y^{(r^*+1)}(x) + O(h^{r^*+2}). \quad (4.56)$$

For convenience, suppose $\alpha_0 = \alpha_0^* = 1$. Under the local assumption, we know

$$\mathcal{L}_h(y(x)) = y(x_n) - y_n^{(0)}, \quad (4.57)$$

and

$$\begin{aligned} \mathcal{L}_h^*(y(x)) = & y(x_n) + \sum_{i=1}^{k^*} (\alpha_i^* y(x_{n-i}) \\ & + h\beta_i^* f(x_{n-i}, y(x_{n-i}))) + h\beta_0^* f(x_n, y(x_n)). \end{aligned} \quad (4.58)$$

Furthermore, we know

$$y_n^{(s+1)} + \sum_{i=1}^{k^*} (\alpha_i^* y_{n-i} + h\beta_i^* f_{n-i}^{(m-t)}) + h\beta_0^* f(x_n, y_n^{(s)}) = 0 \quad (4.59)$$

for $s = 0, 1, \dots, m-1$ where $t = 0$ or 1 depending on the mode being considered. It follows that

$$\begin{aligned} y(x_n) - y_n^{(s+1)} &= \mathcal{L}_h^* - h\beta_0^* (f(x_n, y(x_n)) - f(x_n, y_n^{(s)})) \\ &= \mathcal{L}_h^*(y(x)) - h\beta_0^* \frac{\partial f}{\partial y}(x_n, \xi_n)(y(x_n) - y_n^{(s)}) \end{aligned} \quad (4.60)$$

where ξ_n lies between $y(x_n)$ and $y_n^{(s)}$. Depending upon the orders r and r^* involved, we analyze the following cases:

When $r \geq r^*$: The predictor has higher order than the corrector. We have

$$\begin{aligned} y(x_n) - y_n^{(1)} &= \mathcal{L}_h^* + O(h^{r^*+2}) = O(h^{r^*+1}) \\ y(x_n) - y_n^{(s+1)} &= \mathcal{L}_h^* + O(h^{r^*+2}) = O(h^{r^*+1}) \text{ for } s \geq 1. \end{aligned}$$

In this case, therefore, the principal local truncation error of a *PC* method when $m \geq 1$ is that of the corrector alone.

When $r = r^* - 1$: We have

$$\begin{aligned} y(x_n) - y_n^{(1)} &= \left(c_{r^*+1}^* y^{(r^*+1)}(x_n) - \beta_0^* \frac{\partial f}{\partial y} C_{r^*} y^{(r^*-1)}(x_n) \right) h^{r^*+1} \\ &\quad + O(h^{r^*+2}) = O(h^{r^*+1}) \\ y(x_n) - y_n^{(s+1)} &= \mathcal{L}_h^* + O(h^{r^*+2}) = O(h^{r^*+1}) \text{ for } s \geq 1. \end{aligned}$$

Thus, for $m = 1$ the principal local truncation error of the *PC* method, is of the same order as that of the corrector, but with different constants. For $m \geq 2$, the principal local truncation error is that of the corrector alone.

When $r = r^* - 2$: We have

$$\begin{aligned} y(x_n) - y_n^{(1)} &= -\beta_0^* \frac{\partial f}{\partial y} C_{r^*-1} y^{(r^*-1)}(x_n) h^{r^*} + O(h^{r^*+1}), \\ y(x_n) - y_n^{(2)} &= \left(c_{r^*+1}^* y^{(r^*+1)}(x_n) - (\beta_0^* \frac{\partial f}{\partial y})^2 C_{r^*-1} y^{(r^*-1)}(x_n) \right) h^{r^*+1} \\ &\quad + O(h^{r^*+2}) = O(h^{r^*+1}) \\ y(x_n) - y_n^{(s+1)} &= \mathcal{L}_h^* + O(h^{r^*+2}) \text{ for } s \geq 2. \end{aligned}$$

Therefore for $m \geq 3$, the principal local truncation error is that of the corrector alone.

The trend is quite clear:

1. If $r \geq r^*$, then the principal local truncation error is that of the corrector alone.
2. If $r = r^* - q$ where $0 < q \leq r^*$, the the principal local truncation error is
 - (a) That of the corrector alone if $m \geq q + 1$.
 - (b) Of the same order as the corrector but with different constants when $m = q$.
 - (c) Of the form $Kh^{r^*-q+m+1} + O(h^{r^*-q+m+2})$ when $m < q$.

Based on the above analysis, we make the following conclusions:

- In the mode of correction to convergence, the principal local truncation error is that of the corrector, regardless what the predictor is.
- The case $r > r^*$ is out of question.
- The case $r = r^* - m$ is the best bet, especially when $m = 1$.
- The case $r = r^*$ together with the so called Milne's device is also advisable.

Suppose now that in a *PECE* mode, due to round-off errors, we have

$$\hat{y}_n^{(0)} + \sum_{i=1}^k (\alpha_i \hat{y}_{n-i}^{(1)} + h\beta_i f(x_{n-i}, \hat{y}_{n-i}^{(1)})) = R_n \quad (4.61)$$

and

$$\hat{y}_n^{(1)} + \sum_{i=1}^{k^*} (\alpha_i^* \hat{y}_{n-i}^{(1)} + h\beta_i^* f(x_{n-i}, \hat{y}_{n-i}^{(1)})) + h\beta_0^* f(x_n, \hat{y}_n^{(0)}) = R_n^*. \quad (4.62)$$

where R_n and R_n^* are the local round-off errors. We also know by definitions that

$$y(x_n) + \sum_{i=1}^k (\alpha_i y(x_{n-i}) + h\beta_i f(x_{n-i}, y(x_{n-i}))) = \mathcal{L}_h(y(x_n)), \quad (4.63)$$

and

$$y(x_n) + \sum_{i=1}^{k^*} (\alpha_i^* y(x_{n-i}) + h\beta_i^* f(x_{n-i}, y(x_{n-i}))) + h\beta_0^* f(x_n, y(x_n)) = \mathcal{L}_h^*(y(x_n)). \quad (4.64)$$

Define

$$\begin{aligned} e_n^{(0)} &:= y(x_n) - \hat{y}_n^{(0)} \\ e_n^{(1)} &:= y(x_n) - \hat{y}_n^{(1)}. \end{aligned}$$

When applying this mode to the test problem $y' = \lambda y$ and assuming that $\mathcal{L}_h(y(x_n)) - R_n$ and $\mathcal{L}_h^*(y(x_n)) - R_n^*$ are constants, we obtain

$$e_n^{(0)} + \sum_{i=1}^k (\alpha_i + \bar{h}\beta_i) e_{n-i}^{(1)} = C_1, \quad (4.65)$$

$$e_n^{(1)} + \sum_{i=1}^{k^*} (\alpha_i^* + \bar{h}\beta_i^*) e_{n-i}^{(1)} + \bar{h}\beta_0^* e_n^{(0)} = C_2. \quad (4.66)$$

It follows that

$$e_n^{(1)} + \sum_{i=1}^{k^*} (\alpha_i^* + \bar{h}\beta_i^*) e_{n-i}^{(1)} + \bar{h}\beta_0^* \left\{ - \sum_{i=1}^k (\alpha_i + \bar{h}\beta_i) e_{n-i}^{(1)} \right\} = C_3. \quad (4.67)$$

Adding $\bar{h}\beta_0^* e_n^{(1)}$ to the first sum and then subtracting it from the second sum, we may rewrite (assuming $\alpha_0 = 1, \beta_0 = 1$) (4.67) as

$$\sum_{i=0}^{k^*} (\alpha_i^* + \bar{h}\beta_i^*) e_{n-i}^{(1)} - \bar{h}\beta_0^* \sum_{i=0}^k (\alpha_i + \bar{h}\beta_i) e_{n-i}^{(1)} = C_3. \quad (4.68)$$

Therefore, the stability polynomial of a *PECE* method is given by

$$\Pi_{PECE}(r, \bar{h}) := \rho^*(r) + \bar{h}\sigma^*(r) - \bar{h}\beta_0^*(\rho(r) + \bar{h}\sigma(r)) = 0. \quad (4.69)$$

The analysis can be extended to give the following stability polynomial for the general $P(EC)^m E$ method:

$$\Pi_{P(EC)^m E}(r, \bar{h}) := \rho^*(r) + \bar{h}\sigma^*(r) + M_m(\bar{h}(\rho(r) + \bar{h}\sigma(r))) = 0 \quad (4.70)$$

where

$$M_m(\bar{h}) := \frac{(-\bar{h}\beta_0^*)^m (1 + \bar{h}\beta_0)}{1 - (-\bar{h}\beta_0)^m} \quad (4.71)$$

for $m = 1, 2, \dots$. A similar but somewhat more complicated analysis leads to

$$\Pi_{P(EC)^m}(r, \bar{h}) := -\beta_0^* r^k (\rho^*(r) + \bar{h}\sigma^*(r)) + M_m(\bar{h}(\rho^*(r)\sigma(r) - \sigma^*(r)\rho(r))) = 0. \quad (4.72)$$

Remark. Note that if $|\bar{h}\beta_0^*| < 1$, then $M_m(\bar{h}) \rightarrow 0$ as $m \rightarrow \infty$. It follows that in the mode of correction to convergence the weak stability characteristics are the same as those of the corrector alone.

Remark. It can be shown that both $\Pi_{P(EC)^m E}$ and $\Pi_{P(EC)^m}$ possess a root $r_1 = e^{\bar{h}} + O(\bar{h}^{p+1})$ and hence both modes are absolutely unstable for $\bar{h} > 0$.

4.5 Local Error Estimator

In the previous section, we have suggested that it would be advantageous to choose a *PC* method for which $r = r^* - m > 0$. On the other hand, by choosing $r = r^*$, there is a convenient way to estimate the principal local truncation error of the *PC* method without calculating higher derivatives of $y(x)$. In this section, we explore the so call Milne's device as follows:

Suppose $r = r^*$. Then we have

$$y(x_n) - y_n^{(0)} = c_{r+1} h^{r+1} y^{(r+1)}(x_n) + O(h^{r+2}), \quad (4.73)$$

$$y(x_n) - y_n^{(m)} = c_{r+1}^* h^{r+1} y^{(r+1)}(x_n) + O(h^{r+2}). \quad (4.74)$$

Upon subtracting, we see

$$y_n^{(m)} - y_n^{(0)} = (c_{r+1} - c_{r+1}^*) h^{r+1} y^{(r+1)}(x_n) + O(h^{r+2}).$$

Therefore, the principal local truncation error can be estimated as

$$c_{r+1}^* h^{r+1} y^{(r+1)} \approx \frac{c_{r+1}^*}{c_{r+1} - c_{r+1}^*} \left(y_n^{(m)} - y_n^{(0)} \right). \quad (4.75)$$

That is, the right-hand side of (4.75) which is immediately available from computation, can be used to estimate the local truncation error. One must not be too liberal in using (4.75), however, since the values $y_{n-i}^{(m)}$ for $i = 1, \dots, k$ used in practice are not exact $y(x_{n-i})$.

One alternative application of the Milne's device is to improve $y_n^{(0)}$ and $y_n^{(m)}$. Note that similar to (4.75) we also have

$$c_{r+1} h^{r+1} y^{(r+1)} \approx \frac{c_{r+1}}{c_{r+1} - c_{r+1}^*} \left(y_n^{(m)} - y_n^{(0)} \right). \quad (4.76)$$

Note also that

$$c_{r+1} h^{r+1} y^{(r+1)}(x_n) = c_{r+1} h^{r+1} y^{(r+1)}(x_{n-1}) + O(h^{r+1}). \quad (4.77)$$

Therefore,

$$c_{r+1} h^{r+1} y^{(r+1)}(x_n) \approx \frac{c_{r+1}}{c_{r+1} - c_{r+1}^*} \left(y_{n-1}^{(m)} - y_{n-1}^{(0)} \right). \quad (4.78)$$

Comparing with (4.73), we find that

$$\hat{y}^{(0)} := y_n^{(0)} + \frac{c_{r+1}}{c_{r+1} - c_{r+1}^*} \left(y_{n-1}^{(m)} - y_{n-1}^{(0)} \right) \quad (4.79)$$

is a better approximation to $y(x_n)$ than $y_n^{(0)}$. Similarly,

$$\hat{y}^{(m)} := y_n^{(m)} + \frac{c_{r+1}^*}{c_{r+1} - c_{r+1}^*} \left(y_n^{(m)} - y_n^{(0)} \right) \quad (4.80)$$

is better than $y_n^{(m)}$. If we call such a step a modifier denoted by M , then we may introduce the so called $PM(EC)^mME$ or $PM(EC)^mM$ modes of methods.

Remark. There are several other types of local error estimators. For example, Shampine and Gordon have used the difference between a (k, k) -pair and a $(k-1, k)$ -pair ABM methods to estimate the local error. We shall see that the Nordsieck's transformation also provides some information about the local error.

4.6 Nordsieck's Transformation

In this section we discuss the Nordsieck's transformation and its application to variable-step methods. The basic idea of Nordsieck's transformation is as follows: Locally a linear multistep method represents the solution $y(x)$ by a polynomial $\tilde{y}(x)$. Suppose the degree of $\tilde{y}(x)$ is k . We should know $\tilde{y}(x)$ completely if we know all derivatives of $\tilde{y}(x)$ of order $0, 1, \dots, k$ at a single point. (In approximation theory, this is called an osculatory interpolation.)

We illustrate the Nordsieck's transformation by considering an ABM pair of the same step number in the $P(EC)^m$ mode. Suppose

$$P: \quad y_n^{(0)} = y_{n-1}^{(m)} + h \sum_{i=1}^k \beta_i f_{n-i}^{(m-1)}, \quad (4.81)$$

$$C: \quad y_n^{(s+1)} = y_{n-1}^{(m)} + h \sum_{i=1}^k \beta_i^* f_{n-i}^{(m-1)} + h\beta_0^* f_n^{(s)}, \quad s = 0, \dots, m-1 \quad (4.82)$$

where $\beta_0^* \neq 0$. Note that

$$y_n^{(s+1)} - y_n^{(s)} = h\beta_0^*(f_n^{(s)} - f_n^{(s-1)}) \quad (4.83)$$

for $s = 1, \dots, m-1$ and that

$$y_n^{(1)} - y_n^{(0)} = h\beta_n^* \left(f_n^{(0)} - \sum_{i=1}^k \frac{\beta_i - \beta_i^*}{\beta_0^*} f_{n-i}^{(m-1)} \right). \quad (4.84)$$

Denote

$$\delta_j := \frac{\beta_j - \beta_j^*}{\beta_0^*},$$

$$f_n^{(-1)} := \sum_{i=1}^k \delta_i f_{n-i}^{(m-1)}.$$

Then (4.84) can be written as

$$y_n^{(1)} - y_n^{(0)} = h\beta_0^*(f_n^{(0)} - f_n^{(-1)}). \quad (4.85)$$

Define vectors

$$Y_n^{(s)} := [y_n^{(s)}, hf_n^{(s-1)}, hf_n^{(m-1)}, \dots, hf_n^{(m-1)}]^{T} \quad (4.86)$$

for $s = 0, 1, \dots, m$. Let

$$B := \begin{bmatrix} 1 & \beta_1 & \beta_2 & \dots & \beta_k \\ 0 & \delta_1 & \delta_2 & \dots & \delta_k \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & & & \ddots & \\ 0 & \ddots & & 1 & 0 \end{bmatrix}. \quad (4.87)$$

Then the $P(EC)^m$ method can be written as

$$P : Y_n^{(0)} = BY_{n-1}^{(m)} \quad (4.88)$$

$$C : Y_n^{(s+1)} = Y_n^{(s)} + h(f_n^s - f_n^{(s-1)}) \begin{bmatrix} \beta_0^* \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (4.89)$$

We rewrite (4.89) as

$$C : Y_n^{(s+1)} = Y_n^{(s)} + F(Y_n^{(s)})\mathbf{c} \quad (4.90)$$

where

$$F(Y_n^{(s)}) := h(f_n^{(s)} - f_n^{(s-1)}) \quad (4.91)$$

$$\mathbf{c} := \begin{bmatrix} \beta_0^* \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}^T. \quad (4.92)$$

Note that (4.88) and (4.90) are in the one-step method format, but a change of step size will cause trouble because the vector $Y_n^{(s)}$ involves past values. We now incorporate into Nordsieck's idea, namely, we store the interpolating polynomial via its derivatives at one single point rather than past values.

We begin with an example.

Example. Consider the case $k = 3$. Let the interpolating polynomial be denoted as

$$I(x) = ax^3 + bx^2 + cx + d \quad (4.93)$$

so that

$$\begin{aligned}
 I(x_n) &= y_n^{(0)}, \\
 I(x_{n-1}) &= y_{n-1}^{(m)}, \\
 I'(x_{n-1}) &= f_{n-1}^{(m-1)} \\
 I'(x_{n-2}) &= f_{n-2}^{(m-1)} \\
 I'(x_{n-3}) &= f_{n-3}^{(m-1)}
 \end{aligned} \tag{4.94}$$

whereas the elimination of the four coefficients a, b, c and d among the five equations in (4.94) will yield the 3-step Adams-Bashforth method

$$y_n^{(0)} = y_{n-1}^{(m)} + \frac{h}{12} \left(23f_{n-1}^{(m-1)} - 16f_{n-2}^{(m-1)} + 5f_{n-3}^{(m-1)} \right). \tag{4.95}$$

The specification of past values in

$$Y_{n-1}^{(m)} = [y_{n-1}^{(m)}, hf_{n-1}^{(m-1)}, hf_{n-2}^{(m-2)}, hf_{n-3}^{(m-1)}]^T \tag{4.96}$$

clearly determines $I(x)$ uniquely. Alternatively, we now determine $I(x)$ by specifying its values and its first three derivatives at x_{n-1} , i.e., through values in the vector

$$Z_{n-1}^{(m)} := [I(x_{n-1}), hI'(x_{n-1}), \frac{h^2}{2!}I''(x_{n-1}), \frac{h^3}{3!}I'''(x_{n-1})]^T. \tag{4.97}$$

The key of success is to find how the information stored in $Y_{n-1}^{(m)}$ can be restored from that in $Z_{n-1}^{(m)}$ and vice versa. The answer lies in the Nordsieck's transformation:

Theorem 4.6.1 *With*

$$Q := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \frac{3}{4} & -1 & \frac{1}{4} \\ 0 & \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \end{bmatrix},$$

we have

$$Z_{n-1}^{(m)} = QY_{n-1}^{(m)}. \tag{4.98}$$

Remark. The scaling of derivatives by powers of h has resulted in Q being independent of h .

Going back to the general case, we define

$$Z_{n-1}^{(m)} := [I(x_{n-1}), hI'(x_{n-1}), \dots, \frac{h^k}{k!}I^{(k)}(x_{n-1})]^T. \tag{4.99}$$

Then there exists a constant matrix Q such that

$$Z_{n-1}^{(m)} = QY_{n-1}^{(m)}. \tag{4.100}$$

Note that the first and the second components of $Z_{n-1}^{(m)}$ are always the same as those of $Y_{n-1}^{(m)}$ by definition. Hence the first two rows of Q are always of the form

$$\begin{array}{cccc} 1 & 0 & 0 & \dots 0 \\ 0 & 1 & 0 & \dots 0 \end{array}$$

The $P(EC)^m$ method (4.88) and (4.90) may now be transformed into

$$P: \quad Z_n^{(0)} = QBQ^{-1}Z_{n-1}^{(m)}, \quad (4.101)$$

$$C: \quad Z_n^{(s+1)} = Z_n^{(s)} + F(Z_n^{(s)})\mathbf{d} \quad (4.102)$$

for $s = 0, \dots, m-1$ with $\mathbf{d} := Q\mathbf{c}$. Note that F depends upon only the first two components of its argument and hence $F(Z_n^{(s)}) = F(Y_n^{(s)})$.

Remark. When the step size is changed from h to αh , the $Z_{n-1}^{(m)}$ vector can be updated by multiplying the i -th component of $Z_{n-1}^{(m)}$ by a factor α^i for $i = 0, 1, \dots, k$. Thus the Nordsieck's transformation facilitates the change of step size for a multistep method.