Chapter 3
Multi-Step Methods

3.1 Definition

Remark 3.1. Multi-step methods. The characteristic feature of one-step methods is that they need for computing y_{k+1} only the value from the previous approximation y_k of the solution. A straightforward extension consists in constructing methods that use for computing y_{k+1} more than one of the previous approximations $y_k, y_{k-1}, \ldots$. Such methods are called multi-step methods. □

Definition 3.2. q-step method, linear q-step method. A q-step method with $q \geq 1$ is a numerical method for approximately solving

$$y'(x) = f\left(x, y(x)\right), \quad y(x_0) = y_0, \tag{3.1}$$

where y_{k+1} depends on y_{k+1-q} but not on y_i with $i < k+1-q$.

A q-step method is called linear, if it has the form

$$y_{k+1} = \sum_{j=0}^{q-1} a_j y_{k-j} + h \sum_{j=0}^{q-1} b_j f\left(x_{k-j}, y_{k-j}\right) + h b_{-1} f\left(x_{k+1}, y_{k+1}\right), \tag{3.2}$$

$k = q-1, q, \ldots$, with $q \geq 1$, $a_0, \ldots, a_{q-1}, b_{-1}, \ldots, b_{q-1} \in \mathbb{R}$, $a_{q-1} \neq 0$ or $b_{q-1} \neq 0$. For $q = 1$, the method is called a one-step method. If $b_{-1} \neq 0$, then the linear q-step method is an implicit method, otherwise it is an explicit method. □

Remark 3.3. Initial values for a q-step method. A q-step method needs q initial values. However, the initial value problem (3.1) provides only the initial value y_0. The second initial value y_1 can be computed with a one-step method, the next initial value y_2 with a one-step method or with a two-step method and so on. It follows that all initial values y_i, $i > 0$, are already numerical approximations. This aspect has to be taken into account in the error analysis of multi-step methods, see Remark 3.23. □

3.2 Predictor-Corrector Methods

Remark 3.4. Construction. Starting point of the construction of predictor-corrector methods is the equivalent integral form of the initial value problem (3.1)

$$y(x) = y_0 + \int_{x_0}^{x} f\left(t, y(t)\right) \, \mathrm{d}t. \tag{3.3}$$

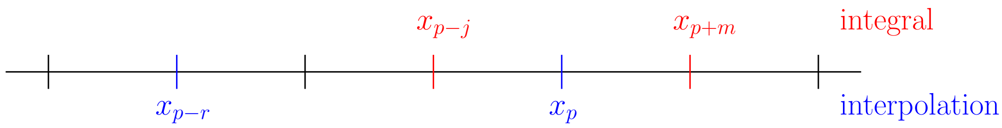

Fig. 3.1 Parameters in the derivation of predictor-corrector schemes.

Denote the solution at $\tilde{x}$ by $y(\tilde{x})$, then it holds that

$$y(x) = y(\tilde{x}) + \int_{\tilde{x}}^{x} f\left(t, y(t)\right) \, \mathrm{d}t. \tag{3.4}$$

The main idea of predictor-corrector methods consists in approximating the integral on the right-hand side of (3.4) in an appropriate way. There are two principal difficulties:
- The dependency of the term in the integral on t is generally not known since the function $y(t)$ is unknown.
- Even if the dependency of the function in the integral on t is known, generally it will be impossible to find an analytic expression of the solution.

Consider an equidistant grid with nodes

$$x_i = x_0 + ih, \quad i = 0, 1, \ldots.$$

For the derivation of the methods, assume that the term in the integral is known. Then, the derivation is similar to the derivation of the Newton[1]–Cotes[2] formulas for numerical quadrature. In this approach, the term in the integral of (3.4) is replaced by a polynomial interpolant throughrestaurants given nodes. Let the boundaries of the integral be the nodes

$$\begin{aligned} \tilde{x} &= x_{p-j}, && \text{starting point with parameter } j, \\ x &= x_{p+m} && \text{end point with parameter } m, \end{aligned} \tag{3.5}$$

with parameters $j, m \in \mathbb{N}_0$ that need yet to be determined. It will be required that the interpolation polynomial $p_r(x)$ satisfies the following properties:
- the degree of $p_r(x)$ is lower than or equal to r,
- $p_r(x_i) = f(x_i, y(x_i))$ for $i = p, p-1, \ldots, p-r$.

Thus, x_p is the most right-hand side node for computing the interpolation polynomial. The value r is a third parameter, compare Figure 3.1. Notice that two sets of nodes are involved in the construction, namely the nodes that determine the boundaries of the integral and the nodes that are used to define the interpolation polynomial. The solution of this interpolation problem is given by the Lagrange[3] interpolation polynomial

$$p_r(x) = \sum_{i=0}^{r} f\left(x_{p-i}, y(x_{p-i})\right) L_i(x)$$

with

$$L_i(x) = \prod_{l=0, l\neq i}^{r} \frac{x - x_{p-l}}{x_{p-i} - x_{p-l}}, \ i = 0, 1, \ldots, r. \tag{3.6}$$

It follows by using (3.4), (3.5), (3.6), and by replacing the unknown values $y(x_{p-i})$ by their computed approximations y_{p-i} that

[1] Isaac Newton (1642 – 1727)
[2] Roger Cotes (1682 – 1716)
[3] Joseph Louis Lagrange (1736 – 1813)

$$y_{p+m} \approx y_{p-j} + \sum_{i=0}^{r} f\left(x_{p-i}, y_{p-i}\right) \int_{x_{p-j}}^{x_{p+m}} L_i(t)\ \mathrm{d}t = y_{p-j} + h \sum_{i=0}^{r} \beta_i f\left(x_{p-i}, y_{p-i}\right) \tag{3.7}$$

with

$$\beta_i = \frac{1}{h} \int_{x_{p-j}}^{x_{p+m}} L_i(t)\ \mathrm{d}t = \frac{1}{h} \int_{x_{p-j}}^{x_{p+m}} \left(\prod_{l=0, l\neq i}^{r} \frac{t - x_{p-l}}{x_{p-i} - x_{p-l}} \right) \mathrm{d}t.$$

The constructed method is in particular linear. Notice that so far the assumption of having an equidistant grid was not used.

Finally, the formula for β_i should be simplified. To this end, notice that all fixed values from the interval are nodes of the equidistant grid, such that, e.g., $x_p = x_0 + ph$. Replacing these values and using the substitution

$$t = x_p + sh \quad \Longrightarrow \quad \mathrm{d}t = h\mathrm{d}s,$$

yields

$$\begin{aligned} \beta_i &= \frac{1}{h} \int_{-j}^{m} \left(\prod_{l=0, l\neq i}^{r} \frac{x_p + sh - x_{p-l}}{x_{p-i} - x_{p-l}} \right) h\ \mathrm{d}s = \int_{-j}^{m} \left(\prod_{l=0, l\neq i}^{r} \frac{x_0 + ph + sh - x_0 - ph + lh}{x_0 + ph - ih - x_0 - ph + lh} \right) \mathrm{d}s \\ &= \int_{-j}^{m} \left(\prod_{l=0, l\neq i}^{r} \frac{s + l}{-i + l} \right) \mathrm{d}s. \end{aligned} \tag{3.8}$$

Now, different methods can be obtained depending on the choice of m, j, and r. There are four important classes of methods. □

Example 3.5. Adams[4]–Bashforth[5] methods. The class of q-step Adams–Bashforth methods is given by $m = 1$, $j = 0$, and $r = q - 1$. It follows that the q-step Adams–Bashforth method uses the nodes $x_{k+1-q}, \ldots, x_k$ for computing the Lagrangian interpolation polynomial. These are q nodes and $p_q(x)$ is at most of degree $q - 1$. Adams–Bashforth methods are explicit methods. They have the general form

$$y_{k+1} = y_k + h \sum_{i=0}^{q-1} \beta_i f\left(x_{k-i}, y_{k-i}\right), \tag{3.9}$$

see (3.7), with

$$\beta_i = \int_0^1 \left(\prod_{l=0, l\neq i}^{q-1} \frac{s + l}{-i + l} \right) \mathrm{d}s, \tag{3.10}$$

compare (3.8).

In the case $q = 1$, the term in the integral in (3.4) is replaced by a constant interpolation polynomial with the node $(x_k, f(x_k, y_k))$. Using the convention that the product is 1 if there is formally no factor in (3.10), this approach yields

$$y_{k+1} = y_k + h \left(\int_0^1 \mathrm{d}s \right) f\left(x_k, y_k\right) = y_k + hf\left(x_k, y_k\right),$$

i.e., one obtains the explicit Euler method.

If $q = 2$, then the term in the integral is approximated by a linear interpolation polynomial with the nodes $(x_{k-1}, f(x_{k-1}, y_{k-1}))$ and $(x_k, f(x_k, y_k))$. Using (3.9) and (3.10), one obtains

[4] John Couch Adams (1819 – 1892)

[5] Francis Bashforth (1819 – 1912)

$$
\begin{aligned}
y_{k+1} &= y_k + h\left[\left(\int_0^1 \frac{s+1}{1}\,\mathrm{d}s\right) f\left(x_k, y_k\right) + \left(\int_0^1 \frac{s}{-1}\,\mathrm{d}s\right) f\left(x_{k-1}, y_{k-1}\right)\right)\right] \\
&= y_k + h\left[\frac{3}{2} f\left(x_k, y_k\right) - \frac{1}{2} f\left(x_{k-1}, y_{k-1}\right)\right] \\
&= y_k + \frac{h}{2}\left[3 f\left(x_k, y_k\right) - f\left(x_{k-1}, y_{k-1}\right)\right].
\end{aligned}
$$

$q \geq 3$, *exercise* □

Example 3.6. Adams–Moulton[6] *methods.* Adams–Moulton methods are defined by $m = 0$, $j = 1$, and $r = q$. Hence, it follows that

$$
\beta_i = \int_{-1}^0 \left(\prod_{l=0, l\neq i}^{q} \frac{s+l}{-i+l}\right) \mathrm{d}s
$$

and from (3.7) that

$$
y_k = y_{k-1} + h \sum_{i=0}^{q} \beta_i f\left(x_{k-i}, y_{k-i}\right)
$$

or, by transforming the index,

$$
y_{k+1} = y_k + h \sum_{i=0}^{q} \beta_i f\left(x_{k+1-i}, y_{k+1-i}\right).
$$

The $q+1$ nodes of these methods are given by $x_{k+1-q}, \ldots, x_k, x_{k+1}$. That means, Adams–Moulton methods are implicit methods.

This class contains two one-step methods that are obtained for $q = 0$ (which can be used in contrast to the requirement in Definition 3.2) and $q = 1$. Notice that the parameter q in (3.2) determines both the previous approximations to be used and the previous arguments of the function f. But in the construction of the methods, three independent parameters were introduced to determine these values. This construction introduces some freedom which allows here to set $q = 0$.

Considering the case $q = 0$, then the term in the integral is replaced by a constant interpolation polynomial with the node at $(x_{k+1}, f(x_{k+1}, y_{k+1}))$. This approach results in the method

$$
y_{k+1} = y_k + h\left(\int_{-1}^0 \mathrm{d}s\right) f\left(x_{k+1}, y_{k+1}\right) = y_k + h f\left(x_{k+1}, y_{k+1}\right),
$$

which is the implicit Euler method.

For $q = 1$, one uses a linear interpolation polynomial with the points $(x_k, f(x_k, y_k))$ and $(x_{k+1}, f(x_{k+1}, y_{k+1}))$. One gets

$$
\begin{aligned}
y_{k+1} &= y_k + h\left[\left(\int_{-1}^0 \frac{s+1}{1}\,\mathrm{d}s\right) f\left(x_{k+1}, y_{k+1}\right) + \left(\int_{-1}^0 \frac{s}{-1}\,\mathrm{d}s\right) f\left(x_k, y_k\right)\right] \\
&= y_k + h\left[\frac{1}{2} f\left(x_{k+1}, y_{k+1}\right) + \frac{1}{2} f\left(x_k, y_k\right)\right] \\
&= y_k + \frac{h}{2}\left[f\left(x_{k+1}, y_{k+1}\right) + f\left(x_k, y_k\right)\right].
\end{aligned}
$$

This method is the trapezoidal rule. □

[6] Forest Ray Moulton (1872 – 1952)

Example 3.7. Nyström[7] methods. The class of Nyström methods is obtained by using $m = 1$, $j = 1$, and $r = q - 1$. They have the form

$$y_{k+1} = y_{k-1} + h \sum_{i=0}^{q-1} \beta_i f\left(x_{k-i}, y_{k-i}\right)$$

with

$$\beta_i = \int_{-1}^{1} \left(\prod_{l=0, l \neq i}^{q-1} \frac{s+l}{-i+l} \right) \mathrm{d}s.$$

These methods are explicit and one uses the q nodes $x_{k+1-q}, \ldots, x_k$.

One gets, e.g., for $q = 1$, the method

$$y_{k+1} = y_{k-1} + h \left(\int_{-1}^{1} \mathrm{d}s \right) f\left(x_k, y_k\right) = y_{k-1} + 2hf\left(x_k, y_k\right).$$

□

Example 3.8. Milne[8] method. Milne methods are defined by $m = 0$, $j = 2$, and $r = q$. Using a transform of the index, one finds that they have the form

$$y_{k+1} = y_{k-1} + h \sum_{i=0}^{q} \beta_i f\left(x_{k+1-i}, y_{k+1-i}\right)$$

with

$$\beta_i = \int_{-2}^{0} \left(\prod_{l=0, l \neq i}^{q} \frac{s+l}{-i+l} \right) \mathrm{d}s.$$

Thus, these are implicit methods. □

Remark 3.9. On the coefficients of multi-step methods. One can find tables with the coefficients for multi-step methods in the literature. □

Remark 3.10. Using implicit methods in practice, predictor-corrector methods. If implicit methods are used, then one has to solve in each node x_{k+1} an equation that is generally nonlinear. This step can be performed with some kind of fixed-point iteration, e.g., with a method of Newton-type. To achieve a good efficiency of the method, a good initial iterate is of importance. To obtain a good initial iterate, one can use an explicit (multi-step) method. For this reason, explicit multi-step methods are called predictor methods and implicit multi-step methods are called corrector methods. The combination of a predictor method with a corrector method is called predictor-corrector method.

Often, it is sufficient for computing the next iterate to perform the predictor step and one or two corrector steps. □

Remark 3.11. Nordsieck[9] form. It is possible to transform multi-step methods in a one-step form, the so-called Nordsieck form. This form uses instead of

$$y_k, \ldots, y_{k-q+1}, f\left(x_k, y_k\right), \ldots, f\left(x_{k-q+1}, y_{k-q+1}\right),$$

the values

[7] Evert J. Nyström (1895 – 1960)
[8] William Edwin Milne (1890 – 1971)
[9] Arnold Nordsieck (1911 – 1971)

$$y_k, y'(x_k), y''(x_k), \ldots, y^{(q)}(x_k),$$

e.g., see (Strehmel *et al.*, 2012, Section 4.4.3). The advantage of the Nordsieck form consists in the possibility of applying a step length control as it is known from one-step methods, Section 1.3. Otherwise, a step length control for form (3.2) of multi-step methods becomes rather complicated. On the other hand, using the Nordsieck form requires that the solution of the initial value problem is q times continuously differentiable. □

3.3 Convergence of Multi-Step Methods

Remark 3.12. Generalities. In this section, linear multi-step methods of the form (3.2) will be considered. Similarly to one-step methods, notations like local error, consistency, or order of convergence will be introduced. The extension of these notations to nonlinear multi-step methods is straightforward. □

Definition 3.13. Local error. Let y_{k+1} be the results of (3.2), $k \ge q$, where the initial values are exactly the values of the solution

$$y_{k+1-q} = y(x_{k+1-q}), \ldots, y_k = y(x_k).$$

Then, the local error is defined by

$$\mathrm{le}(x_{k+1}) = \mathrm{le}_{k+1} = y(x_{k+1}) - \left[\sum_{j=0}^{q-1} a_j y\left(x_{k-j}\right) + h \sum_{j=-1}^{q-1} b_j f\left(x_{k-j}, y(x_{k-j})\right)\right]. \tag{3.11}$$

□

Definition 3.14. Consistent method, consistency order. Let $y(x)$ be the solution of the initial value problem (3.1), $S = \{(x,y) \ : \ x \in I = [x_0, x_e], \ y \in \mathbb{R}\}$, and I_N an equidistant mesh on I with N intervals. The multi-step method (3.2) is called consistent if for all $f \in C(S)$, which satisfy in S a Lipschitz condition with respect to y, it holds

$$\lim_{h \to 0} \left(\max_{x_k \in I_N} \frac{\mathrm{le}(x_k + h)}{h} \right) = 0, \quad \text{with} \quad h = \frac{x_e - x_0}{N}. \tag{3.12}$$

If the expression on the left-hand side converges like h^p for $p \ge 1$, then the multi-step scheme has the consistency order p. □

Example 3.15. Consistency order for a Nyström method. The consistency order of a multi-step method can be computed in the same way as for a one-step method by expanding the local error in a Taylor series with respect to h. After having then divided by h, the order of the first non-vanishing term gives the consistency order.

Consider the Nyström method for $q = 3$

$$\begin{aligned} y_{k+1} &= y_{k-1} + h \left[\left(\int_{-1}^{1} \prod_{l=1}^{2} \frac{s+l}{l} \, \mathrm{d}s \right) f\left(x_k, y_k\right) + \left(\int_{-1}^{1} \prod_{l=0, l\ne 1}^{2} \frac{s+l}{-1+l} \, \mathrm{d}s \right) f\left(x_{k-1}, y_{k-1}\right) \right. \\ &\qquad \left. + \left(\int_{-1}^{1} \prod_{l=0}^{1} \frac{s+l}{-2+l} \, \mathrm{d}s \right) f\left(x_{k-2}, y_{k-2}\right) \right] \\ &= y_{k-1} + h \left[\frac{7}{3} f\left(x_k, y_k\right) - \frac{2}{3} f\left(x_{k-1}, y_{k-1}\right) + \frac{1}{3} f\left(x_{k-2}, y_{k-2}\right) \right]. \end{aligned}$$

It follows with (3.11) and (3.1) that

$$\begin{aligned}\mathrm{le}(x_{k+1}) &= y(x_{k+1}) - y(x_{k-1}) - h\left[\frac{7}{3}f\left(x_k, y(x_k)\right) - \frac{2}{3}f\left(x_{k-1}, y(x_{k-1})\right) + \frac{1}{3}f\left(x_{k-2}, y(x_{k-2})\right)\right] \\ &= y(x_{k+1}) - y(x_{k-1}) - h\left[\frac{7}{3}y'(x_k) - \frac{2}{3}y'(x_{k-1}) + \frac{1}{3}y'(x_{k-2})\right]. \end{aligned} \tag{3.13}$$

Now, the the individual terms will be expanded

$$\begin{aligned} y(x_{k+1}) &= y(x_k + h) = y(x_k) + hy'(x_k) + \frac{h^2}{2}y''(x_k) + \frac{h^3}{6}y'''(x_k) + \frac{h^4}{24}y^{(4)}(x_k) + \mathcal{O}(h^5), \\ y(x_{k-1}) &= y(x_k - h) = y(x_k) - hy'(x_k) + \frac{h^2}{2}y''(x_k) - \frac{h^3}{6}y'''(x_k) + \frac{h^4}{24}y^{(4)}(x_k) + \mathcal{O}(h^5), \\ y'(x_{k-1}) &= y'(x_k - h) = y'(x_k) - hy''(x_k) + \frac{h^2}{2}y'''(x_k) - \frac{h^3}{6}y^{(4)}(x_k) + \mathcal{O}(h^4), \\ y'(x_{k-2}) &= y'(x_k - 2h) = y'(x_k) - 2hy''(x_k) + 2h^2y'''(x_k) - \frac{4h^3}{3}y^{(4)}(x_k) + \mathcal{O}(h^4). \end{aligned}$$

Inserting these expressions in formula (3.13) for the local error gives

$$\begin{aligned} \mathrm{le}(x_{k+1}) &= y(x_k) + hy'(x_k) + \frac{h^2}{2}y''(x_k) + \frac{h^3}{6}y'''(x_k) + \frac{h^4}{24}y^{(4)}(x_k) \\ &\quad -y(x_k) + hy'(x_k) - \frac{h^2}{2}y''(x_k) + \frac{h^3}{6}y'''(x_k) - \frac{h^4}{24}y^{(4)}(x_k) \\ &\quad -\frac{7h}{3}y'(x_k) + \frac{2}{3}\left[hy'(x_k) - h^2y''(x_k) + \frac{h^3}{2}y'''(x_k) - \frac{h^4}{6}y^{(4)}(x_k)\right] \\ &\quad -\frac{1}{3}\left[hy'(x_k) - 2h^2y''(x_k) + 2h^3y'''(x_k) - \frac{4h^4}{3}y^{(4)}(x_k)\right] + \mathcal{O}(h^5) \\ &= \frac{h^4}{3}y^{(4)}(x_k) + \mathcal{O}(h^5). \end{aligned}$$

With (3.12), one obtains that this method has consistency order 3. □

Remark 3.16. Linear multi-step methods with a high order of convergence. The goal in constructing multi-step methods consists of course in obtaining convergent methods of high order. A high order of convergence can be expected only if the consistency order is high, i.e., if the local error is small. Using the Taylor series expansion of the terms in the local error and requiring that as many leading terms as possible vanish, one gets a linear system of equations for determining the coefficients a_j, b_j, $j = 0, \ldots, q-1$ and b_{-1} in (3.11). In this way, one obtains a method of the form

$$y_{k+1} - \sum_{j=0}^{q-1} a_j y_{k-j} = h \sum_{j=-1}^{q-1} b_j f\left(x_{k-j}, y_{k-j}\right). \tag{3.14}$$

Constructing one-step methods in this way, one always obtains a convergent one-step method, e.g., compare Example 1.29. However, the situation might be different for multi-step methods. □

Example 3.17. Non-convergent multi-step method. Consider the idea from Remark 3.16 for the construction of an explicit linear multi-step method with $q = 2$ and with maximal order of consistency. That means, the ansatz for the method is as follows, compare (3.14),

$$y_{k+1} - a_0 y_k - a_1 y_{k-1} = h\left[b_0 f\left(x_k, y_k\right) + b_1 f\left(x_{k-1}, y_{k-1}\right)\right].$$