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

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

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. □

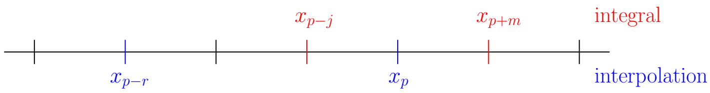

Fig. 3.1 Parameters in the derivation of predictor-corrector schemes.

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)

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

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

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

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. 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 $i = p, p-1, \ldots, p-r$.

[1] Isaac Newton (1642 – 1727)
[2] Roger Cotes (1682 – 1716)

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. Note 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

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

with

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

The constructed method is in particular linear. Note 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, note 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

$$t = x_p + sh \quad \Longrightarrow \quad dt = hds,$$

yields

$$\beta_i = \frac{1}{h} \int_{-j}^{m} \left(\prod_{l=0, l\neq i}^{r} \frac{x_p + sh - x_{p-l}}{x_{p-i} - x_{p-l}} \right) h\ ds$$

[3] Joseph Louis Lagrange (1736 – 1813)

$$= \int_{-j}^{m} \left(\prod_{l=0, l \neq i}^{r} \frac{x_0 + ph + sh - x_0 - ph + lh}{x_0 + ph - ih - x_0 - ph + lh} \right) ds$$

$$= \int_{-j}^{m} \left(\prod_{l=0, l \neq i}^{r} \frac{s + l}{-i + l} \right) ds. \tag{3.8}$$

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

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

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

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

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

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

---

[4] John Couch Adams (1819 – 1892)

[5] Francis Bashforth (1819 – 1912)

$$= y_k + \frac{h}{2}\big[3f(x_k, y_k) - f(x_{k-1}, y_{k-1})\big].$$

$q \geq 3$, *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

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

and from (3.7) that

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

or, by transforming the index,

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

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 $q = 1$. Note 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

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

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

$$y_{k+1} = y_k + h \left[\left(\int_{-1}^{0} \frac{s+1}{1}\, ds \right) f(x_{k+1}, y_{k+1}) + \left(\int_{-1}^{0} \frac{s}{-1}\, ds \right) f(x_k, y_k) \right]$$

[6] Forest Ray Moulton (1872 – 1952)

$$= y_k + h\left[\frac{1}{2}f\left(x_{k+1}, y_{k+1}\right) + \frac{1}{2}f\left(x_k, y_k\right)\right]$$
$$= y_k + \frac{h}{2}\left[f\left(x_{k+1}, y_{k+1}\right) + f\left(x_k, y_k\right)\right].$$

This method is the trapezoidal rule. □

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

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

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

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

□

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

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

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

[7] Evert J. Nyström (1895 – 1960)

[8] William Edwin Milne (1890 – 1971)

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

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

see, e.g., (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), $k \geq q$, 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

[9] Arnold Nordsieck (1911 – 1971)

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

□

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 $p \geq 1$, 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 $q = 3$

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

It follows with (3.11) and (3.1) that

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

Now, the the individual terms will be expanded

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

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),

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

The local error has the form

$$\begin{aligned}
\operatorname{le}(x_{k+1}) &= y\left(x_{k+1}\right) - a_0 y\left(x_k\right) - a_1 y\left(x_{k-1}\right) - h b_0 f\left(x_k, y(x_k)\right) \\
&\quad - h b_1 f\left(x_{k-1}, y(x_{k-1})\right) \\
&= y\left(x_{k+1}\right) - a_0 y\left(x_k\right) - a_1 y\left(x_{k-1}\right) - h b_0 y'(x_k) - h b_1 y'(x_{k-1}).
\end{aligned}$$

Now, the individual terms are expanded in powers of h:

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

Inserting the expansions gives

$$\begin{aligned}
\operatorname{le}(x_{k+1}) &= y\left(x_k\right) + h y'\left(x_k\right) + \frac{h^2}{2} y''\left(x_k\right) + \frac{h^3}{6} y'''\left(x_k\right) - a_0 y\left(x_k\right) \\
&\quad - a_1\left[y\left(x_k\right) - h y'\left(x_k\right) + \frac{h^2}{2} y''\left(x_k\right) - \frac{h^3}{6} y'''\left(x_k\right)\right] - h b_0 y'(x_k) \\
&\quad - h b_1\left[y'\left(x_k\right) - h y''\left(x_k\right) + \frac{h^2}{2} y'''\left(x_k\right)\right] + \mathcal{O}(h^4) \\
&= \left[1 - a_1 - a_0\right] y\left(x_k\right) + \left[1 + a_1 - b_1 - b_0\right] h y'\left(x_k\right) \\
&\quad + \left[\frac{1}{2} - \frac{a_1}{2} + b_1\right] h^2 y''\left(x_k\right) + \left[\frac{1}{6} + \frac{a_1}{6} - \frac{b_1}{2}\right] h^3 y'''\left(x_k\right) + \mathcal{O}(h^4).
\end{aligned}$$

Requiring that the first four terms should vanish leads to the following linear system of equations

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 1 \\ 1/2 & 0 & -1 & 0 \\ -1/6 & 0 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} a_1 \\ a_0 \\ b_1 \\ b_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1/2 \\ 1/6 \end{pmatrix}.$$

The unique solution of this system is $a_1 = 5$, $a_0 = -4$, $b_1 = 2$, $b_0 = 4$. Consequently, one obtains the method

$$y_{k+1} = -4y_k + 5y_{k-1} + h\left[4f\left(x_k, y_k\right) + 2f\left(x_{k-1}, y_{k-1}\right)\right] \tag{3.15}$$

with third order of consistency.

Next, the convergence of the method will be studied at the model initial value problem

$$y'(x) = -y(x), \quad y(0) = 1,$$

with the solution $y(x) = \exp(-x)$. As second initial condition, one takes the value of the solution in the mesh point $x_1 = h$, i.e., $y_1 = \exp(-h)$. Inserting the special form of the right-hand side of the model problem, $f(x_k, y_k) = -y_k$, in (3.15), one can represent the computed solution explicitly. This solution satisfies the homogeneous linear difference equation

$$y_{k+1} + (4 + 4h)y_k + (-5 + 2h)y_{k-1} = 0.$$

The solution of this difference equation can be obtained with the ansatz $y_k = \xi^k$. Inserting this ansatz in the difference equation leads to

$$\xi^{k+1} + (4 + 4h)\xi^k + (-5 + 2h)\xi^{k-1} = 0.$$

This equation is satisfied for $\xi = 0$, which however does not satisfy the initial conditions. Other solutions can be obtained after division by ξ^{k-1} from

$$\xi^2 + (4 + 4h)\xi + (-5 + 2h) = 0. \tag{3.16}$$

One gets the solutions

$$\xi_1(h) = -2 - 2h + 3\sqrt{1 + \frac{2}{3}h + \frac{4}{9}h^2}, \quad \xi_2(h) = -2 - 2h - 3\sqrt{1 + \frac{2}{3}h + \frac{4}{9}h^2}.$$

For simplicity, the dependency on h will be neglected in the notation. The general solution of the difference equations can be represented as a linear combination of the special solutions (superposition principle)

$$y_k = C_1\xi_1^k + C_2\xi_2^k.$$

Now, the constants can be determined from the initial conditions. It holds

$$y_0 = C_1 + C_2 = 1, \quad y_1 = e^{-h} = C_1\xi_1 + C_2\xi_2,$$

from what follows that

$$C_1(h) = \frac{e^{-h} - \xi_2}{\xi_1 - \xi_2}, \quad C_2(h) = -\frac{e^{-h} - \xi_1}{\xi_1 - \xi_2}.$$

Expanding $\xi_1(h)$, $\xi_2(h)$, $C_1(h)$ and $C_2(h)$ in powers of h and inserting these expansions in the solution *(exercise)*, gives for fixed $x > 0$ and $h_N := x/N$

$$\begin{aligned} y_N &= \left[1 + \mathcal{O}\left(\frac{x}{N}\right)\right] \left[1 - \frac{x}{N} + \mathcal{O}\left(\left(\frac{x}{N}\right)^2\right)\right]^N \\ &\quad - \frac{1}{216}\left(\frac{x}{N}\right)^4 \left[1 + \mathcal{O}\left(\frac{x}{N}\right)\right] \left[-5 - 3\frac{x}{N} + \mathcal{O}\left(\left(\frac{x}{N}\right)^2\right)\right]^N. \end{aligned}$$

Considering now the convergence of the method, i.e., $h_N \to 0 \iff N \to \infty$. Then, one obtains for the first term, using well known properties of the exponential, that

$$\lim_{N\to\infty} \left[1 + \mathcal{O}\left(\frac{x}{N}\right)\right] \left[1 - \frac{x}{N} + \mathcal{O}\left(\left(\frac{x}{N}\right)^2\right)\right]^N = e^{-x}.$$

This part approximates the solution of the model problem. For the second term, it holds that

$$\begin{aligned} &- \frac{1}{216}\left(\frac{x}{N}\right)^4 \left[1 + \mathcal{O}\left(\frac{x}{N}\right)\right] \left[-5 - 3\frac{x}{N} + \mathcal{O}\left(\left(\frac{x}{N}\right)^2\right)\right]^N \\ &= -\frac{(-5)^N}{216}\left(\frac{x}{N}\right)^4 \left[1 + \mathcal{O}\left(\frac{x}{N}\right)\right] \left[1 + \frac{3}{5}\frac{x}{N} + \mathcal{O}\left(\left(\frac{x}{N}\right)^2\right)\right]^N. \end{aligned}$$

Since

$$\lim_{N\to\infty} \left[1 + \frac{3}{5}\frac{x}{N} + \mathcal{O}\left(\left(\frac{x}{N}\right)^2\right)\right]^N = e^{3x/5},$$

one finds that the second term behaves for large N as follows

$$-\frac{(-5)^N}{216}\left(\frac{x}{N}\right)^4 e^{3x/5}. \tag{3.17}$$

This expression oscillates with increasing N and the modulus is increasing for finer grids ('exponential $(-5)^N$ is stronger than polynomial $(x/N)^4$'), compare the values for $x = 1$ in the following table

N	value of (3.17)
1	0.0421787
2	- 0.1054467
3	0.3514890
4	- 1.3180836
5	5.2723345
6	- 21.96806
7	94.14883
8	- 411.90113
9	1830.6717
10	- 8238.0226

It follows that the method does not converge.

Such an oscillatory behavior can be observed also if the method is applied for solving other initial value problems. For the considered example, the reason for this behavior is that the general solution of the difference equation contains a term that becomes arbitrarily large for large k and for small h (or large N). For the considered method it is

$$\lim_{h\to 0} \xi_2(h) = -5 \quad \Longrightarrow \quad \lim_{k\to\infty} \left|\xi_2^k(h)\right| = \infty.$$

The solution of the difference equation was obtained from the roots of the polynomial (3.16). It can be guessed that the roots of this polynomial will be of importance for the convergence of multi-step methods. □

Definition 3.18. Null stable linear multi-step method. A linear q-step method is called null stable if the first characteristic polynomial

$$\Psi(\xi) = \xi^q - a_0\xi^{q-1} - \ldots - a_{q-1} \tag{3.18}$$

possesses only roots ξ_q with $\left|\xi_q\right| \le 1$ that are simple in the case that $\left|\xi_q\right| = 1$. For the notation 'null stable', compare Remark 3.37 below. □

Example 3.19. Null stability for predictor-corrector methods. The methods from the four most important classes of predictor-corrector methods are null stable.

- *Adams–Bashforth methods, Adams–Moulton methods.* The first characteristic polynomial has the form

$$\Psi(\xi) = \xi^q - \xi^{q-1} = (\xi - 1)\,\xi^{q-1}.$$

  The only non-trivial root is $\xi_q = 1$ and this root is simple.
- *Nyström methods, Milne methods.* For these methods, the first characteristic polynomial is

$$\Psi(\xi) = \xi^q - \xi^{q-2} = (\xi + 1)\,(\xi - 1)\,\xi^{q-2}.$$

Hence, the only non-trivial roots are $\xi_q = 1$ and $\xi_q = -1$. They are simple. Null stability does not mean stable in the sense that the method can be applied for the numerical solution of stiff problems, see Example 3.22. □

Theorem 3.20. First Dahlquist[10] barrier. *The maximal order of consistency of a null stable linear q-step method is*

$$p = \begin{cases} q+1 & \textit{for } q \textit{ odd}, \\ q+2 & \textit{for } q \textit{ even}, \\ q & \textit{if } b_{-1} \leq 0, \textit{in particular, if the method is explicit.} \end{cases}$$

Proof. Only a sketch of the proof is given here, for details see the literature, e.g., (Strehmel *et al.*, 2012, Section 4.2.3) or (Hairer *et al.*, 1993, Section III.3).

First, one sets for $\xi \in \mathbb{C}$, $|\xi| < 1$,

$$z = \frac{\xi - 1}{\xi + 1}.$$

Then, one defines the polynomials

$$R(z) = \left(\frac{1-z}{2}\right)^q \Psi(\xi) = \sum_{l=0}^{q} \alpha_l z^l,$$
$$S(z) = \left(\frac{1-z}{2}\right)^q \sigma(\xi) = \sum_{l=0}^{q} \beta_l z^l,$$

with

$$\sigma(\xi) = b_{-1}\xi^q + b_0\xi^{q-1} + \ldots + b_{q-1}. \tag{3.19}$$

As next step, one can prove that a linear multi-step method has consistency order p if and only if

$$R(z)\left(\ln\frac{1+z}{1-z}\right)^{-1} - S(z) = \mathcal{O}\left(z^p\right) \quad \text{for } z \to 0.$$

Using a Taylor series expansion of the term with the logarithm, one has on the left-hand side of this statement a polynomial. Now, one studies which coefficients of this polynomial might vanish such that the method is null stable in the individual cases given in the theorem. ■

Example 3.21. Consistency order of some predictor-corrector methods.

- Adams–Bashforth methods with q steps have the consistency order q and Adams–Moulton methods with q steps possess the consistency order $q+1$. Thus, Adams–Moulton methods where q is even have an order that is less than the maximal possible order according to Theorem 3.20.
- The 2-step Milne method (also Milne–Simpson method)

$$y_{k+1} = y_{k-1} + h\left(\frac{1}{3}f(x_{k+1}, y_{k+1}) + \frac{4}{3}f(x_k, y_k) + \frac{1}{3}f(x_{k-1}, y_{k-1})\right) \tag{3.20}$$

[10] Germund Dahlquist (1925 – 2005)

has the consistency order 4. This method achieves the maximal order of consistency for a null stable 2-step method.

□

Example 3.22. Convergence and stability of the 2*-step Milne method.* This implicit method is null stable, see Example 3.19, and it possesses the maximal possible order of consistency for a null stable method, see Example 3.21. Thus, so far it shows favorable properties. But having a closer look on its stability reveals that this method has a severe drawback.

Consider again the model initial value problem

$$y'(x) = \lambda y(x), \quad y(0) = 1,$$

with the solution $y(x) = \exp(\lambda x)$. Applying the 2-step Milne method for the solution of this problem, then the method (3.20) has the form

$$y_{k+1} = y_{k-1} + h\lambda \left(\frac{1}{3} y_{k+1} + \frac{4}{3} y_k + \frac{1}{3} y_{k-1} \right).$$

This equation can be rewritten as a linear difference equation

$$\left(1 - \frac{h\lambda}{3}\right) y_{k+1} - \frac{4h\lambda}{3} y_k - \left(1 + \frac{h\lambda}{3}\right) y_{k-1} = 0.$$

The general solution of this difference equation can be represented in the form

$$y_k = C_1 \xi_1^k + C_2 \xi_2^k, \tag{3.21}$$

where $\xi_1(h)$ and $\xi_2(h)$ are the solutions of the quadratic equation

$$\left(1 - \frac{h\lambda}{3}\right) \xi^2 - \frac{4h\lambda}{3} \xi - \left(1 + \frac{h\lambda}{3}\right) = 0.$$

One obtains

$$\xi_1(h) = \frac{3}{3 - h\lambda} \left(\frac{2h\lambda}{3} + \sqrt{1 + \frac{(h\lambda)^2}{3}} \right),$$

$$\xi_2(h) = \frac{3}{3 - h\lambda} \left(\frac{2h\lambda}{3} - \sqrt{1 + \frac{(h\lambda)^2}{3}} \right).$$

Now, the constants C_1, C_2 can be determined from the initial condition and from the value of the first step. It is for $x = 0$

$$C_1 + C_2 = 1 \tag{3.22}$$

and for $x = h$, taking the exact value,

$$e^{\lambda h} = C_1\xi_1 + (1 - C_1)\xi_2 = (1 - C_2)\,\xi_1 + C_2\xi_2. \qquad (3.23)$$

Expanding $\xi_1(h)$ and $\xi_2(h)$ in powers of h at $h = 0$, one obtains as first order approximation by computing the derivatives of $\xi_1(h)$ and $\xi_2(h)$, respectively, and inserting zero

$$\xi_1(h) = 1 + \lambda h + \mathcal{O}\left(h^2\right), \quad \xi_2(h) = -1 + \frac{\lambda}{3}h + \mathcal{O}\left(h^2\right). \qquad (3.24)$$

In the interesting case, $\lambda < 0$, $\xi_1(h)$ approaches 1 from left, with values smaller than 1, and $\xi_2(h)$ approaches -1 also from left, but here the modulus of $\xi_2(h)$ is larger than 1. The last property leads to undesired effects.

For the approximate solution in the node $x_k = kh$, $k = 0, 1, \ldots$, one gets with (3.21)

$$y_k = C_1\left(1 + \lambda h + \mathcal{O}\left(h^2\right)\right)^{x_k/h} + C_2\left(-1 + \frac{\lambda}{3}h + \mathcal{O}\left(h^2\right)\right)^{x_k/h}. \qquad (3.25)$$

The first term converges to $\exp(\lambda x_k)$ for $h \to 0$, since

$$\lim_{h\to 0}\,(1 + \lambda h)^{x_k/h} = \lim_{h\to 0}\left(1 + \lambda x_k\frac{h}{x_k}\right)^{x_k/h} = \exp(\lambda x_k).$$

It behaves like the solution of the model initial value problem. The second term behaves for small h like

$$(-1)^{x_k/h}\left(1 - \frac{\lambda}{3}h\right)^{x_k/h}.$$

Here, the second factor converges to $\exp(-\lambda x_k/3)$, but the first factor oscillates for $x_k/h \in \mathbb{N}$. That means, for the stable initial value problem with $\lambda < 0$, this term gives an oscillatory, bounded (for fixed x_k), but exponentially large perturbation.

The behavior of the solution depends on the constants C_1 and C_2. Inserting the expansion (3.24) in condition (3.23) gives

$$\begin{aligned}
e^{\lambda h} &= (1 - C_2)\left(1 + \lambda h + \mathcal{O}\left(h^2\right)\right) + C_2\left(-1 + \frac{\lambda}{3}h + \mathcal{O}\left(h^2\right)\right)\\
&= 1 + \lambda h - 2C_2 - \frac{2\lambda h}{3}C_2 + \mathcal{O}\left(h^2\right).
\end{aligned}$$

An expansion of the exponential yields

$$\begin{aligned}
1 + \lambda h + \mathcal{O}\left(h^2\right) &= 1 + \lambda h - 2C_2 - \frac{2\lambda h}{3}C_2 + \mathcal{O}\left(h^2\right) \quad \Longrightarrow\\
\mathcal{O}\left(h^2\right) &= -2C_2 - \frac{2\lambda h}{3}C_2.
\end{aligned}$$

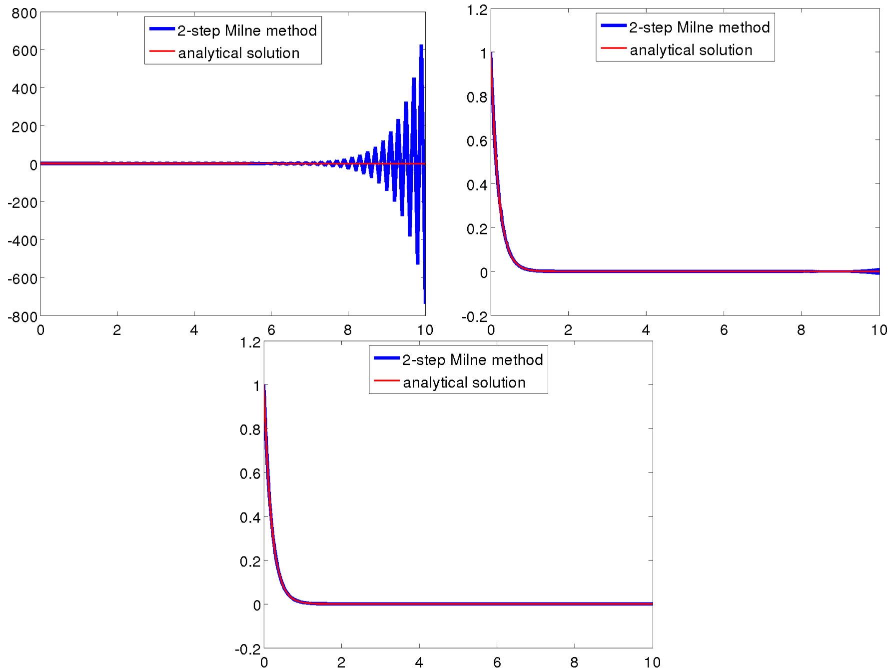

Fig. 3.2 Example 3.22: application of the 2-step Milne method to the model problem with $\lambda = -5$ and $h \in \{0.1, 0.01, 0.001\}$ (left to right, top to bottom).

It follows that $C_2(h) = \mathcal{O}\left(h^2\right)$ and from (3.22), it follows that $C_1 = \mathcal{O}(1)$. In summary, it is for the second term of (3.25)

$$\lim_{h\to 0} C_2(h) \underbrace{\left(-1 + \frac{\lambda}{3}h + \mathcal{O}\left(h^2\right)\right)^{x_k/h}}_{\text{bounded}} = 0.$$

The method converges.

However, the term

$$C_2(h)\left(-1 + \frac{\lambda}{3}h + \mathcal{O}\left(h^2\right)\right)^{x_k/h} \approx \pm h^2 \exp\left(-\frac{\lambda x_k}{3}\right)$$

becomes small in the case $\lambda \ll -1$ and large x_k only if the step size h is very small, see Figure 3.2.

The behavior found for this method can be observed in practice for all q-step methods of consistency order $q+2$ if these methods are applied to initial value problems with exponentially decaying solution. This kind of instability is a strong restriction of the usefulness of these methods. $\square$

Remark 3.23. Start of multi-step methods and convergence. Apart of the consistency of multi-step methods, one is above all interested in their convergence. For one-step methods, convergence follows from consistency under rather general assumptions and the order of consistency and convergence are the same, see Theorem 1.19. The situation becomes more complicated for multi-step methods.

First of all, one needs for starting a q-step method besides the known initial value $y_0 = y(x_0)$ still $(q-1)$ further approximations $y_1, \ldots, y_{q-1}$ for $y(x_1), \ldots, y(x_{q-1})$. These values can be computed, for instance by a one-step method. The accuracy of these approximations has a strong impact on the accuracy of the q-step method that uses these values. Assume that the approximations behave as follows

$$y_0 = y(x_0) + \varepsilon_0(h), \quad y_1 = y(x_1) + \varepsilon_1(h), \quad \ldots, \quad y_{q-1} = y(x_{q-1}) + \varepsilon_{q-1}(h).$$

Then, the values that are computed with the q-step method depend also on the perturbations[11] $\varepsilon_1(h), \ldots, \varepsilon_{q-1}(h)$ and one should write for the computed solution in the node x_k more exactly $y_k(\varepsilon, h)$, where $\varepsilon(x, h)$ is a function for which $\varepsilon_i(h) = \varepsilon(x_i, h)$, $i = 1, \ldots, q-1$, holds. □

Definition 3.24. Global error. Let $y(x)$ be the solution of the initial value problem (3.1). Denote the approximations of $y(x)$ that are computed with a multi-step method with step length h by $y_k(\varepsilon, h)$, where the accuracy of the initial approximations is given by the function $\varepsilon(x, h)$. Then, the quantity

$$\mathrm{e}(x_k, \varepsilon, h) := y_k(\varepsilon, h) - y(x_k)$$

is called global error or global discretization error at the node x_k with respect to the step length h and the perturbations $\varepsilon(x, h)$. □

Definition 3.25. Convergence of a multi-step method. Consider the ordinary differential equation of the initial value problem (3.1) in $[a, b]$ and let $x_0 \in [a, b]$. A multi-step method for solving initial value problems of form (3.1) is called convergent if

$$\lim_{n\to\infty} \mathrm{e}(x, \varepsilon, h_n) = 0, \quad \text{with } h_n = \frac{x - x_0}{n},$$

for all $x \in [a, b]$, for all $f \in C^1([a, b] \times \mathbb{R})$, and for all functions $\varepsilon(x, h)$ with

$$\lim_{n\to\infty} |\varepsilon(x, h_n)| = 0, \quad \text{for } x = x_0 + ih_n,\ i = 0, \ldots, q-1.$$

□

[11] This is meant in the sense that strictly speaking, it is not complete to say, e.g., 'the value at $x_k = 1$ obtained with the Adams-Bashforth method is ...', but one should say 'the value at $x_k = 1$ obtained with the Adams-Bashforth method, where y_1 was computed in this and this way, is ...'.

Lemma 3.26. A convergent linear multi-step methods is null stable. *A convergent linear multi-step method* (3.2) *is null stable.*

Proof. The proof is performed by contradiction. Assume that the linear multi-step method is convergent but not null stable. Consider the initial value problem

$$y'(x) = 0, \quad y(0) = 0,$$

whose solution is $y(x) = 0$. Applying a linear multi-step method of form (3.2) to this problem yields the homogeneous linear difference equation

$$y_{k+1} - \sum_{j=0}^{q-1} a_j y_{k-j} = 0. \tag{3.26}$$

Since the method is assumed to be not null stable, the corresponding first characteristic polynomial $\Psi(\xi)$ has a root with $|\xi_1| > 1$ or a root $|\xi_2| = 1$ that is not simple. Without loss of generality, let the multiplicity of ξ_1 be one and of ξ_2 be two. Similarly to Example 3.17, one finds that the solution of (3.26) in the node $x_k = kh$ is given by

$$y_k = C_1\xi_1^k + C_2 k\xi_2^k, \quad C_1, C_2 \in \mathbb{R},$$

where one of these coefficients is not zero.

Consider a fixed $\overline{x}$ with $\overline{x} = mh$, $m \in \mathbb{N}$. Choosing $C_1 = C_2 = \sqrt{h}$, where it will be discussed below that this is an admissible choice, so that the perturbations of the initial values satisfy the requirements of Definition 3.25, the solution in $\overline{x}$ is given by

$$\sqrt{h}\xi_1^{\overline{x}/h} + \frac{\overline{x}}{\sqrt{h}}\xi_2^{\overline{x}/h}. \tag{3.27}$$

For the initial value $\overline{x} = 0$, the value of (3.27) is $\sqrt{h}$ and for the initial value $\overline{x} = h$, it is $\sqrt{h}\xi_1 + \sqrt{h}\xi_2$. Thus, for the initial values, (3.27) converges to the analytic solution as $h \to 0$, so that the choices of C_1 and C_2 were admissible. However, for other values of $\overline{x}$, which is assumed to be fix, both terms in (3.27) diverge to plus infinity. This observation contradicts the assumed convergence of the linear multi-step method. Hence, it is null stable. ■

Theorem 3.27. Connection of convergence and null stability. *Let*

$$y_{k+1} = \sum_{j=0}^{q-1} a_j y_{k-j} + h\Phi(x_{k+1}, \ldots, x_{k+1-q}, y_{k+1}, \ldots, y_{k+1-q}, h) \tag{3.28}$$

be a consistent multi-step method for the solution of initial value problems of form (3.1)*, which is more general than a linear multi-step method. Assume that the incremental function satisfies the following conditions:*

i) $\Phi(x_{k+1}, \ldots, x_{k+1-q}, y_{k+1}, \ldots, y_{k+1-q}, h) \equiv 0$ *for all* $x \in [a, b]$*, all* $y_k \in \mathbb{R}$*, and all* $h \in \mathbb{R}$ *if* $f(x, y) \equiv 0$.

ii) Lipschitz continuity with respect to the y-components, i.e., there are constants $h_0 > 0$ *and* M *such that*

$$\left|\Phi(x_q, \ldots, x_0, v_q, \ldots, v_0, h) - \Phi(x_q, \ldots, x_0, w_q, \ldots, w_0, h)\right|$$

$$\leq M \sum_{i=0}^{q} |v_i - w_i|$$

for all $x_q, \ldots, x_0 \in [a, b]$, *all* $v_i, w_i \in \mathbb{R}$, $i = 0, \ldots, q$, *and all step sizes* h *with* $h < h_0$.

Then, the multi-step method converges if and only if it is null stable.

Proof. For the proof, it is referred to the literature, e.g., (Strehmel *et al.*, 2012, Section 4.2.5). ■

Remark 3.28. To Theorem 3.27.

- The first assumption and the null stability guarantee that the multi-step method solves the trivial initial value problem

$$y'(x) = 0, \quad y(x_0) = 0,$$

exactly if $\varepsilon_0 = \varepsilon_1 = \ldots = \varepsilon_{q-1} = 0$.

- A linear multi-step method is a special case of (3.28). For linear multi-step methods, the first assumption is always satisfied, since the incremental function is a linear combination of values of the right-hand side $f(x, y)$ of the ordinary differential equation. Due to the same reason, the incremental function of these methods satisfies the second assumption if the right-hand side $f(x, y)$ is Lipschitz continuous with respect to the second argument. Altogether, if the right-hand side of the initial value problem is sufficiently smooth, then a consistent linear multi-step method is convergent if and only if it is null stable.

□

Theorem 3.29. Order of convergence. *Consider a multi-step method of the form* (3.28) *that satisfies the assumptions stated in Theorem 3.27 and which possesses the order of consistency* p. *Then, it holds for all* $f \in C^p([a, b] \times \mathbb{R})$ *and for all* $x \in [a, b]$ *that*

$$|\mathrm{e}(x, \varepsilon, h)| = \mathcal{O}(h^p),$$

if for the accuracy of the initial values it holds

$$|\varepsilon_i(h)| = \mathcal{O}(h^p) \quad \text{for} \quad i = 0, \ldots, q-1.$$

Proof. See literature, e.g., (Strehmel *et al.*, 2012, Section 4.2.5) or (Hairer *et al.*, 1993, Chapter III.4). ■

Remark 3.30. Interpretation of Theorem 3.29. If a multi-step method with consistency order p should also have convergence order p, then it is necessary to compute the initial approximations sufficiently accurately, e.g., with a one-step method of order p. Considering the complete method, which consists of the starting method for computing the approximations $y_1, \ldots, y_{q-1}$ and a predictor-corrector method for computing the other values, then the order of the complete method is determined by the partial method with the lowest order. □

3.4 Backward Difference Formula (BDF) Methods

Remark 3.31. Construction. The construction of Backward Difference Formula (BDF) methods is based on the original initial value problem (3.1) and not on the integral form (3.3) as it is the case for predictor-corrector methods.

Given $q+1$ nodes $x_{k+1-q},\ldots,x_{k+1}$ and $q \geq 1$ known approximations of the solution $y_{k+1-q},\ldots,y_k$. Then, the idea of BDF methods consists in approximating the solution by an interpolation polynomial $p_q(x)$ of degree q with the nodes $\left(x_{k+1-q},y_{k+1-q}\right),\ldots,(x_k,y_k)$. Now, another condition is needed in order to define a polynomial of degree q and this condition shall also allow to compute y_{k+1}. For BDF methods, one uses the requirement that this polynomial should satisfy the differential equation (3.1) in x_{k+1}, i.e.,

$$p_q'\left(x_{k+1}\right) = f\left(x_{k+1},y_{k+1}\right), \tag{3.29}$$

which leads to an interpolation of Hermite type. It follows from this requirement that BDF methods are implicit methods. □

Example 3.32. BDF methods. Consider an equidistant grid with grid size h.

- $q=1$. The linear interpolation polynomial through the points (x_k,y_k) and (x_{k+1},y_{k+1}) is given by the Newton representation (using divided differences)
$$p_1(x) = y_{k+1} + (x-x_{k+1})\frac{y_k-y_{k+1}}{x_k-x_{k+1}}.$$
It is
$$p_1'(x) = \frac{y_k-y_{k+1}}{x_k-x_{k+1}}$$
such that requirement (3.29) and $x_k - x_{k+1} = -h$ leads to
$$\frac{y_k-y_{k+1}}{-h} = f\left(x_{k+1},y_{k+1}\right) \quad \Longleftrightarrow \quad y_{k+1} = y_k + hf\left(x_{k+1},y_{k+1}\right).$$
Hence, BDF(1) is just the implicit Euler method.
- $q=2$. The Newton representation of the quadratic interpolation polynomial through $(x_{k-1},y_{k-1}),(x_k,y_k),(x_{k+1},y_{k+1})$ is given by
$$\begin{aligned} p_2(x) &= y_{k+1} + (x-x_{k+1})\frac{y_k-y_{k+1}}{x_k-x_{k+1}} \\ &\quad + \frac{(x-x_{k+1})(x-x_k)}{x_{k-1}-x_{k+1}}\left(\frac{y_{k-1}-y_k}{x_{k-1}-x_k} - \frac{y_k-y_{k+1}}{x_k-x_{k+1}}\right). \end{aligned} \tag{3.30}$$

Computing the derivative of this polynomial and using that the grid is equidistant yields

$$p_2'(x) = \frac{y_k - y_{k+1}}{-h} + \frac{(x - x_{k+1}) + (x - x_k)}{-2h}\left(\frac{y_{k-1} - y_k}{-h} - \frac{y_k - y_{k+1}}{-h}\right),$$

such that requirement (3.29) leads to

$$p_2'(x_{k+1}) = \frac{y_{k+1} - y_k}{h} + \frac{h}{2h}\left(\frac{y_{k+1} - 2y_k + y_{k-1}}{h}\right) = f\left(x_{k+1}, y_{k+1}\right).$$

Collecting terms gives the BDF(2) method

$$\frac{3}{2}y_{k+1} - 2y_k + \frac{1}{2}y_{k-1} = hf\left(x_{k+1}, y_{k+1}\right). \tag{3.31}$$

BDF(2) is the most popular multi-step method for stiff problems.
- $q \geq 3$. The derivation of higher order methods proceeds in the same way. One obtains, e.g., as BDF(3) method

$$\frac{11}{6}y_{k+1} - \frac{18}{6}y_k + \frac{9}{6}y_{k-1} - \frac{2}{6}y_{k-2} = hf\left(x_{k+1}, y_{k+1}\right). \tag{3.32}$$

It should be emphasized that in BDF methods the right-hand side of the initial value problem appears only in one term, namely $f\left(x_{k+1}, y_{k+1}\right)$. This situation is in contrast to the predictor-corrector methods from Section 3.2. This property of BDF methods is of advantage if the computation of the right-hand side is complicated or numerically expensive, like for special discretizations of partial differential equations. □

Lemma 3.33. Null stability of BDF(1), BDF(2), and BDF(3). *The methods BDF(1), BDF(2), and BDF(3) are null stable.*

Proof. The statement of the lemma is obtained by computing the roots of the first characteristic polynomial.
- $q = 1$. The characteristic polynomial is $\lambda - 1$ with the root $\lambda_1 = 1$.
- $q = 2$. The characteristic polynomial of BDF(2) (3.31) is

$$\lambda^2 - \frac{4}{3}\lambda + \frac{1}{3}.$$

A straightforward calculation gives

$$\lambda_1 = \frac{2}{3} + \frac{1}{3} = 1, \quad \lambda_2 = \frac{2}{3} - \frac{1}{3} = \frac{1}{3}.$$

- $q = 3$. For BDF(3), see (3.32), one obtains the characteristic polynomial

$$\lambda^3 - \frac{18}{11}\lambda^2 + \frac{9}{11}\lambda - \frac{2}{11}.$$

By inserting, one checks that $\lambda_1 = 1$ is a root of this polynomial. Extracting the linear factor with this root yields

$$\frac{\lambda^3 - \frac{18}{11}\lambda^2 + \frac{9}{11}\lambda - \frac{2}{11}}{\lambda - 1} = \lambda^2 - \frac{7}{11}\lambda + \frac{2}{11}.$$

The remaining roots are given by the roots of the quadratic polynomial, which are

$$\lambda_2 = \frac{7 + i\sqrt{39}}{22}, \quad \lambda_3 = \frac{7 - i\sqrt{39}}{22},$$

such that $|\lambda_2| = |\lambda_3| = \sqrt{22}/11 \approx 0.4264$.

■

Remark 3.34. Null stability of BDF(q) methods. It can be shown that BDF(q) methods are null stable only for $q \le 6$, e.g., see Cryer (1972). Hence, BDF(q) methods for $q > 6$ are not of interest. □

Lemma 3.35. Consistency of BDF(q) methods. *BDF(q) methods with $q \le 6$ are consistent of order q.*

Proof. The proof is obtained by a Taylor series expansion (*exercise*). ■

Theorem 3.36. Convergence of BDF(q) methods. *Let $f \in C^q([a,b]\times\mathbb{R})$ and Lipschitz continuous with respect to the second argument and let the initial values be computed sufficiently accurately, then the BDF(q) methods with $q \le 6$ are convergent of order q.*

Proof. The incremental function of BDF(q) methods is

$$\Phi(x_{k+1}, \ldots, x_{k+1-q}, y_{k+1}, \ldots, y_{k+1-q}, h) = f\left(x_{k+1}, y_{k+1}\right),$$

so that the assumptions of Theorem 3.27 are satisfied. Because BDF(q) methods with $q \le 6$ are null stable and consistent of order q, the other assumptions of Theorem 3.29 are also satisfied and the statement of the theorem follows now from Theorem 3.29. ■

Remark 3.37. On the stability. Stability of multi-step methods is studied at the same initial value problem (2.6) as it was used for one-step methods. Assuming that all initial values are computed exactly, one obtains in the same way as in Example 3.17 a homogeneous difference equation

$$y_{k+1} - \sum_{j=0}^{q-1} a_j y_{k-j} = z \sum_{j=-1}^{q-1} b_j y_{k-j}$$

with $z = \lambda h$. With the ansatz $y_k = \xi^k$ and after division by ξ^{k+1-q}, one obtains a characteristic equation

$$\Psi(\xi) - z\sigma(\xi) = 0, \tag{3.33}$$

compare (3.16), where $\Psi(\xi)$ is the first characteristic polynomial (3.18). The polynomial σ has the coefficients b_j, compare (3.19).

Note that for $z = 0$, only the roots of the first characteristic polynomial are considered, which are important for the null stability of the method. This relation might be the reason for the notion 'null' stable. □

Table 3.1 Values of α (in degree) for the A(α)-stability of BDF(q) methods.

q	1	2	3	4	5	6
α	90	90	86.03	73.35	51.84	17.84

Definition 3.38. Stability domain. The set

$$S = \big\{ z \in \mathbb{C} \ : \ \text{for all roots } \xi_l \text{ of (3.33) it holds } |\xi_l| \le 1; \\ \text{if } \xi_l \text{ is a multiple root, then it holds } |\xi_l| < 1 \ \big\}$$

is called stability domain of a linear multi-step method. □

Definition 3.39. A-stability, $A(\alpha)$-stability. A linear multi-step method is called A-stable if $\mathbb{C}^- \subset S$. It is called $A(\alpha)$-stable with $\alpha \in (0, \pi/2)$ if

$$\big\{ z \in \mathbb{C}^- \text{ with } |\arg(z) - \pi| \le \alpha \big\} \subset S,$$

with $\arg(z) \in [0, 2\pi)$. □

Theorem 3.40. Second Dahlquist barrier. *An A-stable linear multi-step method is at most of second order.*

Proof. See literature, e.g., (Strehmel *et al.*, 2012, Section 9.1). ■

Remark 3.41. A(α)-stability of BDF(q) methods. BDF(q) methods are $A(\alpha)$-stable for $q \le 6$ and even A-stable for $q \le 2$. The values of α for BDF(q) methods are given in Table 3.1. Because of the small value of α for $q = 6$, the method BDF(6) is not used in practice. □

Remark 3.42. Variable step size for BDF(q) methods. BDF(q) methods can be used on non-equidistant grids, e.g., a formula for BDF(2) with variable step size can be derived on the basis of the quadratic interpolation polynomial (3.30), *exercise problem.* For $q > 1$ there is some restriction on the admissible change of the mesh size from one mesh cell to its neighbor, e.g., for BDF(2) stability is guaranteed as long as $h_{k+1}/h_k \le 2.41421$, see (Strehmel *et al.*, 2012, p. 328) for more details. □