Chapter 3

Multi-Step Methods

3.1 Generalities

Remark 3.1 *Multi-step methods.* The characteristic feature of one-step is that they need for computing y_{k+1} only the value from the previous approximation of the solution y_k. 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

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

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)

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

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

$$y(x) = y(\underline{x}) + \int_{\underline{x}}^{x} f(t, y(t)) \; dt. \tag{3.3}$$

The main idea of predictor-corrector methods consists in approximating the integral on the right hand side of (3.3) 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 is the dependency of the function in the integral on t is known, generally it will not be possible to find an analytic expression of the solution.

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 (3.3) is replaced by a polynomial interpolant. For the construction of this polynomial, one uses equidistant nodes

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

Let in particular the boundaries of the integral be the nodes

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

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 solution of this interpolation problem is given by the Lagrange[3] interpolation polynomial

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

with

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

It follows by using (3.3) that

$$y_{p+m} \approx y_{p-j} + \sum_{i=0}^{r} f(x_{p-i}, y(x_{p-i})) \int_{\underline{x}}^{x} L_i(t) \; dt = y_{p-j} + h \sum_{i=0}^{r} \beta_i f(x_{p-i}, y(x_{p-i})) \tag{3.4}$$

with

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

The constructed method is in particular linear. Using the substitution

$$t = x_p + sh = x_0 + (p+s)h,$$

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

and the equidistant mesh yields

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

Now, different methods can be obtained, depending on the choice of m, j, and r and by replacing $y(x_{p-i})$ in (3.4) with y_{p-i}. 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

$$y_{k+1} = y_k + h\sum_{i=0}^{q-1}\beta_i f\left(x_{k-i}, y_{k-i}\right) \tag{3.5}$$

with

$$\beta_i = \int_0^1\left(\prod_{l=0,l\neq i}^{q-1}\frac{s+l}{-i+l}\right)ds. \tag{3.6}$$

In the case $q = 1$, the term in the integral in (3.3) is replaced by a constant interpolation polynomial with the node $(x_k, f(x_k, y_k))$. 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 (3.5) and (3.6), one obtains

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

$q \geq 3$, *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

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

[4] John Couch Adams (1819 – 1892)
[5] Francis Bashforth (1819 – 1912)
[6] Forest Ray Moulton (1872 – 1952)

and

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

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 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.

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

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

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

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

This method is the trapezoidal rule. □

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

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

□

[7] Evert J. Nyström (1895 – 1960)

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} a generally nonlinear equation. 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 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, the step length control for form (3.2) of multi-step methods becomes rather complicated. □

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

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

[8]William Edwin Milne (1890 – 1971)

Then, the local error is defined by

$$\mathrm{le}(x_{k+1}) = \mathrm{le}_{k+1} = y(x_{k+1}) - \left[\sum_{j=0}^{q-1} a_j y\left(x_{k-j}\right) + h \sum_{j=-1}^{q-1} b_j f\left(x_{k-j}, y(x_{k-j})\right)\right]. \quad (3.7)$$

□

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

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

If the expression on the left hand side converges for sufficiently small step sizes h 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 $q = 3$

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

It follows with (3.7) and (3.1) that

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

Now, the the individual terms will be expanded

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

Inserting these expressions into the formula for the local error gives

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

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 high order methods. On the first glance, this looks rather simple. If one replaces the solution in the definition of the local error by its numerical approximation, and if one requires that the right hand side of this equation vanishes, one obtains the following ansatz for the method

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

Using now the Taylor series expansion of the local error, one gets a linear system of equations for determining the coefficients a_j, b_j, $j = 0, \ldots, q-1$ and b_{-1}. Constructing one-step methods in this way, one always obtains a convergent one-step method, e.g., compare Example 1.29. However, for multi-step methods the situation might be different. □

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 convergence. That means, the ansatz for the method is as follows

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

The local error has the form

$$
\begin{aligned}
\mathrm{le}(x_{k+1}) &= y\left(x_{k+1}\right) - a_0 y\left(x_k\right) - a_1 y\left(x_{k-1}\right) - hb_0 f\left(x_k, y(x_k)\right) \\
&\quad - hb_1 f\left(x_{k-1}, y(x_{k-1})\right) \\
&= y\left(x_{k+1}\right) - a_0 y\left(x_k\right) - a_1 y\left(x_{k-1}\right) - hb_0 y'(x_k) - hb_1 y'(x_{k-1}).
\end{aligned}
$$

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

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

Inserting the expansions gives

$$
\begin{aligned}
\mathrm{le}(x_{k+1}) &= y\left(x_k\right) + hy'\left(x_k\right) + \frac{h^2}{2} y''\left(x_k\right) + \frac{h^3}{6} y'''\left(x_k\right) - a_0 y\left(x_k\right) \\
&\quad - a_1 \left[y\left(x_k\right) - hy'\left(x_k\right) + \frac{h^2}{2} y''\left(x_k\right) - \frac{h^3}{6} y'''\left(x_k\right) \right] - hb_0 y'(x_k) \\
&\quad - hb_1 \left[y'\left(x_k\right) - hy''\left(x_k\right) + \frac{h^2}{2} y'''\left(x_k\right) \right] + \mathcal{O}(h^4) \\
&= \left[1 - a_1 - a_0\right] y\left(x_k\right) + \left[1 + a_1 - b_1 - b_0\right] hy'\left(x_k\right) \\
&\quad + \left[\frac{1}{2} - \frac{a_1}{2} + b_1 \right] h^2 y''\left(x_k\right) + \left[\frac{1}{6} + \frac{a_1}{6} - \frac{b_1}{2} \right] h^3 y'''\left(x_k\right) + \mathcal{O}(h^4).
\end{aligned}
$$

Thus, one obtains the following linear system of equations, if one requires that the first four terms should vanish

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

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

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

with third order of consistency.

Next, the stability 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., $y_1 = exp(-h)$. Due to the special form of the right hand side of the model problem, $f(x_k, y_k) = -y_k$, one can represent the computed solution explicitly. This solution fulfills 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 $y_k = \lambda^k$. Inserting this ansatz into the difference equation leads to

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

This equation is satisfied for $\lambda = 0$. Other solutions can be obtained after division by λ^{k-1} from

$$\lambda^2 + (4+4h)\lambda + (-5+2h) = 0. \tag{3.8}$$

One gets the solutions

$$\lambda_1(h) = -2 - 2h + 3\sqrt{1+\frac{2}{3}h+\frac{4}{9}h^2}, \quad \lambda_2(h) = -2 - 2h - 3\sqrt{1+\frac{2}{3}h+\frac{4}{9}h^2}.$$

Using the ansatz, the general solution of the difference equations can be represented as a linear combination of the special solutions (superposition principle)

$$y_k = c_1\lambda_1^k + c_2\lambda_2^k.$$

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

$$y_0 = c_1 + c_2 = 1, \quad y_1 = e^{-h} = c_1\lambda_1 + c_2\lambda_2,$$

from what follows that

$$c_1(h) = \frac{e^{-h}-\lambda_2}{\lambda_1-\lambda_2}, \quad c_2(h) = -\frac{e^{-h}-\lambda_1}{\lambda_1-\lambda_2}.$$

Expanding $\lambda_1(h)$, $\lambda_2(h)$, $c_1(h)$ and $c_2(h)$ in powers of h and inserting these expansions into the solution *(exercise)*, gives for fixed $x > 0$ and $h_N := x/N$

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

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 describes 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

$$-\frac{(-5)^N}{216}\left(\frac{x}{N}\right)^4 e^{3x/5}. \tag{3.9}$$

This expression oscillates with increasing N and the modulus is increasing for finer grids, compare the values for $x = 1$ in the following table

N	value of (3.9)
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. The reason for this behavior is that the general solution of the difference equation contains a term which becomes arbitrarily large for large k, e.g., it holds that

$$\lim_{h\to 0} \lambda_2(h) = -5 \quad \Longrightarrow \quad \lim_{k\to\infty} \left|\lambda_2^k(h)\right| = \infty$$

for small h or large N.

The solution of the difference equation was obtained from the roots of the polynomial (3.8). It can be guessed that the roots of this polynomial will be of importance for the convergence of the multi-step method. □

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

$$\Psi(\lambda) = \lambda^q - a_0\lambda^{q-1} - \ldots - a_{q-1}$$

possesses only roots λ_q with $|\lambda_q| \le 1$ which are simple in the case that $|\lambda_q| = 1$. □

Example 3.19 *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 characteristic polynomial has the form

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

  The only non-trivial root is $\lambda_q = 1$.
- *Nyström methods, Milne methods.* For these methods, the characteristic polynomial is

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

  Hence, the only non-trivial roots are $\lambda_q = 1$ and $\lambda_q = -1$.

□

Theorem 3.20 First Dahlquist[9] barrier. *The maximal order of consistency of a null stable linear q-step method is*

$$p = \begin{cases} q+1 & \text{for } q \text{ odd,} \\ q+2 & \text{for } q \text{ even.} \end{cases}$$

[9]Germund Dahlquist (1925 – 2005)

Proof: See literature, e.g., (Strehmel et al., 2012, Section 4.2.3). ■

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 $q+1$.
- The 2-step Milne method (also Milne–Simpson method)

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

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

□

Example 3.22 Stability of the 2-step Milne method. This 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

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

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

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

This equation can be rewritten as a linear difference equation

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

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

$$y_k = c_1\lambda_1^k + c_2\lambda_2^k,$$

where $\lambda_1(h)$ and $\lambda_2(h)$ are the solutions of the quadratic equation

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

One obtains

$$\begin{aligned}\lambda_1(h) &= \frac{3}{3 - h\kappa}\left(\frac{2h\kappa}{3} + \sqrt{1 + \frac{(h\kappa)^2}{3}}\right),\\ \lambda_2(h) &= \frac{3}{3 - h\kappa}\left(\frac{2h\kappa}{3} - \sqrt{1 + \frac{(h\kappa)^2}{3}}\right).\end{aligned}$$

Now, the constants c_1, c_2 can be determined from the initial condition and from the value after the first step. It is for $x = 0$

$$c_1 + c_2 = 1 \tag{3.10}$$

and for $x = h$

$$e^{\kappa h} = c_1\lambda_1 + (1 - c_1)\lambda_2. \tag{3.11}$$

Expanding $\lambda_1(h)$ and $\lambda_2(h)$ in powers of h at $h = 0$, one obtains as first order approximation by computing the derivative and inserting zero

$$\lambda_1(h) = 1 + \kappa h + \mathcal{O}\left(h^2\right), \quad \lambda_2(h) = -1 + \frac{\kappa}{3}h + \mathcal{O}\left(h^2\right). \tag{3.12}$$

In the interesting case, $\kappa < 0$, $\lambda_1(h)$ approaches to 1 from left, with values smaller than 1, and $\lambda_2(h)$ approaches to -1 also from left, but here the modulus of $\lambda_2(h)$ 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

$$y_k = c_1\left(1 + \kappa h + \mathcal{O}\left(h^2\right)\right)^{x_k/h} + c_2\left(-1 + \frac{\kappa}{3}h + \mathcal{O}\left(h^2\right)\right)^{x_k/h}. \tag{3.13}$$

The first term converges for $h \to 0$ to $\exp(\kappa x_k)$. It behaves like the solution of the model initial problem. The second term behaves for small h as

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

Here, the second factor converges to $\exp(-\kappa x_k/3)$, but the first factor oscillates the faster the smaller h. That means, for the stable initial value problem with $\kappa < 0$, 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 (3.12) into the condition (3.11) and using (3.10) gives

$$\begin{aligned} e^{\kappa h} &= (1 - c_2)\left(1 + \kappa h + \mathcal{O}\left(h^2\right)\right) + c_2\left(-1 + \frac{\kappa}{3}h + \mathcal{O}\left(h^2\right)\right) \\ &= 1 + \kappa h - 2c_2 - \frac{2\kappa h}{3}c_2 + \mathcal{O}\left(h^2\right). \end{aligned}$$

An expansion of the exponential yields

$$\begin{aligned} 1 + \kappa h + \mathcal{O}\left(h^2\right) &= 1 + \kappa h - 2c_2 - \frac{2\kappa h}{3}c_2 + \mathcal{O}\left(h^2\right) \quad \Longrightarrow \\ \mathcal{O}\left(h^2\right) &= -2c_2 - \frac{2\kappa h}{3}c_2 + \mathcal{O}\left(h^2\right). \end{aligned}$$

It follows that $c_2(h) = \mathcal{O}\left(h^2\right)$ and from (3.10) it follows that $c_1 = \mathcal{O}(1)$. In summary, it is for the second term of (3.13)

$$\lim_{h\to 0} c_2(h)\left(-1 + \frac{\kappa}{3}h + \mathcal{O}\left(h^2\right)\right)^{x_k/h} = 0.$$

The method converges.

However, the term

$$h^2 \exp -\frac{\kappa x_k}{3}$$

becomes small in the case $\kappa \ll -1$ and large x_k only if the step size h is very small, see Figure 3.1.

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 property is a strong restriction of the usefulness of these methods. □

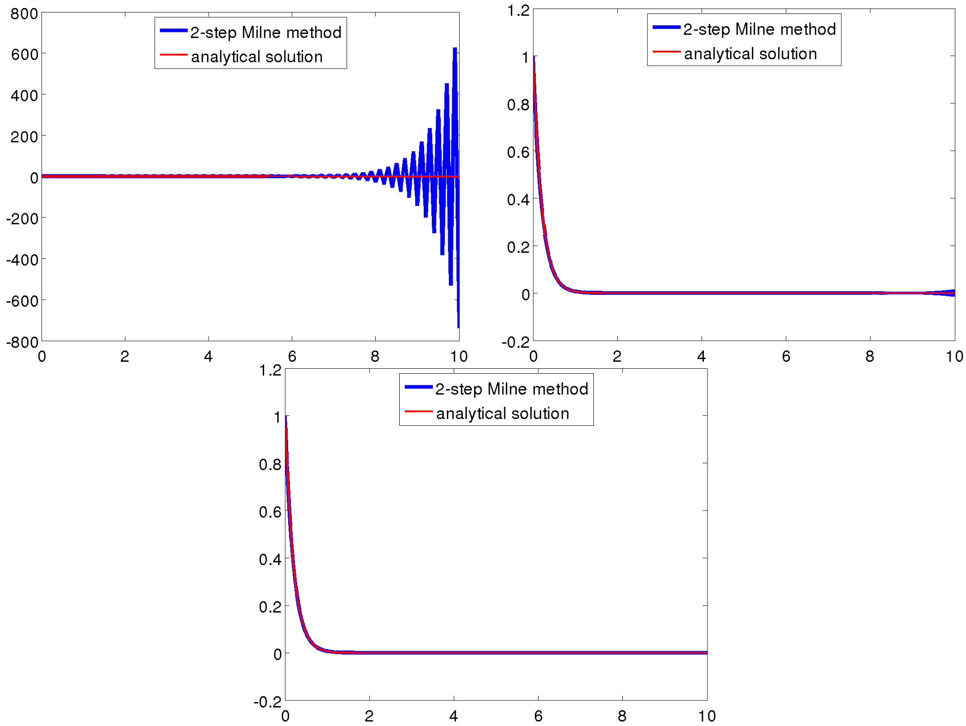

Figure 3.1: Example 3.22: application of the 2-step Milne method to the model problem with $\kappa = -5$ and $h \in \{0.1, 0.01, 0.001\}$ (left to right, top to bottom).

Remark 3.23 *Start of multi-step methods and convergence.* Apart of the consistency of multi-step, 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 influence on the accuracy of the q-step method that uses these approximations. Assume that the approximations behave as follows

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

Then, the values which are computed with the q-step method depend also on the perturbations $\varepsilon_1(h), \ldots, \varepsilon_{q-1}(h)$ 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 $y(x)$ which 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. □

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

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

□

Theorem 3.26 Connection of convergence and null stability. *Let*

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

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

i) $\Phi(x_j, y_{k+1}, \ldots, y_{k+1-q}, h, f) \equiv 0$ *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 of the 2^{nd} *till* $(q + 1)^{th}$ *component, i.e., there are constants* $h_0 > 0$ *and* M *such that*

$$|\Phi(x, v_q, \ldots, v_0, h, f) - \Phi(x, w_q, \ldots, w_0, h, f)| \le M \sum_{i=0}^{q} |v_i - w_i|$$

for all $x \in [a, b]$*, 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: See literature, e.g., (Strehmel et al., 2012, Section 4.2.5). ■

Remark 3.27 *To Theorem 3.26.*

- 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 $\varepsilon_1 = \varepsilon_2 = \ldots = \varepsilon_{q-1} = 0$.
- 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 fulfills the second assumption if the right hand side $f(x, y)$ is partially Lipschitz continuous with respect to the second argument.

□

Theorem 3.28 Order of convergence. *Consider a multi-step method of the form given in Theorem 3.26 that satisfies the assumptions stated in this theorem 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 holds

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

Proof: See literature, e.g., (Strehmel et al., 2012, Section 4.2.5). ■

Remark 3.29 *Interpretation of Theorem 3.28.* 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 accurate, 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. □