2 Multistep Methods

Up to now, all methods we studied were single step methods, i.e., the value $\boldsymbol{y}_{n+1}$ was found using information only from the previous time level t_n. Now we will consider so-called *multistep methods*, i.e., more of the history of the solution will affect the value $\boldsymbol{y}_{n+1}$.

2.1 Adams Methods

Consider the first-order ODE

$$\boldsymbol{y}'(t) = \boldsymbol{f}(t, \boldsymbol{y}(t)).$$

If we integrate from t_{n+1} to t_{n+2} we have

$$\int_{t_{n+1}}^{t_{n+2}} \boldsymbol{y}'(\tau) d\tau = \int_{t_{n+1}}^{t_{n+2}} \boldsymbol{f}(\tau, \boldsymbol{y}(\tau)) d\tau$$

or

$$\boldsymbol{y}(t_{n+2}) - \boldsymbol{y}(t_{n+1}) = \int_{t_{n+1}}^{t_{n+2}} \boldsymbol{f}(\tau, \boldsymbol{y}(\tau)) d\tau. \tag{31}$$

As we saw earlier for Euler's method and for the trapezoidal rule, different numerical integration rules lead to different ODE solvers. In particular, the left-endpoint rule yields Euler's method, while the trapezoidal rule for integration gives rise to the trapezoidal rule for IVPs. Incidentally, the right-endpoint rule provides us with the backward Euler method.

We now use a different quadrature formula for the integral in (31).

Example Instead of viewing the slope $\boldsymbol{f}$ as a constant on the interval $[t_n, t_{n+1}]$ we now represent $\boldsymbol{f}$ by its *linear interpolating polynomial* at the points $\tau = t_n$ and $\tau = t_{n+1}$ given in Lagrange form, i.e.,

$$\begin{aligned} \boldsymbol{p}(\tau) &= \frac{\tau - t_{n+1}}{t_n - t_{n+1}} \boldsymbol{f}(t_n, \boldsymbol{y}(t_n)) + \frac{\tau - t_n}{t_{n+1} - t_n} \boldsymbol{f}(t_{n+1}, \boldsymbol{y}(t_{n+1})) \\ &= \frac{t_{n+1} - \tau}{h} \boldsymbol{f}(t_n, \boldsymbol{y}(t_n)) + \frac{\tau - t_n}{h} \boldsymbol{f}(t_{n+1}, \boldsymbol{y}(t_{n+1})), \end{aligned}$$

where we have used the stepsize $t_{n+1} - t_n = h$.

FIGURE

Therefore, the integral becomes

$$\begin{aligned} \int_{t_{n+1}}^{t_{n+2}} \boldsymbol{f}(\tau, \boldsymbol{y}(\tau)) d\tau &\approx \int_{t_{n+1}}^{t_{n+2}} \boldsymbol{p}(\tau) d\tau \\ &= \int_{t_{n+1}}^{t_{n+2}} \left[\frac{t_{n+1} - \tau}{h} \boldsymbol{f}(t_n, \boldsymbol{y}(t_n)) + \frac{\tau - t_n}{h} \boldsymbol{f}(t_{n+1}, \boldsymbol{y}(t_{n+1})) \right] d\tau \\ &= \left[\boldsymbol{f}(t_n, \boldsymbol{y}(t_n)) \left(-\frac{1}{2} \right) \frac{(t_{n+1} - \tau)^2}{h} + \boldsymbol{f}(t_{n+1}, \boldsymbol{y}(t_{n+1})) \frac{(\tau - t_n)^2}{2h} \right]_{t_{n+1}}^{t_{n+2}} \\ &= \frac{3h}{2} \boldsymbol{f}(t_{n+1}, \boldsymbol{y}(t_{n+1})) - \frac{h}{2} \boldsymbol{f}(t_n, \boldsymbol{y}(t_n)). \end{aligned}$$

Thus (31) motivates the numerical method

$$\boldsymbol{y}_{n+2} = \boldsymbol{y}_{n+1} + \frac{h}{2}\left[3\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1}) - \boldsymbol{f}(t_n, \boldsymbol{y}_n)\right]. \tag{32}$$

Since formula (32) involves two previously computed solution values, this method is known as a *two-step method*. More precisely, is is known as the *second-order Adams-Bashforth method* (or AB method) dating back to 1883.

Remark 1. We will establish later that this method is indeed of second order accuracy.

2. Note that the method (32) requires two initial conditions. Since the IVP will give us only one initial condition, in the Matlab demo script `ABDemo.m` we take the second starting value from the exact solution. This is, of course, not realistic, and in practice one often precedes the Adams-Bashforth method by one step of, e.g., a second-order Runge-Kutta method (see later). However, even a single Euler step (which is also of order $\mathcal{O}(h^2)$) can also be used to start up (and maintain the accuracy of) the second-order AB method. This approach can also be used in `ABDemo.m` by uncommenting the corresponding line.

Example The Matlab script `ABDemo.m` compares the convergence of Euler's method (the one-step AB method) with the two-step AB method (32) for the IVP

$$y'(t) = -y^2(t), \qquad y(0) = 1$$

on the interval $[0, 10]$ with different stepsizes $N = 50, 100, 200$ and 400. The exact solution of this problem is

$$y(t) = \frac{1}{t+1}.$$

Example If we use a linear Lagrange interpolant to the integrand $\boldsymbol{f}$ of (31) at the points $\tau = t_{n+1}$ and $\tau = t_{n+2}$ then we obtain

$$\boldsymbol{y}(t_{n+2}) \approx y(t_{n+1}) + \frac{h}{2}\left[\boldsymbol{f}(t_{n+1}, \boldsymbol{y}(t_{n+1})) + \boldsymbol{f}(t_{n+2}, \boldsymbol{y}(t_{n+2}))\right]$$

or the numerical scheme

$$\boldsymbol{y}_{n+2} = \boldsymbol{y}_{n+1} + \frac{h}{2}\left[\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1}) + \boldsymbol{f}(t_{n+2}, \boldsymbol{y}_{n+2})\right]. \tag{33}$$

FIGURE

This method is known as second-order *Adams-Moulton method* (or AM method). It is a one-step method, and identical to the trapezoidal rule studied earlier (modulo a shift of the indices by one).

Remark In general, a pth-order Adams method is obtained by replacing the integrand $\boldsymbol{f}$ in (31) by a polynomial of degree $p-1$. However, the Adams-Bashforth method is an explicit method that uses the most recent information as well as $p-1$ "historical" points to fit the polynomial to. The pth-order Adams-Moulton method is an implicit method that fits the polynomial to the point to be determined next, the current point, and $p-2$ "historical" points. Therefore, the pth-order AB method is a p-step method, while the pth-order AM method is a $p-1$-step method.

For a general s-step Adams method we start the derivation as usual with the first-order ODE (42) and integrate from the current time t_{n+s-1} to the new time t_{n+s}. This gives us

$$\boldsymbol{y}(t_{n+s}) - \boldsymbol{y}(t_{n+s-1}) = \int_{t_{n+s-1}}^{t_{n+s}} \boldsymbol{f}(\tau, \boldsymbol{y}(\tau))d\tau. \tag{34}$$

As mentioned above, for the s-step AB method we now fit the slope $\boldsymbol{f}$ with a polynomial of degree $s-1$ at the s "known" points $t_n, t_{n+1}, \ldots, t_{n+s-2}, t_{n+s-1}$, i.e., we replace the integrand in (34) by the polynomial

$$\boldsymbol{p}(\tau) = \sum_{m=0}^{s-1} p_m(\tau)\boldsymbol{f}(t_{n+m}, \boldsymbol{y}_{m+n}),$$

where the p_m are the Lagrange functions (cf. Section 1.0.2)

$$p_m(\tau) = \prod_{\substack{\ell=0 \\ \ell \neq m}}^{s-1} \frac{\tau - t_{n+\ell}}{t_{n+m} - t_{n+\ell}}, \qquad m = 0, 1, \ldots, s-1.$$

This gives us

$$\begin{aligned}
\boldsymbol{y}(t_{n+s}) - \boldsymbol{y}(t_{n+s-1}) &\approx \int_{t_{n+s-1}}^{t_{n+s}} \boldsymbol{p}(\tau)d\tau \\
&= \int_{t_{n+s-1}}^{t_{n+s}} \sum_{m=0}^{s-1} p_m(\tau)\boldsymbol{f}(t_{n+m}, \boldsymbol{y}_{m+n})d\tau \\
&= \sum_{m=0}^{s-1} \boldsymbol{f}(t_{n+m}, \boldsymbol{y}_{m+n}) \int_{t_{n+s-1}}^{t_{n+s}} p_m(\tau)d\tau.
\end{aligned}$$

The calculations just performed motivate the numerical method

$$\boldsymbol{y}_{n+s} = \boldsymbol{y}_{n+s-1} + h\sum_{m=0}^{s-1} b_m \boldsymbol{f}(t_{n+m}, \boldsymbol{y}_{m+n}), \tag{35}$$

where the coefficients b_m, $m = 0, 1, \ldots, s-1$, are given by (using the substitution $u = \tau - t_{n+s-1}$, and the fact that $t_{n+s} - t_{n+s-1} = h$)

$$\begin{aligned}
b_m &= \frac{1}{h}\int_{t_{n+s-1}}^{t_{n+s}} p_m(\tau)d\tau \\
&= \frac{1}{h}\int_0^h p_m(t_{n+s-1} + u)du. \qquad (36)
\end{aligned}$$

Formula (35) together with the coefficients (36) is known as the general *s-step Adams-Bashforth method.*

Example The coefficients b_m are independent of the specific stepsize h and timestep n, so they can be computed once and for all. For the case $s = 2$ discussed earlier we compute

$$p_0(\tau) = \frac{\tau - t_{n+1}}{t_n - t_{n+1}},$$

Order	Formula	LTE
1	$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + h\boldsymbol{f}_n$	$\frac{h^2}{2}\boldsymbol{y}''(\eta)$
2	$\boldsymbol{y}_{n+2} = \boldsymbol{y}_{n+1} + \frac{h}{2}\left[3\boldsymbol{f}_{n+1} - \boldsymbol{f}_n\right]$	$\frac{5h^3}{12}\boldsymbol{y}'''(\eta)$
3	$\boldsymbol{y}_{n+3} = \boldsymbol{y}_{n+2} + \frac{h}{12}\left[23\boldsymbol{f}_{n+2} - 16\boldsymbol{f}_{n+1} + 5\boldsymbol{f}_n\right]$	$\frac{3h^4}{8}\boldsymbol{y}^{(4)}(\eta)$
4	$\boldsymbol{y}_{n+4} = \boldsymbol{y}_{n+3} + \frac{h}{24}\left[55\boldsymbol{f}_{n+3} - 59\boldsymbol{f}_{n+2} + 37\boldsymbol{f}_{n+1} - 9\boldsymbol{f}_n\right]$	$\frac{251h^5}{720}\boldsymbol{y}^{(5)}(\eta)$
5	$\boldsymbol{y}_{n+5} = \boldsymbol{y}_{n+4} + \frac{h}{720}\left[1901\boldsymbol{f}_{n+4} - 2774\boldsymbol{f}_{n+3} + 2616\boldsymbol{f}_{n+2} - 1274\boldsymbol{f}_{n+1} + 251\boldsymbol{f}_n\right]$	$\frac{95h^6}{2888}\boldsymbol{y}^{(6)}(\eta)$

Table 1: Adams-Bashforth formulas of different order. Notation: $\boldsymbol{f}_{n+m}$ denotes $\boldsymbol{f}(t_{n+m}, \boldsymbol{y}_{n+m})$, $m = 0, 1, \ldots, 5$, LTE stands for local truncation error.

and

$$\begin{aligned} b_0 &= \frac{1}{h}\int_0^h \frac{t_{n+1} + u - t_{n+1}}{t_n - t_{n+1}} du \\ &= \frac{1}{h}\int_0^h \frac{u}{-h} du \\ &= -\frac{1}{h^2}\left[\frac{u^2}{2}\right]_0^h = -\frac{1}{2}. \end{aligned}$$

Formulas for other choices of s are listed in Table 1.

Remark 1. The technique of using an interpolating polynomial $\boldsymbol{p}$ of degree $s-1$ at s equally spaced nodes with spacing h to replace the integrand $\boldsymbol{f}$ leads to so-called *Newton-Cotes formulas* for numerical integration. It is known that the interpolation error in this case is of the order $\mathcal{O}(h^s)$, and integration of this polynomial over an interval of length h adds another factor of h to the order. Therefore, the s-step Adams-Bashforth method has a local truncation error of order $\mathcal{O}(h^{s+1})$, which — provided the method converges — translates into a global method of order s.

2. General Adams-Moulton formulas can be derived similarly and are listed in Table 2. Note that the backward Euler method does not quite fit the general description of an AM method, since it is a single step method of order 1 (while the other AM methods are s-step methods of order $s+1$). In fact, there are two single step AM methods: the backward Euler method and the trapezoidal rule.

2.2 The Predictor-Corrector Idea

As mentioned earlier, one way to implement an implicit scheme is to couple it with a corresponding explicit scheme of the same order. We will now explain this *predictor-corrector* approach using the 2nd-order AB and AM methods.

Order	Formula	LTE
1	$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + h\boldsymbol{f}_{n+1}$	$-\frac{h^2}{2}\boldsymbol{y}''(\eta)$
2	$\boldsymbol{y}_{n+2} = \boldsymbol{y}_{n+1} + \frac{h}{2}\left[\boldsymbol{f}_{n+2} + \boldsymbol{f}_{n+1}\right]$	$-\frac{h^3}{12}\boldsymbol{y}'''(\eta)$
3	$\boldsymbol{y}_{n+3} = \boldsymbol{y}_{n+2} + \frac{h}{12}\left[5\boldsymbol{f}_{n+3} + 8\boldsymbol{f}_{n+2} - \boldsymbol{f}_{n+1}\right]$	$-\frac{h^4}{24}\boldsymbol{y}^{(4)}(\eta)$
4	$\boldsymbol{y}_{n+4} = \boldsymbol{y}_{n+3} + \frac{h}{24}\left[9\boldsymbol{f}_{n+4} + 19\boldsymbol{f}_{n+3} - 5\boldsymbol{f}_{n+2} + \boldsymbol{f}_{n+1}\right]$	$-\frac{19h^5}{720}\boldsymbol{y}^{(5)}(\eta)$
5	$\boldsymbol{y}_{n+5} = \boldsymbol{y}_{n+4} + \frac{h}{720}\left[251\boldsymbol{f}_{n+5} + 646\boldsymbol{f}_{n+4} - 264\boldsymbol{f}_{n+3} + 106\boldsymbol{f}_{n+2} - 19\boldsymbol{f}_{n+1}\right]$	$-\frac{3h^6}{160}\boldsymbol{y}^{(6)}(\eta)$

Table 2: Adams-Moulton formulas of different order. Notation: $\boldsymbol{f}_{n+m}$ denotes $\boldsymbol{f}(t_{n+m}, \boldsymbol{y}_{n+m})$, $m = 0, 1, \ldots, 5$, LTE stands for local truncation error.

We start with the *predictor* — in our case the second-order AB method. However, we treat its output only as a temporary answer, i.e.,

$$\tilde{\boldsymbol{y}}_{n+2} = \boldsymbol{y}_{n+1} + \frac{h}{2}\left[3\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1}) - \boldsymbol{f}(t_n, \boldsymbol{y}_n)\right].$$

Next we *correct* this value by using it on the right-hand side of the second-order AM method, i.e.,

$$\boldsymbol{y}_{n+2} = \boldsymbol{y}_{n+1} + \frac{h}{2}\left[\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1}) + \boldsymbol{f}(t_{n+2}, \tilde{\boldsymbol{y}}_{n+2})\right].$$

While this approach provides a simple realization of an implicit method, it can also be used to create a scheme that uses a variable stepsize h. The basic idea is to use the difference $|\tilde{\boldsymbol{y}}_{n+2} - \boldsymbol{y}_{n+2}|$ to judge the accuracy of the method. The following algorithm describes the general idea:

Algorithm

```
$\tilde{\boldsymbol{y}}_{n+2} = \boldsymbol{y}_{n+1} + \frac{h}{2}\left[3\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1}) - \boldsymbol{f}(t_n, \boldsymbol{y}_n)\right]$
$\boldsymbol{y}_{n+2} = \boldsymbol{y}_{n+1} + \frac{h}{2}\left[\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1}) + \boldsymbol{f}(t_{n+2}, \tilde{\boldsymbol{y}}_{n+2})\right]$
$\kappa = \frac{1}{6}|\tilde{\boldsymbol{y}}_{n+2} - \boldsymbol{y}_{n+2}|$
if $\kappa$ is relatively large, then
    $h \leftarrow h/2$ (i.e., reduce the stepsize)
    repeat
else if $\kappa$ is relatively small, then
    $h \leftarrow 2h$ (i.e., increase the stepsize)
else
    continue (i.e., keep $h$)
end
```

The specific choice of κ in the algorithm is motivated by the following argument. The local truncation errors for the second-order AB and AM methods, respectively, are

$$\begin{aligned} \boldsymbol{y}(t_{n+2}) - \tilde{\boldsymbol{y}}_{n+2} &= \frac{5}{12}h^3\boldsymbol{y}'''(\eta_{AB}) \\ \boldsymbol{y}(t_{n+2}) - \boldsymbol{y}_{n+2} &= -\frac{1}{12}h^3\boldsymbol{y}'''(\eta_{AM}). \end{aligned}$$

If we assume that $\boldsymbol{y}'''$ is nearly constant over the interval of interest, i.e., $\boldsymbol{y}'''(\eta_{AB}) \approx \boldsymbol{y}'''(\eta_{AM}) \approx \boldsymbol{y}'''(\eta)$, then we can subtract the above two equations from each other to get

$$\boldsymbol{y}_{n+2} - \tilde{\boldsymbol{y}}_{n+2} \approx \frac{1}{2}h^3\boldsymbol{y}'''(\eta),$$

and therefore the error at this time step is

$$|\boldsymbol{y}(t_{n+2}) - \boldsymbol{y}_{n+2}| \approx \frac{1}{12}h^3\boldsymbol{y}'''(\eta) \approx \frac{1}{6}|\boldsymbol{y}_{n+2} - \tilde{\boldsymbol{y}}_{n+2}|.$$

Remark

1. Finding a good way to characterize “relatively large” and “relatively small” in the algorithm can be tricky.
2. Note that it may be necessary to generate additional function values by interpolation if the stepsize is reduced and the predictor has to be evaluated for this new stepsize.
3. A variable stepsize, variable-order AB-AM predictor-corrector scheme is implemented in Matlab in the routine `ode113`.

2.3 Order and Convergence of Multistep Methods

There are even more general multistep methods than the Adams methods. We will write them in the form

$$\sum_{m=0}^{s} a_m\boldsymbol{y}_{n+m} = h\sum_{m=0}^{s} b_m\boldsymbol{f}(t_{n+m}, \boldsymbol{y}_{n+m}), \qquad n = 0, 1, \ldots, \tag{37}$$

where the coefficients a_m and b_m, $m = 0, 1, \ldots, s$ are independent of h, n, and the underlying ODE. Usually, the formula is normalized so that $a_s = 1$. The formula is a true s-step formula if either a_0 or b_0 are different from zero. Different choices of the coefficients a_m and b_m yield different numerical methods. In particular, if $b_s = 0$ the method will be explicit. Otherwise it will be implicit.

Remark The general multistep formula (37) is of the same form as so-called *recursive* or *infinite impulse response (IIR) digital filters* used in digital signal processing.

Example The second-order AB method corresponds to $s = 2$ with

$$a_2 = 1,\ a_1 = -1,\ a_0 = 0,\ b_2 = 0,\ b_1 = 3/2,\ b_0 = -1/2,$$

and the second-order AM method corresponds to $s = 2$ with

$$a_2 = 1,\ a_1 = -1,\ a_0 = 0,\ b_2 = 1/2,\ b_1 = 1/2,\ b_0 = 0.$$

Remark The coefficients a_m and b_m will play a crucial role in our following discussion of order and convergence of multistep methods, as well as later on in our study of stability.

As used many times before, a numerical IVP solver of the form

$$\boldsymbol{y}_{n+1} = \mathcal{Y}_n(\boldsymbol{f}, h, \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$$

is said to be of order p if and only if

$$\boldsymbol{y}(t_{n+1}) - \mathcal{Y}_n(\boldsymbol{f}, h, \boldsymbol{y}(t_0), \boldsymbol{y}(t_1), \ldots, \boldsymbol{y}(t_n)) = \mathcal{O}(h^{p+1}).$$

For the multistep methods (37) we can alternatively consider the linear functional ψ_t defined by

$$\begin{aligned} \psi_t \boldsymbol{y} &= \sum_{m=0}^{s} a_m \boldsymbol{y}(t+mh) - h \sum_{m=0}^{s} b_m \boldsymbol{f}(t+mh, \boldsymbol{y}(t+mh)) \\ &= \sum_{m=0}^{s} a_m \boldsymbol{y}(t+mh) - h \sum_{m=0}^{s} b_m \boldsymbol{y}'(t+mh). \end{aligned}$$

Then the s-step method (37) is of order p if and only if

$$\psi_t \boldsymbol{y} = \mathcal{O}(h^{p+1})$$

for all sufficiently smooth functions $\boldsymbol{y}$.

We now characterize the order p of a multistep method in terms of the coefficients a_m and b_m.

Theorem 2.1 *The multistep method*

$$\sum_{m=0}^{s} a_m \boldsymbol{y}_{n+m} = h \sum_{m=0}^{s} b_m \boldsymbol{f}(t_{n+m}, \boldsymbol{y}_{n+m})$$

is of order $p \geq 1$ if and only if

$$\begin{aligned} \sum_{m=0}^{s} a_m &= 0, \\ \sum_{m=0}^{s} \frac{m^k}{k!} a_m &= \sum_{m=0}^{s} \frac{m^{k-1}}{(k-1)!} b_m, \qquad k = 1, 2, \ldots, p, \\ \sum_{m=0}^{s} \frac{m^{p+1}}{(p+1)!} a_m &\neq \sum_{m=0}^{s} \frac{m^p}{p!} b_m. \end{aligned}$$

Proof We have

$$\psi_t \boldsymbol{y} = \sum_{m=0}^{s} a_m \boldsymbol{y}(t+mh) - h \sum_{m=0}^{s} b_m \boldsymbol{y}'(t+mh).$$

Using Taylor expansions for both $\boldsymbol{y}$ and $\boldsymbol{y}'$ we obtain

$$\begin{aligned}
\psi_t \boldsymbol{y} &= \sum_{m=0}^{s} a_m \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{y}^{(k)}(t)(mh)^k - h\sum_{m=0}^{s} b_m \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{y}^{(k+1)}(t)(mh)^k \\
&= \sum_{m=0}^{s} a_m \sum_{k=0}^{\infty} \frac{1}{k!}\boldsymbol{y}^{(k)}(t)m^k h^k - h\sum_{m=0}^{s} b_m \sum_{k=1}^{\infty} \frac{1}{(k-1)!}\boldsymbol{y}^{(k)}(t)m^{k-1}h^{k-1} \\
&= \sum_{k=0}^{\infty}\left(\sum_{m=0}^{s}\frac{m^k}{k!}a_m\right)\boldsymbol{y}^{(k)}(t)h^k - \sum_{k=1}^{\infty}\left(\sum_{m=0}^{s}\frac{m^{k-1}}{(k-1)!}b_m\right)\boldsymbol{y}^{(k)}(t)h^k \\
&= \left(\sum_{m=0}^{s} a_m\right)\boldsymbol{y}(t) + \sum_{k=1}^{\infty}\left(\sum_{m=0}^{s}\frac{m^k}{k!}a_m\right)\boldsymbol{y}^{(k)}(t)h^k - \sum_{k=1}^{\infty}\left(\sum_{m=0}^{s}\frac{m^{k-1}}{(k-1)!}b_m\right)\boldsymbol{y}^{(k)}(t)h^k \\
&= \left(\sum_{m=0}^{s} a_m\right)\boldsymbol{y}(t) + \sum_{k=1}^{\infty}\left(\sum_{m=0}^{s}\frac{m^k}{k!}a_m - \sum_{m=0}^{s}\frac{m^{k-1}}{(k-1)!}b_m\right)\boldsymbol{y}^{(k)}(t)h^k
\end{aligned}$$

We get $\psi_t \boldsymbol{y} = \mathcal{O}(h^{p+1})$ by satisfying the conditions as claimed. ∎

Remark 1. We can now use the simple conditions in Theorem 2.1 to check the order of *any* multistep method. This is generally *much easier* than the special techniques we used earlier (see the example below).

2. A method is called *consistent* if it is of order $p \geq 1$, i.e., if

$$\sum_{m=0}^{s} a_m = 0 \quad \text{and} \quad \sum_{m=0}^{s} m a_m = \sum_{m=0}^{s} b_m.$$

In this context consistent is short for *the numerical method is consistent with the differential equation*. This means that the discrete numerical method (a difference equation) is an accurate representation of the continuous problem (given in the form of a differential equation), i.e., the local truncation error is small. Remember that we noted earlier that merely establishing the order of a method does not ensure its convergence (see the second example below). With this new terminology we can say that *consistency alone does not imply convergence*.

3. If we introduce the polynomials (often called *characteristic polynomials* or *generating polynomials* of the method)

$$\rho(w) = \sum_{m=0}^{s} a_m w^m, \qquad \sigma(w) = \sum_{m=0}^{s} b_m w^m,$$

then one can show that the general multistep method is of order p if and only if there exists a constant $c \neq 0$ such that

$$\rho(w) - \sigma(w)\ln w = c(w-1)^{p+1} + \mathcal{O}(|w-1|^{p+2}) \qquad \text{as } w \to 1. \tag{38}$$

Note that this condition tells us how well the log-function is approximated by the rational function $\frac{\rho}{\sigma}$ near $w = 1$. In terms of the polynomials ρ and σ the two consistency conditions above correspond to

$$\rho(1) = 0 \quad \text{and} \quad \rho'(1) = \sigma(1).$$

4. The constant c in (38) (as well as the difference of the two sides of the third condition in Theorem 2.1) is in fact the local error constant of the method (cf. the following example and the local truncation errors listed in Tables 1 and 2).

Example We show that the Adams-Bashforth method (32) derived earlier is indeed of second order. The iteration formula was

$$\boldsymbol{y}_{n+2} - \boldsymbol{y}_{n+1} = \frac{h}{2}\left[3\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1}) - \boldsymbol{f}(t_n, \boldsymbol{y}_n)\right],$$

so that — as noted earlier — $s = 2$ and $a_2 = 1$, $a_1 = -1$, $a_0 = 0$, $b_2 = 0$, $b_1 = 3/2$, $b_0 = -1/2$.

Now,

$$\sum_{m=0}^{2} a_m = 0 - 1 + 1 = 0,$$

and for $k = 1$ (note: $0! = 1! = 1$)

$$\begin{aligned} \sum_{m=0}^{2} m a_m &= \sum_{m=0}^{2} b_m \\ \iff 0(0) + (1)(-1) + 2(1) &= -\frac{1}{2} + \frac{3}{2} + 0 \\ \iff 1 &= 1, \end{aligned}$$

for $k = 2$

$$\begin{aligned} \sum_{m=0}^{2} \frac{m^2}{2} a_m &= \sum_{m=0}^{2} m b_m \\ \iff \frac{0}{2}(0) + \frac{1}{2}(-1) + \frac{4}{2}(1) &= (0)(-\frac{1}{2}) + (1)\frac{3}{2} + (2)0 \\ \iff \frac{3}{2} &= \frac{3}{2}, \end{aligned}$$

and for $k = 3$

$$\begin{aligned} \sum_{m=0}^{2} \frac{m^3}{3!} a_m &= \sum_{m=0}^{2} \frac{m^2}{2!} b_m \\ \iff \frac{0}{6}(0) + \frac{1}{6}(-1) + \frac{8}{6}(1) &= \frac{0}{2}(-\frac{1}{2}) + \frac{1}{2}\frac{3}{2} + \frac{4}{2}0 \\ \iff \frac{7}{6} &= \frac{3}{4}. \end{aligned}$$

Therefore, the method is indeed of order $p = 2$. Moreover, the difference $\frac{7}{6} - \frac{3}{4} = \frac{5}{12}$ (the error constant listed in Table 1).

Alternatively, we can check the condition

$$\rho(w) - \sigma(w)\ln w = c(w-1)^{p+1} + \mathcal{O}(|w-1|^{p+2}) \qquad \text{as } w \to 1.$$

For our example

$$\begin{aligned}
\rho(w) &= \sum_{m=0}^{2} a_m w^m = w^2 - w, \\
\sigma(w) &= \sum_{m=0}^{2} b_m w^m = \frac{3}{2}w - \frac{1}{2}.
\end{aligned}$$

Since the right-hand side of our condition is written in terms of $w-1$ we express everything in terms of $\xi = w - 1$ and then take the limit as $\xi \to 0$. Thus

$$\begin{aligned}
\rho(\xi+1) &= (\xi+1)^2 - (\xi+1) = \xi^2 + \xi, \\
\sigma(\xi+1) &= \frac{3}{2}(\xi+1) - \frac{1}{2} = \frac{3}{2}\xi + 1,
\end{aligned}$$

and therefore (using the Taylor expansion of the logarithm)

$$\begin{aligned}
\rho(\xi+1) - \sigma(\xi+1)\ln(\xi+1) &= \left(\xi^2+\xi\right) - \left(\frac{3}{2}\xi + 1\right)\ln(\xi+1) \\
&= \left(\xi^2+\xi\right) - \left(\frac{3}{2}\xi + 1\right)\sum_{k=1}^{\infty}(-1)^{k-1}\frac{\xi^k}{k} \\
&= \left(\xi^2+\xi\right) - \left(\frac{3}{2}\xi + 1\right)\left(\xi - \frac{\xi^2}{2} + \frac{\xi^3}{3} - \cdots\right) \\
&= \left(\xi^2+\xi\right) - \left(\xi + \xi^2 - \frac{5}{12}\xi^3 + \mathcal{O}(\xi^4)\right) = \frac{5}{12}\xi^3 + \mathcal{O}(\xi^4).
\end{aligned}$$

Thus, $c = \frac{5}{12} \neq 0$ (again, the error constant of Table 1), and the order is $p = 2$ as before.

Example The implicit 2-step method

$$\boldsymbol{y}_{n+2} - 3\boldsymbol{y}_{n+1} + 2\boldsymbol{y}_n = h\left[\frac{13}{12}\boldsymbol{f}(t_{n+2}, \boldsymbol{y}_{n+2}) - \frac{5}{3}\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1}) - \frac{5}{12}\boldsymbol{f}(t_n, \boldsymbol{y}_n)\right] \qquad (39)$$

has order two. This can easily be verified using the criteria of Theorem 2.1. However, consider the trivial IVP

$$y'(t) = 0, \qquad y(0) = 1,$$

with solution $y(t) = 1$. For this example the right-hand side of (39) is always zero, so we immediately get the 3-term recurrence relation

$$y_{n+2} - 3y_{n+1} + 2y_n = 0.$$

The general solution of this equation is given by

$$y_n = c_1 + c_2 2^n, \qquad n = 0, 1, \ldots, \qquad (40)$$

with arbitrary constants c_1 and c_2. This can easily be verified using induction.

While the choice $c_1 = 1$ and $c_2 = 0$ does provide the exact solution, in practice we will have to initialize the method with two values y_0 and y_1, and most likely these values will be different (which is equivalent to having $c_2 \neq 0$ in (40) above). In this case, however, the numerical solution blows up (since $2^n \to \infty$ for $n \to \infty$). Thus, only for one very special set of starting values do we have an accurate solution. In general, the method *does not converge* (even though it is consistent, and is even of order 2).

We are now ready to present a necessary and sufficient condition for convergence of a multistep method that is very easy to check. To this end we have

Definition 2.2 *A (complex) polynomial p obeys the* root condition *if*

- *all its zeros, i.e., all z such that $p(z) = 0$, lie in the unit disk, i.e., $|z| \leq 1$, and*
- *all zeros on the unit circle are simple, i.e., if $|z| = 1$ then $p'(z) \neq 0$.*

Theorem 2.3 *(Dahlquist Equivalence Theorem) Consider the general multistep method (cf. (37))*

$$\sum_{m=0}^{s} a_m \boldsymbol{y}_{n+m} = h \sum_{m=0}^{s} b_m \boldsymbol{f}(t_{n+m}, \boldsymbol{y}_{n+m}), \qquad n = 0, 1, \ldots,$$

and assume that the starting values $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{s-1}$ are accurate, i.e., they are determined up to an error that tends to zero for $h \to 0$. Then the multistep method is convergent if and only if it is consistent and the polynomial ρ obeys the root condition.

Proof The proof is too involved to be included here. ∎

Remark

1. An immediate — and very important — consequence of the Equivalence Theorem is that a multistep method whose characteristic polynomial ρ does not satisfy the root condition cannot be convergent. Do not use such methods!
2. We will see later that the fact that ρ satisfies the root condition is equivalent to stability of a multistep method. Therefore, for multistep methods,

$$\text{convergence} \quad \Longleftrightarrow \quad \text{consistency \& stability.}$$

Example Earlier we claimed that the implicit 2-step method

$$\boldsymbol{y}_{n+2} - 3\boldsymbol{y}_{n+1} + 2\boldsymbol{y}_n = h\left[\frac{13}{12}\boldsymbol{f}(t_{n+2}, \boldsymbol{y}_{n+2}) - \frac{5}{3}\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1}) - \frac{5}{12}\boldsymbol{f}(t_n, \boldsymbol{y}_n)\right] \qquad (41)$$

has order two, and gave a counterexample to show that it cannot be convergent. Now we can use the Dahlquist Equivalence Theorem to establish this fact. The characteristic polynomial ρ of the method is given by

$$\rho(w) = \sum_{m=0}^{2} a_m w^m = w^2 - 3w + 2 = (w-1)(w-2),$$

and we see that ρ violates the root condition since one of its zeros, $w = 2$, lies outside the unit disk. According to the Equivalence Theorem it cannot be convergent.

Example The characteristic polynomials ρ for all Adams methods (both AB and AM) are of the same type, namely,

$$\rho(w) = w^s - w^{s-1} = w^{s-1}(w-1),$$

and so they all satisfy the root condition. As a consequence, *all Adams methods are convergent* (since we already established that they are consistent).

If we look at the general s-step method

$$\sum_{m=0}^{s} a_m \boldsymbol{y}_{n+m} = h \sum_{m=0}^{s} b_m \boldsymbol{f}(t_{n+m}, \boldsymbol{y}_{n+m}), \qquad n = 0, 1, \ldots,$$

then we see that this method involves $2s+1$ free parameters (after normalization). Therefore, one might think that it is possible to construct an s-step method that is of order $2s$. Unfortunately, there is another theorem by Dahlquist that states that one cannot have a convergent s-step method of order $2s$ for any $s \geq 3$. More precisely,

Theorem 2.4 *(Dahlquist's First Barrier) The maximal order of a convergent s-step method is at most*

- *$s+2$ for implicit schemes with s even,*
- *$s+1$ for implicit schemes with s odd, and*
- *s for explicit schemes.*

Proof Also too complicated. ∎

Remark 1. A procedure for construction of a convergent s-step method of order $s+1$ is outlined in the Iserles book.

2. Adams-Bashforth methods are *optimal* in the sense that the corresponding order is as high as possible. The same is true for Adams-Moulton formulas with odd s. It can be shown that implicit s-step methods of order $s+2$ are of questionable stability.

2.4 Backward Differentiation Formulae

As mentioned earlier, the "extreme" choice $\rho(w) = w^{s-1}(w-1)$ that always satisfies the root condition and also places as many of the zeros at the origin as possible yields the family of Adams methods. If we, on the other hand, choose an extreme $\sigma(w) = \beta w^s$, then we obtain the so-called *backward differentiation formulae* (BDFs).

These methods have their name from the way in which they can be derived. Starting from the ODE $\boldsymbol{y}'(t) = \boldsymbol{f}(t, \boldsymbol{y})$ we do not integrate as before (and fit the integrand $\boldsymbol{f}$ on the right-hand side by a polynomial), but instead fit the derivative $\boldsymbol{y}'$ on the left-hand side by the derivative of the interpolating polynomial to the $s+1$ data points chosen *backward* from the new point, i.e., $(t_{n+s}, \boldsymbol{y}_{n+s}), (t_{n+s-1}, \boldsymbol{y}_{n+s-1}), \ldots, (t_n, \boldsymbol{y}_n)$.

Example In the simplest case, $s = 1$, we get the *Backward Euler method.* The interpolating polynomial to the data $(t_{n+1}, \boldsymbol{y}_{n+1}), (t_n, \boldsymbol{y}_n)$ is given by

$$\boldsymbol{p}(\tau) = \frac{t_{n+1} - \tau}{t_{n+1} - t_n}\boldsymbol{y}_n + \frac{\tau - t_{n+1}}{t_{n+1} - t_n}\boldsymbol{y}_{n+1},$$

and its derivative is

$$\boldsymbol{p}'(\tau) = -\frac{1}{h}\boldsymbol{y}_n + \frac{1}{h}\boldsymbol{y}_{n+1},$$

where we have use $t_{n+1} - t_n = h$ as usual. If we replace the right-hand side of the ODE by $\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1})$, then we end up with the numerical method

$$-\frac{1}{h}\boldsymbol{y}_n + \frac{1}{h}\boldsymbol{y}_{n+1} = \boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1}) \quad \Longleftrightarrow \quad \boldsymbol{y}_{n+1} = \boldsymbol{y}_n + h\boldsymbol{f}(t_{n+1}, \boldsymbol{y}_{n+1})$$

— the backward Euler method.

In general, one can show that once the form of the polynomial σ and the order s are chosen, then the method is determined.

Lemma 2.5 *For a BDF of order s with $\sigma(w) = \beta w^s$ we have*

$$\beta = \left(\sum_{m=1}^{s} \frac{1}{m}\right)^{-1}, \quad \textit{and} \quad \rho(w) = \beta \sum_{m=1}^{s} \frac{1}{m} w^{s-m}(w-1)^m.$$

Proof Straightforward algebra using (38) and a Taylor series expansion for the logarithm. See the Iserles book. ▮

Example In the simplest case $s = 1$ we have $\sigma(w) = \beta w$ with $\beta = 1$ and $\rho(w) = w - 1$. This gives us the backward Euler method mentioned above.

For $s = 2$ we have $\sigma(w) = \beta w^2$, and

$$\beta = \left(\sum_{m=1}^{2} \frac{1}{m}\right)^{-1} = \frac{1}{1 + \frac{1}{2}} = \frac{2}{3}.$$

With this value

$$\rho(w) = \beta \sum_{m=1}^{2} \frac{1}{m} w^{s-m}(w-1)^m = \frac{2}{3}\left[w(w-1) + \frac{1}{2}(w-1)^2\right] = w^2 - \frac{4}{3}w + \frac{1}{3}.$$

This results in the BDF formula

$$\boldsymbol{y}_{n+2} - \frac{4}{3}\boldsymbol{y}_{n+1} + \frac{1}{3}\boldsymbol{y}_n = \frac{2}{3}h\boldsymbol{f}(t_{n+2}, \boldsymbol{y}_{n+2}).$$

While the Adams methods were constructed so that they satisfy the root condition, this is no longer automatically true for the BDFs. In fact, we have

Theorem 2.6 *The characteristic polynomial ρ for a BDF satisfies the root condition and the underlying BDF method is convergent if and only if $1 \leq s \leq 6$.*

Order	Formula	LTE
1	$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + h\boldsymbol{f}_{n+1}$	$-\frac{h^2}{2}\boldsymbol{y}''(\eta)$
2	$\boldsymbol{y}_{n+2} - \frac{4}{3}\boldsymbol{y}_{n+1} + \frac{1}{3}\boldsymbol{y}_n = \frac{2h}{3}\boldsymbol{f}_{n+2}$	$-\frac{2h^3}{9}\boldsymbol{y}'''(\eta)$
3	$\boldsymbol{y}_{n+3} - \frac{18}{11}\boldsymbol{y}_{n+2} + \frac{9}{11}\boldsymbol{y}_{n+1} - \frac{2}{11}\boldsymbol{y}_n = \frac{6h}{11}\boldsymbol{f}_{n+3}$	$-\frac{3h^4}{22}\boldsymbol{y}^{(4)}(\eta)$
4	$\boldsymbol{y}_{n+4} - \frac{48}{25}\boldsymbol{y}_{n+3} + \frac{36}{25}\boldsymbol{y}_{n+2} - \frac{16}{25}\boldsymbol{y}_{n+1} + \frac{3}{25}\boldsymbol{y}_n = \frac{12h}{25}\boldsymbol{f}_{n+4}$	$-\frac{12h^5}{125}\boldsymbol{y}^{(5)}(\eta)$
5	$\boldsymbol{y}_{n+5} - \frac{300}{137}\boldsymbol{y}_{n+4} + \frac{300}{137}\boldsymbol{y}_{n+3} - \frac{200}{137}\boldsymbol{y}_{n+2} + \frac{75}{137}\boldsymbol{y}_{n+1} - \frac{12}{137}\boldsymbol{y}_n = \frac{60h}{137}\boldsymbol{f}_{n+5}$	$-\frac{10h^6}{137}\boldsymbol{y}^{(6)}(\eta)$
6	$\boldsymbol{y}_{n+6} - \frac{360}{147}\boldsymbol{y}_{n+5} + \frac{450}{147}\boldsymbol{y}_{n+4} - \frac{400}{147}\boldsymbol{y}_{n+3} + \frac{225}{147}\boldsymbol{y}_{n+2} - \frac{72}{147}\boldsymbol{y}_{n+1} + \frac{10}{147}\boldsymbol{y}_n = \frac{60h}{147}\boldsymbol{f}_{n+6}$	$-\frac{20h^7}{343}\boldsymbol{y}^{(7)}(\eta)$

Table 3: Backward differentiation formulas of different order. Notation: $\boldsymbol{f}_{n+m}$ denotes $\boldsymbol{f}(t_{n+m}, \boldsymbol{y}_{n+m})$, $m = 0, 1, \ldots, 6$, LTE stands for local truncation error.

Proof Too involved.

Remark 1. The root condition fails for $s > 6$.

2. The coefficients and local truncation errors for all 6 BDFs of practical interest are listed in Table 3.

3. Note that all BDF methods are *implicit* methods, and therefore special care is needed for their implementation. However, as the following example shows (and as we will see in more detail later), they have better stability properties than, say AB methods, and are therefore better suited for *stiff* problems.

Example The Matlab script `StiffDemo.m` compares the performance of the second-order BDF method with that of the second-order AB method for the linear — but stiff — ODE system

$$\boldsymbol{y}'(t) = \begin{bmatrix} -10 & 1 \\ 0 & -1 \end{bmatrix} \boldsymbol{y}, \qquad \boldsymbol{y}(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

The solution to this problem is

$$\boldsymbol{y}(t) = \begin{bmatrix} \frac{1}{9}e^{-t} + \frac{8}{9}e^{-10t} \\ e^{-t} \end{bmatrix},$$

and the presence of the two different scales in the first component of the solution is what makes this problem *stiff*. We will discuss stiff problems in more detail in Section 4.