

Chapter 3

Linear Multistep Methods



[...] this approach, which was first adopted by Dahlquist, leads to a mathematically well-rounded theory. It also leads to the discovery of new integration formulas which could not be obtained by the heuristic methods.

(Peter Henrici [206]. This quotation has also been highlighted by Ernst Hairer in [188])

One-step methods have been introduced, analyzed and implemented in Chap. 2. Even if they provide the simplest and maybe most intuitive family of step-by-step numerical schemes, enlarging this class with more complex methods could be useful in order to achieve better accuracy and stability properties. For this reason, we present a more general family of methods relying on a multistep structure, defined as follows.

3.1 The Principle of Multistep Numerical Integration

Definition 3.1 The family of *linear multistep methods* (LMMs), with respect to the discretization (2.1), is defined by the order k difference equation

$$\sum_{j=0}^k \alpha_j y_{n+j} = h \sum_{j=0}^k \beta_j f_{n+j}, \quad (3.1)$$

where $f_{n+j} = f(t_{n+j}, y_{n+j})$, $j = 0, 1, \dots, k$, $n = 0, 1, \dots, N$.

The integer k is usually denoted as the number of *steps* of the method and represents the order of the difference equation defining (3.1). Normally, we assume $\alpha_k = 1$ (as

usual, if this is not the case, all coefficients can be normalized in order to fall in this instance) and, moreover, $|\alpha_0| + |\beta_0| \neq 0$ in order to avoid α_0 and β_0 simultaneously equal to zero.

The family of LMMs includes all the one-step methods (hence, $k = 1$) introduced in Chap. 2, namely

- explicit Euler method (2.19), assuming that $\alpha_0 = -1, \alpha_1 = 1, \beta_0 = 1, \beta_1 = 0$;
- implicit Euler method (2.32), for $\alpha_0 = -1, \alpha_1 = 1, \beta_0 = 0, \beta_1 = 1$;
- the trapezoidal method (2.33), imposing $\alpha_0 = -1, \alpha_1 = 1, \beta_0 = \frac{1}{2}, \beta_1 = \frac{1}{2}$.

Let us now provide an example of LMM (3.1) depending on more than one step, obtained by means of proper numerical quadrature.

Example 3.1 We now aim to derive an example of LMM with $k = 2$, i.e., a two-step method. As in the construction of the trapezoidal method (2.33), let us consider, for $t \geq t_n$, the integral form of (1.1), i.e.,

$$y(t) = y(t_n) + \int_{t_n}^t f(s, y(s)) ds,$$

and evaluate it in t_{n+2} , obtaining

$$y(t_{n+2}) = y(t_n) + \int_{t_n}^{t_{n+2}} f(s, y(s)) ds.$$

Approximating the integral in the right-hand side by the Cavalieri-Simpson formula

$$\int_{t_n}^{t_{n+2}} f(s, y(s)) ds \approx \frac{h}{3} (f(t_n, y(t_n)) + 4f(t_{n+1}, y(t_{n+1})) + f(t_{n+2}, y(t_{n+2})))$$

yields

$$y(t_{n+2}) \approx y(t_n) + \frac{h}{3} (f(t_n, y(t_n)) + 4f(t_{n+1}, y(t_{n+1})) + f(t_{n+2}, y(t_{n+2}))).$$

This is an approximate equality involving exact values that can be regarded as an exact equality involving approximate values, i.e.,

$$y_{n+2} = y_n + \frac{h}{3} (f_n + 4f_{n+1} + f_{n+2}), \quad (3.2)$$

that is the so-called *Milne-Simpson* method.

Differently from one-step methods, LMMs are not self-starting (unless $k = 1$). For instance, consider the Milne-Simpson method (3.2) when $n = 0$, leading to

$$y_2 = y_0 + \frac{h}{3} (f_0 + 4f_1 + f_2).$$

The value of y_0 is initial value given by the problem (1.1), but the value of y_1 is missing and it needs to be recovered in order to compute y_2 and launch the step-by-step procedure.

More in general, for the family of k -step methods (3.1) a proper starting method is needed to reconstruct the missing starting values y_1, y_2, \dots, y_{k-1} . Such values can be recovered, for instance, by a suitable one-step method. Just as an example, if we employ Euler method (2.19) as starting method, the step-by-step numerical scheme described by (3.1) can be summarized as follows:

- y_0 is given by the initial value problem (1.1);
- compute the missing starting values y_1, y_2, \dots, y_{k-1} by repeatedly applying (2.19), i.e.,

$$\begin{aligned} y_1 &= y_0 + hf_0, \\ y_2 &= y_1 + hf_1, \\ &\vdots \\ y_{k-1} &= y_{k-2} + hf_{k-2}; \end{aligned}$$

- compute y_k by applying the LMM (3.1)

$$y_k + \sum_{j=0}^{k-1} \alpha_j y_j = h\beta_k f_k + \sum_{j=0}^{k-1} \beta_j f_j.$$

We observe that, if $\beta_k \neq 0$, this step is equivalent to solving a nonlinear system of algebraic equations in y_k ;

- go on applying (3.1) up to the computation of

$$y_N + \sum_{j=0}^{k-1} \alpha_j y_{N-k+j} = h\beta_k f_N + \sum_{j=0}^{k-1} \beta_j f_{N-k+j}.$$

Relevant examples of LMMs (3.1) can be computed, for instance, via polynomial interpolation, where the interpolation points are normally chosen among the grid points. Let us illustrate this idea through the following examples.

Example 3.2 (An Adams-Bashforth Method) Let us consider the integral formulation of (1.1)

$$y(t) = y(t_{n+1}) + \int_{t_{n+1}}^t f(s, y(s)) ds,$$

and evaluate it in t_{n+2} , obtaining

$$y(t_{n+2}) = y(t_{n+1}) + \int_{t_{n+1}}^{t_{n+2}} f(s, y(s)) ds.$$

Let us approximate $f(s, y(s))$ through the linear interpolant with respect to the nodes $(t_n, y(t_n))$ and $(t_{n+1}, y(t_{n+1}))$, leading to

$$f(s, y(s)) \approx f(t_n, y(t_n)) \frac{s - t_{n+1}}{t_n - t_{n+1}} + f(t_{n+1}, y(t_{n+1})) \frac{s - t_n}{t_{n+1} - t_n}.$$

Hence,

$$\int_{t_n}^{t_{n+2}} f(s, y(s)) ds \approx -\frac{h}{2} (f(t_n, y(t_n)) - 3f(t_{n+1}, y(t_{n+1}))),$$

leading to

$$y_{n+2} = y_{n+1} - \frac{h}{2} (f_n - 3f_{n+1}), \quad (3.3)$$

that is the so-called *two-step Adams-Bashforth method*, which is an explicit method.

More in general, Adams-Bashforth methods are obtained by replacing the interpolating polynomial approximating f on a given set of nodes chosen among the grid points, excluding the point related to the advancing term, i.e., t_{n+k} . This choice leads to a family of explicit methods. Including t_{n+k} in the set of interpolation points leads to a family of implicit methods, the so-called *Adams-Moulton* formulae. An example is given below.

Example 3.3 (An Adams-Moulton Method) Let us consider again the integral form of (1.1)

$$y(t) = y(t_n) + \int_{t_n}^t f(s, y(s)) ds,$$

and proceed by approximating the function f with its linear interpolant on the nodes $(t_n, y(t_n))$ and $(t_{n+1}, y(t_{n+1}))$, leading to

$$f(s, y(s)) \approx f(t_n, y(t_n)) \frac{s - t_{n+1}}{t_n - t_{n+1}} + f(t_{n+1}, y(t_{n+1})) \frac{s - t_n}{t_{n+1} - t_n}.$$

Hence,

$$\int_{t_n}^{t_{n+1}} f(s, y(s)) ds \approx \frac{h}{2} (f(t_n, y(t_n)) + f(t_{n+1}, y(t_{n+1}))),$$

obtaining

$$y_{n+1} = y_n + \frac{h}{2} (f_n + f_{n+1}), \quad (3.4)$$

that is the so-called *second order Adams-Moulton method*. Actually, this implicit method is not a novelty for us, since it is the trapezoidal method (2.33).

3.2 Handling Implicitness by Fixed Point Iterations

Looking at the coefficient β_k in (3.1) allows us to distinguish whether the method is explicit or implicit: indeed, if $\beta_k = 0$, the method is explicit; when $\beta_k \neq 0$, the method is implicit. As explained in Chap. 2 for one-step methods, we now aim to handle the implicitness of LMMs via fixed point iterations. To this purpose, let us first recast (3.1) in a different, equivalent form. In particular, let us separate the implicit part from the explicit one in the method, by isolating the terms for $j = k$ in the summations, leading to

$$y_{n+k} = h\beta_k f(t_{n+k}, y_{n+k}) + g_{n+k-1}, \quad (3.5)$$

where

$$g_{n+k-1} = h \sum_{j=0}^{k-1} \beta_j f_{n+j} - \sum_{j=0}^{k-1} \alpha_j y_{n+j}.$$

Then, we treat the implicitness of (3.5) by fixed point iterations as follows: in order to advance from t_n to t_{n+1} ,

- we arbitrarily choose an initial guess $y_{n+k}^{[0]} \in \mathbb{R}^d$. To speed up the convergence of the iterative process, a smart choice may be $y_{n+k}^{[0]} = y_{n+k-1}$;
- we perform fixed point iterations

$$y_{n+k}^{[v]} = h\beta_k f(t_{n+k}, y_{n+k}^{[v-1]}) + g_{n+k-1}, \quad v \geq 1, \quad (3.6)$$

stopping at the iteration M if

$$\|y_{n+k}^{[M]} - y_{n+k}^{[M-1]}\| \leq tol,$$

being tol an a-priori prescribed accuracy. Then $y_{n+k} = y_{n+k}^{[M]}$.

A major issue to address regards the convergence of the above fixed point iterative process. This aspect is object of the following result.

Theorem 3.1 *Consider the initial value problem (1.1), whose vector field satisfies the Lipschitz condition (1.8), and denote by L the Lipschitz constant. If*

$$h|\beta_k|L < 1, \quad (3.7)$$

then (3.5) has a unique solution y_{n+k} such that

$$y_{n+k} = \lim_{v \rightarrow \infty} y_{n+k}^{[v]},$$

with $y_{n+k}^{[v]}$ defined in (3.6), for any arbitrarily chosen initial guess $y_{n+k}^{[0]} \in \mathbb{R}^d$.

Proof Let us introduce the auxiliary map $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}^d$, defined by

$$\varphi(y) = h\beta_k f(t_{n+k}, y) + g_{n+k-1}, \quad y \in \mathbb{R}^d.$$

For any $y, z \in \mathbb{R}^d$, we have

$$\|\varphi(y) - \varphi(z)\| = h |\beta_k| \|f(t_{n+k}, y) - f(t_{n+k}, z)\|$$

and, by the Lipschitz continuity of f , we obtain

$$\|\varphi(y) - \varphi(z)\| \leq h |\beta_k| L \|y - z\|.$$

Since $h|\beta_k|L < 1$,

$$\|\varphi(y) - \varphi(z)\| \leq \|y - z\|,$$

i.e., φ is a contraction in \mathbb{R}^d . Hence, by the contraction mapping theorem, there exists a unique fixed point $\alpha \in \mathbb{R}^d$ such that $\alpha = \varphi(\alpha)$. Since $y_{n+k} = \varphi(y_{n+k})$, y_{n+k} is the unique fixed point of the map φ . By the contraction mapping theorem, such a fixed point is the limit of the fixed point iterations

$$y_{n+k}^{[v]} = \varphi(y_{n+k}^{[v-1]}), \quad v \geq 1,$$

for any arbitrarily chosen initial guess $y_{n+k}^{[0]} \in \mathbb{R}^d$. □

We highlight that (3.7) is the first limitation on the stepsize we have encountered so far: in order to have a convergent fixed point iterative process for implicit LMMs (3.1), h cannot be arbitrarily chosen, but it should satisfy the restriction

$$h < \frac{1}{|\beta_k|L}.$$

We present other relevant stepsize restrictions in the remainder of this book. We will see that, in certain situations, important properties of numerical methods can be translated into proper stepsize restrictions.

3.3 Consistency and Order Conditions

Let us now focus our attention on the analysis of the accuracy of LMMs (3.1). We have seen in Chap. 2 that basic necessary accuracy and stability requirements (i.e., consistency, zero-stability and convergence) have to be fulfilled by any numerical method for (1.1). Our aim is now devoted to developing a theory of multistep methods inspired by the principles presented in the previous chapter. Hence, let us start with the following definition.

Definition 3.2 For a given grid function

$$v : \mathcal{I}_h \rightarrow \mathbb{R}^d,$$

let us denote by v_n its value in $t_n \in \mathcal{I}_h$. We define the *numerical residual operator* associated to (3.1) as

$$R_h(v_n) = \frac{1}{h} \sum_{j=0}^k \alpha_j v_{n+j} - \sum_{j=0}^k \beta_j f_{n+j}, \quad (3.8)$$

for $n = 0, 1, \dots, N - k$, with $f_{n+j} = f(t_{n+j}, v_{n+j})$.

As it happens for one-step methods, when the residual operators (2.24) and (3.8) are respectively evaluated in the exact solution of (1.1) and its numerical approximation computed by (3.1), we have

$$R(y) = 0, \quad R_h(y_n) = 0.$$

The numerical residual operator (3.8) evaluated in the exact solution $y(t)$ of (1.1) provides the local truncation error associated to (3.1), having the following expression:

$$\begin{aligned} T(t_n, y(t_n); h) &= R_h(y(t_n)) = \frac{1}{h} \sum_{j=0}^k \alpha_j y(t_{n+j}) - \sum_{j=0}^k \beta_j f(t_{n+j}, y(t_{n+j})) \\ &= \frac{1}{h} \sum_{j=0}^k \alpha_j y(t_{n+j}) - \sum_{j=0}^k \beta_j y'(t_{n+j}). \end{aligned} \quad (3.9)$$

Correspondingly, we give the following definition.

Definition 3.3 A linear multistep method (3.1) is *consistent* if, for any $(t, y) \in [t_0, T] \times \mathbb{R}^d$,

$$\lim_{h \rightarrow 0} T(t, y; h) = 0.$$

Clearly, a consistent method (3.1) satisfies

$$T(t, y; h) = O(h^p),$$

with $p \geq 1$. In other terms, consistent methods have at least order 1. The notion of order of a LMM is given in the following definition.

Definition 3.4 A linear multistep method (3.1) has *order* p if, for a chosen vector norm $\| \cdot \|$, there exists a real constant $C > 0$ such that

$$\|T(t, y; h)\| \leq Ch^p,$$

for any $(t, y) \in [t_0, T] \times \mathbb{R}^d$, where C is independent on t, y and h .

It is the right moment to address a very crucial point. Analyzing accuracy and stability properties of numerical methods can be very tricky if we only rely on their definitions. However, the work of the pioneers of the numerical approximation of ODEs has led to highly effective tools which make the analysis much simpler, since it only requires algebraic computation involving the coefficients of the methods. As regards LMMs, seminal contributions in this direction have been provided through the talent and the ingenious work of Germund Dahlquist (1925–2005). Let us briefly present his biography, based on the obituary written by Åke Björck, Bill Gear and Gustaf Söderlind in Siam News of May 1st, 2005 and on the information reported in the gifted MacTutor History of Mathematics Archive (<https://mathshistory.st-andrews.ac.uk/Biographies/Dahlquist/>).

A Portrait of Germund Dahlquist

Germund Dahlquist is one of the pioneers in establishing a theory for the numerical discretization of differential equations. He was born in 1925 in Uppsala, son of a minister in the Church of Sweden (his father) and a poet (his mother). He studied mathematics at Stockholm University since 1942 and was strongly influenced by one of his professors, Harald Bohr (brother of Niels Bohr, the famous Danish physicist who achieved the Nobel Prize in 1922). Bohr was a refugee from Denmark during the Second World War and inspired Dahlquist a lot not only as regards Mathematics, but also because of his gifted character that made of him a professor highly dedicated to his students (as well as a gifted soccer player. He was member of the Danish national football team and achieved a silver medal in the 1908 Summer Olympics).

He graduated in 1949, but he did not promptly start a Ph.D. program: indeed, he was appointed at the Swedish Board of Computer Machinery as an

(continued)

applied mathematician and programmer. In 1951 Sweden developed a digital computer (its name was BESK, the acronym for Binary Electronic Sequential Calculator), which came into operation in December 1953: 1951 is a crucial year for us, since Germund Dahlquist started his studies leading to *ground-breaking contributions to the theory of numerical methods for initial value problems in ordinary differential equations*, as written in his obituary.

BESK was employed by Dahlquist to solve differential equations, clearly after a proper study of difference methods. His theoretical study was also accompanied by his membership in with a team working on numerical weather forecasts, guided by the Swedish-American meteorologist Carl Gustaf Rossby at the International Meteorological Institute of Stockholm University. This working group was able to develop in 1954 the first 24-hour weather observations made the same day, carried out on BESK.

That was a very fruitful time for Dahlquist research activity, which led to his first publications in numerical analysis. In 1956, Dahlquist presented his studies on linear multistep methods, giving rise to the beautiful convergence theory for such methods. His theory parallels that of Peter Lax, that introduced his equivalence principle in 1955 had established the Lax principle.

From 1956 to 1959 Dahlquist covered the position head of Mathematical Analysis and Programming Development at the Swedish Board of Computer Machinery. He defended his Ph.D. thesis in 1958, entitled “Stability and Error Bounds in the Numerical Solution of Ordinary Differential Equations”, advised by Fritz Carlson. In his thesis he also introduced the logarithmic norm, independently developed also by Lozinskii in 1958, explained in Definition 1.3. The theory introduced in these years was spread out by Peter Henrici in 1962, through his monograph, a masterpiece for the modern theory of numerical discretization of ODEs.

In 1959 Dahlquist he was appointed to the Royal Institute of Technology in Stockholm, where he spent the rest of his career and where the Department of Numerical Analysis and Computer Science was founded in 1962 as an offshoot the Department of Applied Mathematics. In these years he was pioneer also in establishing the new-born journal BIT Numerical Mathematics, published for the first time in 1961, served by Dahlquist as editor for more than 30 years. BIT published several relevant contributions by Germund Dahlquist, such as the highly cited masterpiece on A-stability (a concept we will later introduce in next chapters) on the famous Dahlquist barriers.

In 1963 he got his position as Full Professor of “Computer Sciences, in particular Numerical Analysis”, actually the first full professorship position of this kind in Sweden. His highly appreciated book *Numeriska Metoder*, co-authored by Åke Björck appeared in 1969 and a revised extended version entitled “Numerical Methods” was published in 1974 by Prentice-Hall [113]. This book had a great success all over the world: it was translated in German

(continued)

in 1972, in Polish in 1983, in Chinese in 1990. Nick Higham writes in his review of the book:

This work is a monumental undertaking and represents the most comprehensive textbook survey of numerical analysis to date. It will be an important reference in the field for many years to come.

During the 1960s and 1970s, Dahlquist visited many institutes in Europe, USA, Australia, New Zealand and China. He visited Stanford University in 1968 and 1977–1978, where he held a five-year part-time position from 1982 to 1986.

The Society for Industrial and Applied Mathematics (SIAM) named him their John von Neumann lecturer in 1988. Germund Dahlquist retired from the Royal Institute of Technology in 1990, but continued in actively working on research. He obtained honorary doctorates from Hamburg University (1981), Helsinki University (1994), and Linköping University (1996). In 1999 he achieved the prestigious Peter Henrici Prize, with the following motivation:

He has created the fundamental concepts of stability, A-stability and the nonlinear G-stability for the numerical solution of ordinary differential equations. He succeeded, in an extraordinary way, to relate stability concepts to accuracy and proved the deep results which are nowadays called the first and second Dahlquist barrier. His interests, like Henrici's, are very broad, and he contributed significantly to many parts of numerical analysis. As a human being and scientist, he gives freely of his talent and knowledge to others and remains a model for many generations of scientists to come.

In 1995, on the occasion of his 70th birthday, SIAM established the Germund Dahlquist Prize to be awarded biennially, normally to a young scientist for original contributions to fields related to the numerical solution of differential equations and numerical methods for scientific computing.

In his obituary, two more significant aspects arises: his active work for Amnesty International and his love of music. Here is an excerpt:

As an active member of Amnesty International during the 1970s, Dahlquist worked to help scientists who were politically persecuted, in some cases traveling to offer his encouragement and recognition in person. He used to tell the story of his intervention on behalf of a Russian mathematician who, in despair, had made a thoughtless public statement to the effect that the Soviet Union was “a land of alcoholics”. Guriy I Marchuk, who had visited Stockholm University in the 1960s, was then president of the USSR Academy of Sciences and vice-chair of the USSR Council of Ministers. Dahlquist wrote to Marchuk pleading the dissident's case. After a long time with no response, two staff members of the Soviet Embassy called at Germund's office one day, bringing greetings from Marchuk and a package, that turned out to contain... two bottles of vodka! Germund had a keen interest in music, mainly classical but also jazz music. He would often happily sit down at the piano and entertain his colleagues with a few old standards, starting with “On the Sunny Side of the Street” and ending with “As Time Goes By”. But his knowledge went much deeper. On one visit to the

(continued)

USA, with a few colleagues in a fine restaurant, Germund heard a female bar pianist whose music was obviously the highlight of the evening for him. When it was time to leave, Germund told the pianist how much he had enjoyed her stylish playing, adding that it had reminded him of one of his favorites, the great jazz pianist Art Tatum. The pianist was duly flattered, but it was Germund who was surprised when she answered: “Art Tatum was my father!”.

Germund Dahlquist died in 2005. His work inspired many researches over several decades. As John Butcher wrote:

When I met him in 1970, I started to appreciate that he was more than a brilliant mathematician and computational scientist: he was a kind and sensitive man and a loyal friend. [...] Everything Germund published was a separate gem, exhibiting deep mathematical insight and, at the same time, a clear understanding of sound computational practice. He was a pioneer who remained a central figure throughout his career; he will be sadly missed.

Let us recast (3.9) in the following form

$$\sum_{j=0}^k \alpha_j y(t_{n+j}) - h \sum_{j=0}^k \beta_j y'(t_{n+j}) = hT(t_n, y(t_n); h).$$

Such a relation, though defined on vector valued functions, actually results to be the same on each component of the involved vectors. For this reasons, it makes sense to analyze it on scalar functions, motivating the following definition.

Definition 3.5 For a given scalar function $z: [t_0, T] \rightarrow \mathbb{R}$ of class $C^1([t_0, T])$, the *linear difference operator* associated to a linear multistep method (3.1) is defined by

$$\mathcal{L}[z(t), h] = \sum_{j=0}^k \alpha_j z(t + jh) - h \sum_{j=0}^k \beta_j z'(t + jh). \quad (3.10)$$

Clearly, if $y(t)$ is the solution of (1.1), we have

$$\mathcal{L}[y_i(t), h] = hT(t, y_i(t); h), \quad i = 1, 2, \dots, d$$

and, as a consequence, if the method (3.1) has order p , $\mathcal{L}[z(t), h] = \mathcal{O}(h^{p+1})$, for any scalar function $z(t)$. This observation leads to the following result.

Theorem 3.2 A linear multistep method (3.1) has order p if and only if $C_\ell = 0$, for any $\ell = 0, 1, \dots, p$, with

$$C_0 = \sum_{j=0}^k \alpha_j, \quad C_\ell = \sum_{j=0}^k \left(\frac{j^\ell}{\ell!} \alpha_j - \frac{j^{\ell-1}}{(\ell-1)!} \beta_j \right), \quad \ell > 0$$

and $C_{p+1} \neq 0$.

Proof We expand in Taylor series around t each evaluation of $z(t)$ appearing in the right-hand side of (3.10), obtaining

$$\begin{aligned} \mathcal{L}[z(t), h] = & \sum_{j=0}^k \alpha_j \left(z(t) + \sum_{\ell \geq 1} \frac{(jh)^\ell}{\ell!} z^{(\ell)}(t) \right) \\ & - h \sum_{j=0}^k \beta_j \left(z'(t) + \sum_{\ell > 1} \frac{(jh)^{\ell-1}}{(\ell-1)!} z^{(\ell)}(t) \right). \end{aligned}$$

Collecting in powers of h leads to

$$\mathcal{L}[z(t), h] = \left(\sum_{j=0}^k \alpha_j \right) z(t) + \sum_{\ell \geq 1} \sum_{j=0}^k \left(\frac{j^\ell}{\ell!} \alpha_j - \frac{j^{\ell-1}}{(\ell-1)!} \beta_j \right) h^\ell z^{(\ell)}(t),$$

i.e.,

$$\mathcal{L}[z(t), h] = C_0 z(t) + \sum_{\ell \geq 1} C_\ell h^\ell z^{(\ell)}(t).$$

In order to have $\mathcal{L}[z(t), h] = \mathcal{O}(h^{p+1})$, we need to satisfy

$$C_0 = C_1 = \dots = C_p = 0, \quad C_{p+1} \neq 0,$$

obtaining the thesis. □

Definition 3.6 The non-zero constant

$$C_{p+1} = \sum_{j=0}^k \left(\frac{j^{p+1}}{(p+1)!} \alpha_j - \frac{j^p}{p!} \beta_j \right)$$

of an order p method (3.1) is denoted as its *error constant*.

In other terms, for an order p linear multistep method we have

$$\mathcal{L}[z(t), h] = C_{p+1} h^{p+1} z^{(p+1)}(t) + \mathcal{O}(h^{p+2}),$$

for all regular scalar functions $z(t)$, while the corresponding local truncation error is

$$T(t, y(t); h) = C_{p+1} h^p y^{(p+1)}(t) + \mathcal{O}(h^{p+1}).$$

The following corollary of Theorem 3.2 gives us a very immediate way to analyze the consistency of a LMM (3.1), only requiring a straightforward algebraic computation involving the coefficients of the method. This way of proceeding is certainly much simpler than proving consistency using Definition 3.3.

Corollary 3.1 A linear multistep method (3.1) is consistent if and only if

$$\sum_{j=0}^k \alpha_j = 0, \quad \sum_{j=0}^k (j\alpha_j - \beta_j) = 0. \quad (3.11)$$

Proof A consistent method has order at least one. Hence, according to Theorem 3.2, it satisfies

$$C_0 = \sum_{j=0}^k \alpha_j = 0, \quad C_1 = \sum_{j=0}^k (j\alpha_j - \beta_j) = 0$$

and the thesis holds true. \square

Let us now present few examples of applications of the notions introduced in this section.

Example 3.4 Let us analyze consistency, orders and error constants of the examples of LMMs (3.1) we have developed so far, depending on one and two steps.

- The explicit Euler method (2.19) satisfies

$$C_0 = \sum_{j=0}^k \alpha_j = 0, \quad C_1 = \sum_{j=0}^k (j\alpha_j - \beta_j) = 0,$$

$$C_2 = \sum_{j=0}^k \left(\frac{j^2}{2} \alpha_j - j\beta_j \right) = \frac{1}{2},$$

so it is consistent, of order 1 and its error constant is equal to $1/2$;

- for the implicit Explicit Euler method (2.32) we have

$$C_0 = \sum_{j=0}^k \alpha_j = 0, \quad C_1 = \sum_{j=0}^k (j\alpha_j - \beta_j) = 0,$$

$$C_2 = \sum_{j=0}^k \left(\frac{j^2}{2} \alpha_j - j\beta_j \right) = -\frac{1}{2}.$$

Then, it is consistent, of order 1 and its error constant is equal to $-1/2$;

- the trapezoidal method (2.33) fulfills

$$C_0 = \sum_{j=0}^k \alpha_j = 0, \quad C_1 = \sum_{j=0}^k (j\alpha_j - \beta_j) = 0,$$

$$C_2 = \sum_{j=0}^k \left(\frac{j^2}{2} \alpha_j - j\beta_j \right) = 0, \quad C_3 = \sum_{j=0}^k \left(\frac{j^3}{6} \alpha_j - \frac{j^2}{2} \beta_j \right) = -\frac{1}{12},$$

so it is consistent, of order 2 and its error constant is $-1/12$;

- Milne-Simpson method (3.2) is a LMM (3.1) with $\alpha_0 = -1$, $\alpha_1 = 0$,

(continued)

Example 3.4 (continued)

$\alpha_2 = 1$, $\beta_0 = 1/3$, $\beta_1 = 4/3$ and $\beta_2 = 1/3$. Hence, it satisfies

$$\begin{aligned} C_0 &= \sum_{j=0}^k \alpha_j = 0, & C_1 &= \sum_{j=0}^k (j\alpha_j - \beta_j) = 0, \\ C_2 &= \sum_{j=0}^k \left(\frac{j^2}{2} \alpha_j - j\beta_j \right) = 0, & C_3 &= \sum_{j=0}^k \left(\frac{j^3}{6} \alpha_j - \frac{j^2}{2} \beta_j \right) = 0, \\ C_4 &= \sum_{j=0}^k \left(\frac{j^4}{24} \alpha_j - \frac{j^3}{6} \beta_j \right) = 0, & C_5 &= \sum_{j=0}^k \left(\frac{j^5}{120} \alpha_j - \frac{j^4}{24} \beta_j \right) = -\frac{1}{90}, \end{aligned}$$

so it is consistent, of order 4, with error constant $-1/90$;

- the two-step Adams-Bashforth method (3.3) is a LMM (3.1) with $\alpha_0 = 0$, $\alpha_1 = -1$, $\alpha_2 = 1$, $\beta_0 = -1/2$, $\beta_1 = 3/2$ and $\beta_2 = 0$. Therefore,

$$\begin{aligned} C_0 &= \sum_{j=0}^k \alpha_j = 0, & C_1 &= \sum_{j=0}^k (j\alpha_j - \beta_j) = 0, \\ C_2 &= \sum_{j=0}^k \left(\frac{j^2}{2} \alpha_j - j\beta_j \right) = 0, & C_3 &= \sum_{j=0}^k \left(\frac{j^3}{6} \alpha_j - \frac{j^2}{2} \beta_j \right) = \frac{5}{12}, \end{aligned}$$

so it is consistent, of order 2 and its error constant is equal to $5/12$.

Example 3.5 We now aim to study the consistency of the following numerical method

$$y_{n+2} - 2y_{n+1} + y_n = hf_n, \quad (3.12)$$

both using conditions (3.11) and through an experimental check. Equation (3.12) provides an explicit two-step method and, according to Corollary 3.1 it is a non-consistent, since

$$C_0 = \sum_{j=0}^k \alpha_j = 0, \quad C_1 = \sum_{j=0}^k (j\alpha_j - \beta_j) = -1.$$

(continued)

Example 3.5 (continued)

Let us experimentally check this property. To this purpose, consider the following scalar test problem

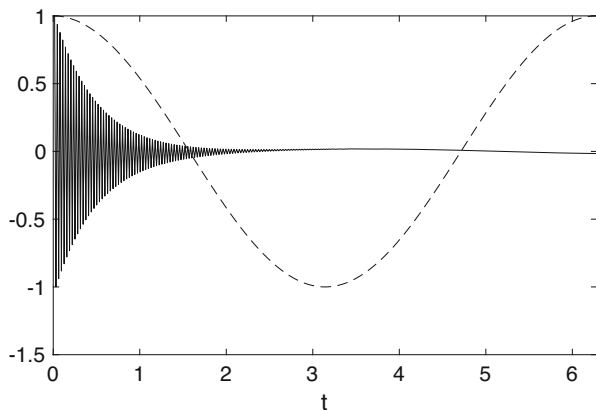
$$\begin{cases} y'(t) = 2(y(t) - \cos(t)) - \sin(t), & t \in [0, 2\pi], \\ y(0) = 1, \end{cases} \quad (3.13)$$

whose exact solution is $y(t) = \cos(t)$. As visible in Fig. 3.1, the numerical solution does not match the exact one and the reader can verify that a similar behavior occurs also if the stepsize is reduced. This is a typical situation of lack of consistency: the application of a non-consistent method leads to the pattern of another function rather than one reproducing the exact solution.

3.4 Zero-Stability

We have learned in Chap. 2 that consistency is a local accuracy property, while zero-stability and convergence are global accuracy properties. We now focus on zero-stability analysis that, according to our analysis in Chap. 2, ensures the boundedness of the error for small values of the stepsize. We first need to introduce the following tools.

Fig. 3.1 Numerical (straight line) vs exact (dashed line) solutions of (3.13). The numerical solution is computed by the non-consistent method (3.12) with stepsize $h = \pi/100$



Definition 3.7 The *first characteristic polynomial* associated to a linear multistep method (3.1) is given by

$$\rho(z) = \sum_{j=0}^k \alpha_j z^j, \quad (3.14)$$

while the *second characteristic polynomial* associated to (3.1) is given by

$$\sigma(z) = \sum_{j=0}^k \beta_j z^j. \quad (3.15)$$

These polynomials are very useful also to analyze consistency of (3.1). Indeed, for a consistent method, we have

$$\rho(1) = 0, \quad \rho'(1) = \sigma(1),$$

since

$$\rho(1) = \sum_{j=0}^k \alpha_j, \quad \rho'(1) = \sum_{j=0}^k j \alpha_j, \quad \sigma(1) = \sum_{j=0}^k \beta_j.$$

The roots of the first characteristic polynomial are important to analyze the zero-stability of (3.1). To this purpose, there is a relevant property that we are going to use, defined as follows.

Definition 3.8 An algebraic polynomial satisfies the *root condition* if each of its roots has modulus strictly less than 1 or has modulus one but it is simple.

Example 3.6 The polynomial

$$\rho(z) = z^2 - \frac{5}{6}z + \frac{1}{6}$$

satisfies the root condition, since its roots are $1/2$ and $1/3$. The polynomial

(continued)

Example 3.6 (continued)

$$\rho(z) = z^2 - \frac{4}{3}z + \frac{1}{3}$$

also satisfies the root condition, since its roots are $1/3$ and 1 . Finally, the polynomial

$$\rho(z) = z^3 - \frac{5}{2}z^2 + 2z - \frac{1}{2}$$

does not satisfy the root condition, since its roots are $1/2$ and 1 , the latter with multiplicity 2 .

As we know from Chap. 2, zero-stability ensures that the numerical solution does not blow-up as h tends to 0 ; in other terms, we need to prove that the general solution of the difference equation describing (3.1) does not blow-up as h goes to 0 . Hence, we first have to analyze what happens to the solution of a linear difference equation; this issue is explained by the following result, reported for the scalar case, whose proof is here omitted (the interested reader can refer, for instance, to [170]).

Theorem 3.3 *Consider the following order k inhomogeneous linear difference equation*

$$\sum_{j=0}^k \alpha_j y_{n+j} = g_{n+k},$$

where $\alpha_j, y_{n+j} \in \mathbb{R}$, $j = 0, 1, \dots, k$, and $g_{n+k} \in \mathbb{R}$. Then, there exists $M > 0$ independent on n such that

$$|y_n| \leq M \left(\max_{0 \leq i \leq k-1} |y_i| + \sum_{m=k}^n |g_m| \right), \quad n \geq 0,$$

if and only if the characteristic polynomial

$$\rho(z) = \sum_{j=0}^k \alpha_j z^j$$

satisfies the root condition.

According to Theorem 3.3, the root condition is a necessary and sufficient condition for the stability of the solutions of linear difference equations. We now prove that the root condition of the first characteristic polynomial (3.14) is exactly what we need to have stable numerical solutions through LMMs (3.1). First of all, let us provide a rigorous definition of zero-stability.

Consider the vector

$$\mathbf{y}_h = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_N \end{bmatrix} \quad (3.16)$$

collecting the numerical solution of the initial value problem (1.1) in each grid point, computed by (3.1) and the vector

$$\tilde{\mathbf{y}}_h = \begin{bmatrix} \tilde{y}_0 \\ \tilde{y}_1 \\ \vdots \\ \tilde{y}_N \end{bmatrix}$$

of the numerical approximations of the solution to the perturbed problem (1.12), obtained by (3.1). As in the one-step case described in Chap. 2, $R_h(y_n) = 0$, while we suppose that $R_h(\tilde{y}_n) = \varepsilon_n$ and collect all these values in the vector

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{bmatrix}.$$

We also denote by $\boldsymbol{\delta}$ the vector collecting the deviations in the initial values, i.e.,

$$\boldsymbol{\delta} = \begin{bmatrix} y_0 - \tilde{y}_0 \\ y_1 - \tilde{y}_1 \\ \vdots \\ y_{k-1} - \tilde{y}_{k-1} \end{bmatrix}.$$

We define zero-stability of LMMs (3.1) as follows.

Definition 3.9 A linear multistep method (3.1) is *zero-stable* if there exists $\Lambda > 0$ such that, for any value of the stepsize $h \in [0, h_0]$, with $h_0 > 0$, the following stability inequality holds true

$$\|\mathbf{y}_h - \tilde{\mathbf{y}}_h\|_\infty \leq \Lambda(\|\delta\|_\infty + \|\varepsilon\|_\infty). \quad (3.17)$$

As already observed in Chap. 2 for one-step methods, the zero-stability inequality (3.17) imitates the corresponding inequality (1.15), useful to study the continuous dependence on the initial data and the vector field of the underlying initial value problem.

Let us now prove the following zero-stability criterion, whose statement recalls the stability result for difference equations presented in Theorem (3.3).

Theorem 3.4 *A linear multistep method (3.1) is zero-stable if and only if its first characteristic polynomial (3.14) satisfies the root condition.*

Proof We separately prove the necessity and the sufficiency of the root condition for the zero-stability of (3.1).

- First part: let us prove that zero-stability implies the root condition for its first characteristic polynomial. Suppose that the vector field f of the continuous problem (1.1) is identically null. The corresponding LMM (3.1) reads

$$\sum_{j=0}^k \alpha_j y_{n+j} = 0. \quad (3.18)$$

Let us collect the numerical solution arising from (3.18) in the vector \mathbf{y}_h given by (3.16) and consider that the homogeneous difference equation (3.18) also admits the zero solution. Then, from the zero-stability hypothesis, there exists $\Lambda > 0$ such that

$$\|\mathbf{y}_h\|_\infty \leq \Lambda \max_{0 \leq j \leq k-1} |y_j|.$$

Hence, the solution \mathbf{y}_h of the difference equation (3.18) is uniformly bounded and, by Theorem 3.3, its characteristic polynomial

$$\rho(z) = \sum_{j=0}^k \alpha_j z^j$$

satisfies the root condition.

- Second part: let us prove that the root condition for the first characteristic polynomial of (3.1) implies its zero-stability. Let us consider two given grid functions

$$u, v : \mathcal{I}_h \rightarrow \mathbb{R}^d$$

and denote by u_n and v_n their values in $t_n \in \mathcal{I}_h$, respectively. Correspondingly,

$$\begin{aligned} \sum_{j=0}^k \alpha_j u_{n+j} &= h \sum_{j=0}^k \beta_j f(t_{n+j}, u_{n+j}) + h R_h(u_n), \\ \sum_{j=0}^k \alpha_j v_{n+j} &= h \sum_{j=0}^k \beta_j f(t_{n+j}, v_{n+j}) + h R_h(v_n). \end{aligned}$$

By subtraction, we have

$$\sum_{j=0}^k \alpha_j (u_{n+j} - v_{n+j}) = g_{n+k}, \quad (3.19)$$

where

$$g_{n+k} = h \sum_{j=0}^k \beta_j (f(t_{n+j}, u_{n+j}) - f(t_{n+j}, v_{n+j})) + h (R_h(u_n) - R_h(v_n)).$$

In other terms, $\{u_n - v_n\}_{n \in \mathbb{N}}$ is solution of the inhomogeneous difference equation (3.19). Since the root condition holds true, by Theorem 3.3 there exists $M > 0$, independent on n , such that

$$\|u_n - v_n\|_\infty \leq M \left(\max_{0 \leq i \leq k-1} \|u_i - v_i\|_\infty + \sum_{m=k}^n \|g_m\|_\infty \right).$$

Certainly, by defining $(r_h)_n = R_h(u_n) - R_h(v_n)$, we have

$$\begin{aligned} \|g_m\|_\infty &= \left\| h \sum_{j=0}^k \beta_j (f(t_{m+j-k}, u_{m+j-k}) - f(t_{m+j-k}, v_{m+j-k})) + h (r_h)_{m-k} \right\|_\infty \\ &\leq h L \beta \sum_{j=0}^k \|u_{m+j-k} - v_{m+j-k}\|_\infty + h \|r_h\|_\infty, \end{aligned}$$

being $\beta = \max_{0 \leq j \leq k} \beta_j$. By defining $e_n = u_n - v_n$, we have

$$\begin{aligned} \|e_n\|_\infty &\leq M \left(\max_{0 \leq i \leq k-1} \|e_i\|_\infty + \sum_{m=k}^n \left(h L \beta \sum_{j=0}^k \|e_{m+j-k}\|_\infty + h \|r_h\|_\infty \right) \right) \\ &\leq M \left(\max_{0 \leq i \leq k-1} \|e_i\|_\infty + h L \beta \sum_{m=k}^n \sum_{j=0}^k \|e_{m+j-k}\|_\infty + N h \|r_h\|_\infty \right). \end{aligned}$$

Clearly,

$$\sum_{m=k}^n \sum_{j=0}^k \|e_{m+j-k}\|_\infty \leq \sum_{j=0}^k \sum_{m=0}^n \|e_m\|_\infty = (k+1) \sum_{m=0}^n \|e_m\|_\infty.$$

Hence,

$$\|e_n\|_\infty \leq M \left(\max_{0 \leq i \leq k-1} \|e_i\|_\infty + h L \beta (k+1) \sum_{m=0}^n \|e_m\|_\infty + (T-t_0) \|r_h\|_\infty \right)$$

or, equivalently,

$$\mu \|e_n\|_\infty \leq M \left(\max_{0 \leq i \leq k-1} \|e_i\|_\infty + h L \beta (k+1) \sum_{m=0}^{n-1} \|e_m\|_\infty + (T-t_0) \|r_h\|_\infty \right),$$

with $\mu = (1 - h L \beta (k+1))$. Suppose to choose h small enough in order to make $\mu > 0$ (it is enough to choose $h < 1/(h L \beta (k+1))$ to make this possible), so that

$$\|e_n\|_\infty \leq h A \sum_{m=0}^{n-1} \|e_m\|_\infty + B,$$

with

$$A = \frac{M}{\mu} L \beta (k+1), \quad B = \frac{M}{\mu} \left(\max_{0 \leq i \leq k-1} \|e_i\|_\infty + (T-t_0) \|r_h\|_\infty \right).$$

Let us consider the corresponding difference equation

$$w_n = h A \sum_{m=0}^{n-1} w_m + B,$$

with initial value $w_0 = B$. The reader can easily prove by induction that its solution is

$$w_n = B(1 + hA)^n, \quad n \geq 0.$$

Then

$$\|e_n\|_\infty - w_n \leq h A \sum_{m=0}^{n-1} (\|e_m\|_\infty - w_m)$$

and the reader can again prove by induction that

$$\|e_n\|_\infty \leq w_n.$$

Therefore,

$$\|e_n\|_\infty \leq B(1 + hA)^n \leq B e^{nhA} \leq B e^{NhA} = B e^{(T-t_0)A}$$

and replacing the value of B gives

$$\|e_n\|_\infty \leq \frac{M}{\mu} e^{(T-t_0)A} \left(\max_{0 \leq i \leq k-1} \|e_i\|_\infty + (T - t_0) \|r_h\|_\infty \right).$$

The choice

$$\Lambda = \frac{M}{\mu} e^{(T-t_0)A} \max\{1, T - t_0\}$$

let the thesis hold true. □

Analyzing zero-stability through Definition 3.9 may be very tricky. However, Theorem 3.4 provides a practical condition that makes the analysis of zero-stability much easier. Let us test such a condition on few examples.

Example 3.7 According to Theorem 3.4,

- Euler methods (2.19) and (2.32) are zero-stable, since their first characteristic polynomial is

$$\rho(z) = z - 1;$$

(continued)

Example 3.7 (continued)

- the trapezoidal method (2.33) is zero-stable, since its first characteristic polynomial is

$$\rho(z) = z - 1;$$

- Milne-Simpson method (3.2) is zero-stable, since its first characteristic polynomial is

$$\rho(z) = z^2 - 1;$$

- the two-step Adams-Bashforth method (3.3) is zero-stable, since its first characteristic polynomial is

$$\rho(z) = z^2 - z.$$

Example 3.8 We now aim to study the zero-stability of the following numerical method

$$y_{n+2} - 3y_{n+1} + 2y_n = hf_n, \quad (3.20)$$

both checking the root condition and experimentally. This method is not zero-stable, since the roots of its first characteristic polynomial

$$\rho(z) = z^2 - 3z + 2$$

are 1 and 2, so the root condition is not satisfied. Let us experimentally check this lack of zero-stability on the test problem (3.13). As visible in Fig. 3.2, the numerical solution does not have a stable behavior, so it does not match the stable character of the exact solution. Hence, the applied method is clearly not zero-stable.

For a zero-stable linear multistep method (3.1), the following result holds true.

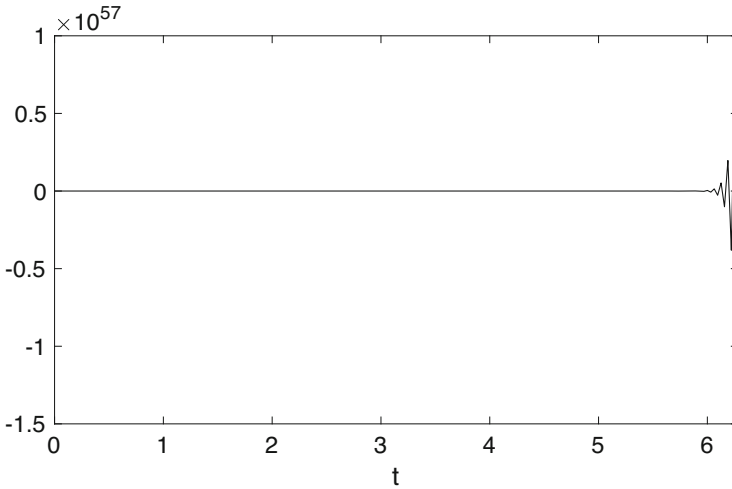


Fig. 3.2 Numerical solution of (3.13) computed by method (3.20), that is not zero-stable, with stepsize $h = \pi/100$

Theorem 3.5 *The order p of a zero-stable linear multistep method (3.1) depending on k steps satisfies*

$$p \leq \begin{cases} k + 1, & k \text{ odd,} \\ k + 2, & k \text{ even.} \end{cases}$$

This results, whose proof can be found, for instance, in [67], is an order barrier for linear multistep methods, well-known in the literature as *first Dahlquist barrier*. In other terms, the number of steps provides an upper bound for the order of convergence of the corresponding method. Maximal order methods are those of order $k + 1$ if k is odd, $k + 2$ if k is even. An example of maximal order method is the Milne-Simpson method (3.2), which depends on 2 steps and has order 4. However, maximal order methods can have poor stability properties, as it is the case of Milne-Simpson method itself, so we should provide a reasonable balance between accuracy and stability properties. We will analyze this aspect in Chap. 6.

3.5 Convergence

Let us now turn our attention to analysis of convergence for linear multistep methods (3.1), starting from the following definition.

Definition 3.10 Suppose that (3.16) is the vector collecting the numerical approximations of the solution to the continuous problem (1.1) in each grid point of the uniform grid (2.1), obtained by the method (3.1) and also consider the vector of the corresponding exact values

$$\mathbf{v}_h = \begin{bmatrix} y(t_0) \\ y(t_1) \\ \vdots \\ y(t_N) \end{bmatrix}.$$

Denote by

$$s_h = \begin{bmatrix} y_1 \\ \vdots \\ y_{k-1} \end{bmatrix}$$

the vector collecting the missing starting values computed by a proper starting procedure. Then, the LMM (3.1) is *convergent* if

$$\lim_{h \rightarrow 0} \|\mathbf{v}_h - \mathbf{y}_h\|_\infty = 0,$$

whenever

$$\lim_{h \rightarrow 0} (s_h)_i = y(t_i), \quad i = 1, 2, \dots, k - 1.$$

It is worth underlining that a significant role is played by the starting procedure, which is assumed to provide accurate starting values in the above given definition of convergence. Again, as aforementioned for consistency and zero-stability, convergence analysis through its definition may not be an easy task. However, there is a very powerful result, according to which convergence is equivalent to consistency plus zero-stability. In such a way, convergence analysis only relies on very simple calculations involving the coefficients of the method. This result originally belongs to a huge masterpiece due to Peter D. Lax (Budapest, 1926), specialized to LMMs by Germund Dahlquist. Peter Lax is a mathematician born in Hungary, Professor at

Courant Institute of Mathematical Sciences at New York University, winner of the prestigious Abel Prize in 2005.

Theorem 3.6 (Lax Equivalence Theorem) *A linear multistep method (3.1) is convergent if and only if it is consistent and zero-stable.*

Proof We separately prove the necessity and the sufficiency parts of the theorem.

- First part: we prove that convergence implies consistency and zero-stability. Consider the initial value problem (1.1) with f identically zero and $y_0 = 0$. Then, $y(t) = 0$. By contradiction, suppose that the method is not zero-stable: then, there exists a root \bar{z} of the first characteristic polynomial of (3.1) such that $|\bar{z}| > 1$ or $|\bar{z}| = 1$ and its multiplicity is greater than 1. In the first case, the solution of (3.1), thought as a difference equation, contains a term

$$c\bar{z}^n, \quad c \in \mathbb{R}$$

tending to infinity, which contradicts the hypothesis of convergence. The case $|\bar{z}| = 1$ with multiplicity greater than 1 is similar and left to the reader.

Let us now prove that convergence implies consistency, by proving that consistency conditions arise from the exactness of the method on the monomial basis $\{1, t\}$ (this choice is connected to Exercise 4, Sect. 3.6).

- (i) Consider the initial value problem (1.1) with f identically zero and $y_0 = 1$. Then, $y(t) = 1$. The corresponding LMM (3.1) is given by

$$\sum_{j=0}^k \alpha_j y_{n+j} = 0$$

and, assuming $y_1 = y_2 = \dots = y_{k-1}$, convergence yields

$$\sum_{j=0}^k \alpha_j = 0,$$

that is the first consistency condition.

- (ii) Consider the initial value problem (1.1) with f identically equal to 1 and $y_0 = 0$. Then, $y(t) = t - t_0$. The corresponding LMM (3.1) is given by

$$\sum_{j=0}^k \alpha_j y_{n+j} = h \sum_{j=0}^k \beta_j,$$

i.e.,

$$\sum_{j=0}^k \alpha_j y_{n+j} = h\sigma(1). \quad (3.21)$$

Certainly,

$$y_n = \frac{\sigma(1)}{\rho'(1)}nh$$

is solution of (3.21), as the reader can easily check. By convergence, y_n converges to nh and, as a consequence,

$$\frac{\sigma(1)}{\rho'(1)} = 1$$

that is the second consistency condition.

- Second part: we prove that consistency and zero-stability imply convergence. Since $R_h(y_n) = 0$ and $R_h(y(t)) = T(t, y(t); h)$, the zero-stability inequality (3.17) reads

$$\|\mathbf{v}_h - \mathbf{y}_h\|_\infty \leq \Lambda(\|\delta\|_\infty + \max_{0 \leq i \leq N} |T(t_i, y(t_i); h)|).$$

The right-hand side of last inequality goes to 0, because of the preliminary assumption on the starting procedure given in Definition 3.10 and the consistency assumption, leading to the thesis. \square

Example 3.9 Let us analyze the family of two-step implicit LMMs (3.1), given by

$$y_{n+2} + \alpha_1 y_{n+1} + \alpha_0 y_n = h(\beta_2 f_{n+2} + \beta_1 f_{n+1} + \beta_0 f_n). \quad (3.22)$$

- *Consistency.* Let us give consistency conditions (3.11) for this class of methods. We have

$$\alpha_0 + \alpha_1 + 1 = 0,$$

$$\alpha_1 + 2 = \beta_0 + \beta_1 + \beta_2.$$

(continued)

Example 3.9 (continued)

- *Order 2.* We obtain the constraints on the coefficients of the methods ensuring second order, by imposing the additional condition $C_2 = 0$ in Theorem (3.2), i.e.,

$$\frac{1}{2}\alpha_1 + 2 = \beta_1 + 2\beta_2.$$

Hence, order 2 methods (3.22) satisfy

$$\alpha_0 = 3 - 2\beta_1 - 4\beta_2,$$

$$\alpha_1 = -4 + 2\beta_1 + 4\beta_2,$$

$$\beta_0 = -2 + \beta_1 - 3\beta_2.$$

- *Order 3.* Third order methods also satisfy $C_3 = 0$ in Theorem (3.2), i.e.,

$$\frac{1}{6}\alpha_1 + \frac{4}{3} = \frac{1}{2}\beta_1 + 2\beta_2.$$

In summary, third order methods satisfy

$$\alpha_0 = -5 + 12\beta_2,$$

$$\alpha_1 = 4 - 12\beta_2,$$

$$\beta_0 = 2 - 11\beta_2.$$

$$\beta_1 = 4 - 8\beta_2.$$

- *Order 4.* Fourth order methods also fulfill $C_4 = 0$ in Theorem (3.2), i.e.,

$$\frac{1}{24}\alpha_1 + \frac{2}{3} = \frac{1}{6}\beta_1 + \frac{4}{3}\beta_2.$$

In summary, fourth order methods satisfy

$$\alpha_0 = 7, \quad \alpha_1 = -8, \quad \beta_0 = -9, \quad \beta_1 = -4, \quad \beta_2 = -1.$$

We observe that the maximal order method

$$y_{n+2} - 8y_{n+1} + 7y_n = -h(f_{n+2} + 4f_{n+1} + 9f_n)$$

(continued)

Example 3.9 (continued)

is not convergent because it is not-zero stable, since the zeros of the first characteristic polynomial

$$\rho(z) = z^2 - 8z + 7$$

are 1 and 7, then the root condition is not satisfied.

Clearly, due to Lax equivalence theorem, if a method is not convergent then consistency and/or zero-stability are missing. A numerical evidence of this aspect has already been given in Examples 3.5 and 3.8. We now aim to experimentally analyze the performances of a convergent method, namely the second-order Adams-Bashforth method (3.3), implemented in Program 3.1 by using the explicit Euler method (2.19) as a starting procedure.

Program 3.1 (Adams-Bashforth Method)

```
% Function implementing the second order Adams-Bashforth
% method on a uniform grid, for the numerical solution
% of a d-dimensional ODE.

% Inputs
% - problem: label of the problem to be solved;
% - tspan: vector of two components, storing the extrema
%         of the integration interval;
% - y0: initial value;
% - h: constant stepsize.

% Outputs
% - t: set of N equidistant grid points (tspan(1) excluded);
% - y: d×N matrix whose i-th column y(:,i) stores the
%     approximate value in the i-th grid point, i=1,2,...,N.

function [t,y]=AdamsBashforth(problem,tspan,y0,h)
N=(tspan(2)-tspan(1))/h;
t=linspace(h,tspan(2),N);
d=length(y0);
y=zeros(d,N);
fold=f(problem,tspan(1),y0);
% Starting value y(:,1) recovered by explicit Euler method
y(:,1)=y0+h*fold;
fnew=f(problem,t(1),y(:,1));
y(:,2)=y(:,1)-h*(fold-3*fnew)/2;
% The variables fold and fnew are introduced to reuse
% already computed function evaluations.
```

(continued)

Program 3.1 (continued)

```

fold=fnew;
fnew=f(problem,t(2),y(:,2));
for i=3:N
    y(:,i)=y(:,i-1)-h*(fold-3*fnew)/2;
    fold=fnew;
    fnew=f(problem,t(i),y(:,i));
end

```

Next example gives an experimental confirmation of the order of a numerical method. The involved order estimate is based on the application of the method with stepsize h , whose principal error term is given by $err(h) \approx C_{p+1}h^p$, and halved stepsize $h/2$, with principal error term $err(h/2) \approx C_{p+1}(h/2)^p$. The ratio between the two errors gives

$$\frac{err(h)}{err(h/2)} \approx 2^p,$$

i.e.,

$$p \approx \log_2 \left(\frac{err(h)}{err(h/2)} \right), \quad (3.23)$$

that provides an estimate of the order of convergence. Clearly, the smaller is h , the more the estimate of p is accurate.

Example 3.10 Consider the following van der Pol oscillator

$$\begin{cases} y_1'(t) = y_2(t), \\ y_2'(t) = (1 - y_1(t)^2)y_2(t) - y_1(t), \end{cases} \quad (3.24)$$

with $t \in [0, 10]$ and initial value $y_0 = [2 \ -2/3]^\top$. Let us numerically solve this problem by using the second order Adams-Bashforth method (3.3): the pattern of the solution is displayed in Fig. 3.3. We know that (3.4) is a convergent method; let us give an experimental evidence of this property in Table 3.1, where it is visible that the more the stepsize diminishes, the more the error decreases. The error is computed as

$$\|y_{AB} - y_{ODE45}\|_\infty,$$

(continued)

Example 3.10 (continued)

where $y_{AB} \approx y(10)$ is computed by (3.3) and y_{ODE45} is the solution in $t = 10$ computed by the Matlab built-in function `ode45`, with high accuracy, given by

$$y_{ODE45} = [-1.914027891764918 \quad 0.446099168376746]^T.$$

As visible, both the convergence and the order of convergence are confirmed in the experiments.

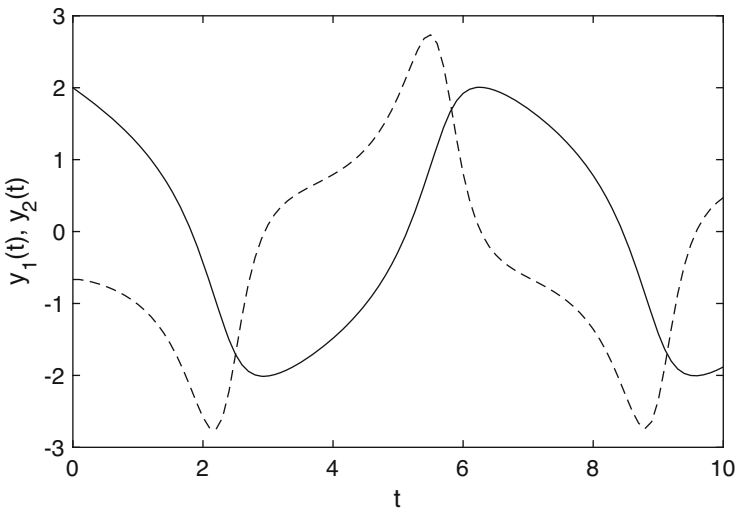


Fig. 3.3 Pattern of the solution of (3.24) computed by the Adams-Bashforth method (3.3), with $h = 0.1$. The plot of the component $y_1(t)$ is the straight line, while $y_2(t)$ is the dashed line

Table 3.1 Example 3.10: error in the final integration point associated to the application of the Adams-Bashforth method (3.4) to the van der Pol problem (3.24) and order estimation, computed as suggested by Eq. (3.23)

h	$\ y_{AB} - y_{ODE45}\ _\infty$	p
$h_0 = 0.1$	$3.10 \cdot 10^{-2}$	
$h_0/2$	$8.23 \cdot 10^{-3}$	1.91
$h_0/4$	$2.13 \cdot 10^{-3}$	1.95
$h_0/8$	$5.41 \cdot 10^{-4}$	1.97
$h_0/16$	$1.36 \cdot 10^{-4}$	1.99

3.6 Exercises

1. Analyze the family of explicit three-step methods (3.1), providing families of convergent methods of various orders. Is the maximal order method of the family also convergent?
2. Compute the values of $\vartheta_1, \vartheta_2 \in \mathbb{R}$ such that the two-parameter family of methods

$$y_{n+2} + (2\vartheta_1 - 3\vartheta_2)y_{n+1} - \left(\frac{5\vartheta_1}{2} - 2\right)y_n = \vartheta_1 h (f_n + f_{n+1})$$

reduces to a single convergent method. Does the corresponding method also achieve maximal order?

3. As seen in Example 3.4, Milne-Simpson method (3.2) is an implicit two-step method of order 4. Provide an experimental confirmation of its order on the van der Pol problem (3.24).
4. Prove that a LMM (3.1) of order p exactly solves all the differential problems whose solution is a polynomial of degree at most p .
Hint: prove that, for methods of order p , the linear difference operator (3.10) annihilates on the set of functions $\{1, t, t^2, \dots, t^p\}$.
5. Construct an explicit two-step method (3.1) exactly solving all differential problems whose solution belongs to the functional space spanned by

$$\{1, t, \cos(\omega t), \sin(\omega t)\}.$$

Then, compute the limit as ωh tends to 0, being h the stepsize, of the coefficients of the obtained method and check if they fulfill the set of order conditions described in Theorem 3.2 up to a certain integer p .

6. As regards Example 3.9, find the values of β_2 such that third order methods are zero-stable.
7. Write a code in your favorite programming language implementing the three-step method

$$y_{n+3} - \frac{18}{11}y_{n+2} + \frac{9}{11}y_{n+1} - \frac{2}{11}y_n = \frac{6}{11}hf(y_{n+3}, t_{n+3})$$

for the numerical solution of d -dimensional systems (1.1). You need to recover the two-missing starting values y_1 and y_2 , hence you can implement two options: recover both y_1 and y_2 by two consecutive steps of a one-step method; recover y_1 by a one-step method and then y_2 by a two-step method. Provide an experimental evidence of the convergence of the method and of its order.

8. The bound (3.7) estimating the stepsize restriction for the convergence of fixed point iterations in the case implicit LMMs (3.1) is fully computable if the value of the Lipschitz constant L of the vector field is known. Estimating the Lipschitz constant of a function is very important in many fields: for instance, several

estimation algorithms for the Lipschitz constant have been developed in papers on global optimization.

The exercise consists in two parts:

- write a code in your favorite programming language implementing the following estimation algorithm for the Lipschitz constant L of a given scalar function [346].

Step 1. Compute the approximation P solutions of a scalar initial value problem (1.17) in autonomous form, by a chosen LMM (3.1), corresponding to P different initial values. Denote by $y_n^{i,j}$ the i -th component of the j -th solution in the point $t_n \in \mathcal{I}_h$, $i = 1, 2, \dots, d$, $j = 1, 2, \dots, P$. Then, define

$$a_i = \min_{j=1, \dots, P} \min_{t_n \in \mathcal{I}_h} X_n^{i,j}, \quad b_i = \max_{j=1, \dots, P} \max_{t_n \in \mathcal{I}_h} X_n^{i,j},$$

$i = 1, 2, \dots, d$.

Step 2. Generate Q couples of vectors

$$x_k = [x_k^1, x_k^2, \dots, x_k^d]^T, \quad y_k = [y_k^1, y_k^2, \dots, y_k^d]^T,$$

with $k = 1, 2, \dots, Q$, such that (x_k^i, y_k^i) is uniformly distributed in $[a_i, b_i] \times [a_i, b_i]$, $i = 1, 2, \dots, d$.

Step 3. Compute

$$s_k = \frac{|f(x_k) - f(y_k)|^2}{|x_k - y_k|^2}, \quad k = 1, 2, \dots, Q.$$

Step 4. Assume as estimate of L the value of $\max\{s_1, \dots, s_Q\}$;

- solve the same differential problem by using the implicit Euler method (2.32) and give an experimental confirmation of the sharpness of the bound (3.7), by choosing values of the stepsize above and below this bound (fully computable using the aforementioned estimation algorithm for the Lipschitz constant of the vector field) and checking the convergence of fixed point iterations in correspondence of the chosen values of the stepsize.
9. Using Definition 3.3 of consistency, provide a consistency analysis of Milne-Simpson (3.2) computing its local truncation error and, consequently, infer its order.
 10. Develop maximal order convergent LMMs depending on 4 and 5 steps.