Milan Práger; Jiří Taufer; Emil Vitásek
Overimplicit methods for the solution of evolution problems

Terms of use:

Overimplicit Methods for the Solution of Evolution Problems

M. PRÁGER, J. TAUFER, E. VITÁSEK
Mathematical Institute, Czechoslovak Academy of Sciences, Prague

I. Introduction

It is known that in deriving the numerical methods for the approximate solution of many technical problems leading to systems of ordinary differential equations it is very often important not to aim only at the convergence or the high asymptotic accuracy but also to satisfy other requirements. Typical example of such problems are so-called stiff problems where we are interested in the component of the vector of the solution containing a strongly stifled term which is, consequently, negligible but which must be, when using an unconvenient method, still approximated. The similar problems arise in the solution of partial differential equations of parabolic type. This fact can be most easily comprehended assuming that the parabolic equation is solved by transformation on the system of ordinary differential equations (discretizing only the space variables) and observing that the resulting system of ordinary differential equations is the more stiff the more fine the mesh-size is. One of such above mentioned requirements is the Dahlquist's A-stability which has often proved very reasonable. It is well known, however, that in the class of basic methods for the numerical solution of ordinary differential equations (linear multistep methods, Runge-Kutta methods) there do not exist A-stable methods of order higher than 2. The aim of this paper is therefore to introduce a larger class of methods for the solution of ordinary differential equations, a class which will contain A-stable methods of arbitrarily high orders. Since it is also known that A-stable linear multistep methods are necessarily implicit, the implicit character of our methods will be emphasized in such a way that instead of computing the approximate solution in one point from the known approximate solution in l preceding points (as it is in the case of linear l-step method) we will compute the approximate solution in k successive points simultaneously from some system of equations, supposing that the solution is known in l successive points. From this reason our methods will be called *overimplicit* methods.

2. The Definition of an Overimplicit Multistep Method

For the sake of simplicity we will treat only one differential equation of the first order

$$y' = f(x, y) \quad \text{in} \quad \langle a, b \rangle \tag{2.1}$$

with the initial condition

$$y(a) = \eta . \tag{2.2}$$

The generalization for systems of first order differential equations will be obvious in what follows. Also for the simplicity, we will assume that the points x_i in which the approximate solution is sought are equidistant, i.e., $x_i = a + ih, i = 0, 1, \ldots$ and $h > 0$ is constant. The right-hand term of the given differential equation is assumed to be defined, continuous and satisfying the Lipschitz condition with respect to y in the strip $a \leqq x \leqq b$, $-\infty < y < \infty$ so that the solution of the problem (2.1), (2.2) exists and is unique in the whole interval $\langle a, b \rangle$. If we denote the approximate solution in the point x_i by y_i one step of our method consists — as it was already mentioned — in computing the values $y_{n+1}, \ldots, y_{n+k}$ of the approximate solution in the points $x_{n+1}, \ldots, x_{n+k}$ (assuming $y_{n-l+1}, \ldots, y_n$ to be known) simultaneously from the system

$$\begin{bmatrix} y_{n+1} \\ \vdots \\ y_{n+k} \end{bmatrix} + \mathbf{B} \begin{bmatrix} y_{n-l+1} \\ \vdots \\ y_n \end{bmatrix} = h\mathbf{C} \begin{bmatrix} f_{n+1} \\ \vdots \\ f_{n+k} \end{bmatrix} + h\mathbf{D} \begin{bmatrix} f_{n-l+1} \\ \vdots \\ f_n \end{bmatrix} \tag{2.3}$$

where $f_i = f(x_i, y_i)$, $\mathbf{C}$ is a square matrix of order k, and $\mathbf{B}$, $\mathbf{D}$ are $k \times l$ matrices.

The fact that the function $f(x, y)$ satisfies the Lipschitz condition guarantees the existence and the uniqueness of the solution of (2.3) for any sufficiently small h so that one step of our method is well defined. In order to describe the whole method it is necessary, moreover, to indicate how to continue in the following step, i.e. how to choose l new initial values. The method will be practicable obviously only in that case when the new initial values will be chosen from the values $y_{n-l+2}, \ldots, y_{n+k}$. In order to specify them exactly let us take an integer s, $1 \leqq s \leqq k$ and the new initial values let be $y_{n-l+1+s}, \ldots, y_{n+s}$. Hence, the overimplicit method is characterized not only by the matrices $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ but also by the parameter s.

Let us note that if $s < k$ it is necessary to forget just computed values $y_{n+s+1}, \ldots, y_{n+k}$ and to recompute them in the following step. For the simple notation we will always denote the value of the approximate solution in the point x_i by only one symbol y_i even though this value need not be the same in different phases of the computation.

3. Convergence of Overimplicit Methods

Our first task is to indicate the conditions under which the method just described is convergent. In order to be able to formulate the basic theorems concerning this problem it is necessary to introduce some concepts and definitions.

Definition 3.1. *The overimplicit method given by the matrices* $\mathbf{B} = \{b_{ij}\}$, $\mathbf{C} = \{c_{ij}\}$, $\mathbf{D} = \{d_{ij}\}$ *and by the parameter* s *is said to be of order* p (p *positive integer) if the following conditions are satisfied:*

$$1 + \sum_{j=1}^{l} b_{ij} = 0 , \quad i - \sum_{j=1}^{l} b_{ij}(l-j) = \sum_{j=1}^{k} c_{ij} + \sum_{j=1}^{l} d_{ij} ,$$

$$i^{\nu} + (-1)^{\nu} \sum_{j=1}^{l} b_{ij}(l-j)^{\nu} = \nu \left[\sum_{j=1}^{k} c_{ij} j^{\nu-1} + (-1)^{\nu-1} \sum_{j=1}^{l} d_{ij}(l-j)^{\nu-1} \right] ,$$

$$\nu = 2, \ldots, p; \quad i = 1, \ldots, k. \tag{3.1}$$

Definition 3.2. *The method* (2.3) *is said to be consistent if it is of order at least one.*

Let us draw the attention to the fact that both the consistency and the order of the method are local properties of the method, i.e., they depend only on the matrices $\mathbf{B}, \mathbf{C}, \mathbf{D}$ and do not depend on s.

The conditions (3.1) express that the local error $\mathbf{L}(y(x); h)$ defined by

$$\mathbf{L}(y(x); h) = \begin{bmatrix} y(x+h) \\ \vdots \\ y(x+kh) \end{bmatrix} + \mathbf{B} \begin{bmatrix} y(x-(l-1)h) \\ \vdots \\ y(x) \end{bmatrix} - h\mathbf{C} \begin{bmatrix} y'(x+h) \\ \vdots \\ y'(x+kh) \end{bmatrix} - $$

$$- h\mathbf{D} \begin{bmatrix} y'(x-(l-1)h) \\ \vdots \\ y'(x) \end{bmatrix} \tag{3.2}$$

is of order h^{p+1}.

Because we are dealing with the multistep method it can be expected that the convergence will not be guaranteed only by the assumption that the local error is small but that some other conditions analogical to Dahlquist's stability conditions will have to be fulfilled. In order to be able to formulate them let us introduce further notations.

Let there be given an overimplicit method and let firstly $l \leq s$. Define the matrix $\mathbf{R}$ by

$$\mathbf{R} = [\mathbf{O}_{l, s-l}, \quad \mathbf{I}_l, \quad \mathbf{O}_{l, k-s}] \tag{3.3}$$

where $\mathbf{O}_{m,n}$ is $m \times n$ null matrix and $\mathbf{I}_l$ is the identity matrix of order l. Further, define the matrix $\mathbf{E}$ by

$$\mathbf{E} = -\mathbf{R}\mathbf{B} . \tag{3.4}$$

Secondly, let $l > s$; define now the matrix $\mathbf{S}$ by

$$\mathbf{S} = [\mathbf{I}_s, \quad \mathbf{O}_{s, k-s}] . \tag{3.5}$$

Further, let

$$i = \left[\frac{l-1}{s} \right] \tag{3.6}$$

and construct the matrix

$$B^{(1)} = [O_{k,\ (i+1)\ s-l}, \quad B]\ ;$$

(3.7)

let us divide $SB^{(1)}$ into $i+1$ square blocks in such a way that

$$SB^{(1)} = [B_0, \ldots, B_i]$$

(3.8)

and define

$$(3.9)$$

After having introduced the matrix E we are able to define the stability of an overimplicit method (2.3).

Definition 3.3. *The overimplicit method (2.3) is said to be stable if there exists a constant Γ such that for any n*

$$\|E^n\| \leq \Gamma.$$

(3.10)

Here E is defined by (3.4) or (3.9) for $l \leq s$ or $l > s$ respectively and $\|E\|$ is an arbitrary norm of E as the linear mapping in the corresponding vector space; for the definiteness let us consider the spectral norm.

Let us note that while the consistency depends on the matrices B, C, D and does not depend on s the stability depends only on B and s.

Now we have all prepared to be able to formulate the basic theorems concerning the convergence of an overimplicit method.

Theorem 3.1. *A stable and consistent overimplicit method is convergent.*

Let us note to this theorem that if the convergence is understood in the usual sense, i.e., if any approximate solution the initial conditions of which converge by $h \to 0$ to the initial condition of the given differential equation converges to the exact solution then the conditions of Theorem 3.1 are not only sufficient but also necessary.

Theorem 3.2. *Let the solution of the given differential equation have continuous derivatives up to order $p+1$. Then the approximate solution computed by an overimplicit stable method of order p the initial conditions of which are approximated with the accuracy h^p converges to the exact solution with the rate h^p.*

4. A-stability of Overimplicit Methods

Let us investigate now the A-stability of overimplicit methods. The main result concerning this problem can be formulated very simply.

Theorem 4.1. *In the class of overimplicit methods there exist A-stable methods of arbitrarily high orders.*

Let us indicate very briefly one of the possibilities how to prove *constructively* this theorem.

First of all, in order to facilitate our task, let us seek A-stable methods of arbitrarily high orders only in the subset of the class of overimplicit methods for which $l = 1$ and $\mathbf{B} = (-1, \ldots, -1)^T$ and which are of the order at least k. The methods of this subset will be called selfstarting overimplicit almost optimal methods from the reason that they need no starting procedure and that their order is at least k. Thus, in what follows we will deal only with the formulae

$$\begin{bmatrix} y_{n+1} \\ \vdots \\ y_{n+k} \end{bmatrix} = \begin{bmatrix} y_n \\ \vdots \\ y_n \end{bmatrix} + h\mathbf{C} \begin{bmatrix} f_{n+1} \\ \vdots \\ f_{n+k} \end{bmatrix} + hf_n\mathbf{d} \tag{4.1}$$

with $\mathbf{C}$ and $\mathbf{d}$ given such that

$$\begin{bmatrix} y(x + h) \\ \vdots \\ y(x + kh) \end{bmatrix} - \begin{bmatrix} y(x) \\ \vdots \\ y(x) \end{bmatrix} - h\mathbf{C} \begin{bmatrix} y'(x + h) \\ \vdots \\ y'(x + kh) \end{bmatrix} - hy'(x)\,\mathbf{d} = O(h^{k+1}) \tag{4.2}$$

holds for any sufficiently smooth function $y(x)$. Thus, any matrix $\mathbf{C}$ and any vector $\mathbf{d}$ satisfying (4.2) define a selfstarting overimplicit almost optimal method.

Let us note that this class is not empty. Any formula of the interpolation type

$$y_{n+i} = y_n + h \sum_{j=0}^{k} \gamma_{ij} f_{n+j}, \quad i = 1, \ldots, k \tag{4.3}$$

where

$$\gamma_{ij} = \int_0^i l_j(t)\,\mathrm{d}t \tag{4.4}$$

and $l_j(t)$ is the elementary Lagrange interpolating polynomial for the points $t = 0, \ldots, k$ is an element of our class since it is of order $k + 1$. The method (4.3) will be called the self-starting method of *Adams* type.

Let us return now to the general selfstarting almost optimal method (4.1) and let us investigate its A-stability. If we use this formula for solving the differential equation $y' = \alpha y$, where α is a (complex) constant, we get

$$(\mathbf{I} - z\mathbf{C}) \begin{bmatrix} y_{n+1} \\ \vdots \\ y_{n+k} \end{bmatrix} = y_n(\mathbf{e} + z\mathbf{d}) \tag{4.5}$$

where

$$\mathbf{e} = (1, \ldots, 1)^T \tag{4.6}$$

and

$$z = \alpha h . \tag{4.7}$$

129

Because only the value y_{n+s} is used as the initial value in the following step of the method we are interested only in the values y_{rs}, $r = 0, 1, ...$ in fact. For them we obtain immediately by Cramer's rule

$$y_{(r+1)s} = \frac{P_s(z)}{Q(z)} \ y_{rs}, \ r = 0, 1, ... \tag{4.8}$$

where

$$Q(z) = \det (I - zC) \tag{4.9}$$

and $P_s(z)$ is the determinant of the matrix which arises from the matrix $I - zC$ by replacing its s-th column by the vector $e + zd$. From (4.8) it follows now immediately that the necessary and sufficient condition for the A-stability of a self-starting formula (4.1) is that

$$\left| \frac{P_s(z)}{Q(z)} \right| < 1 \tag{4.10}$$

for any z with negative real part.

The construction of a selfstarting A-stable method is now based on the following statement.

Lemma 4.1. *Let* $\sum_{i=0}^{k} a_i z^i$ *be any polynomial with* $a_0 = 1$. *Let* t *be the vector defined by*

$$t = k! \sum_{i=0}^{k} a_{k-i} \frac{1}{(i+1)!} \ M^{i+1} e \tag{4.11}$$

where e *is defined by* (4.6) *and* M *is the diagonal matrix of order* k *having the numbers* $1, ..., k$ *on the diagonal. Further, let the matrix* C *be the solution of the equation*

$$M^2 V = CMV(I + M) + [\overbrace{0, \ ..., \ 0}^{(k-1) \ times}, \ (k+1) \ t] \tag{4.12}$$

where 0 *is the null* k-dimensional vector and V is' the Vandermond matrix for the numbers $1, ..., k$, i.e.,

$$V = \begin{bmatrix} 1^0 \ 1^1 \ ... \ 1^{k-1} \\ 2^0 \ 1^1 \ ... \ 2^{k-1} \\ \vdots \\ k^0 \ k^1 \ ... \ k^{k-1} \end{bmatrix} \tag{4.13}$$

Finally, let d *be the vector defined by*

$$d = Me - Ce \ . \tag{4.14}$$

Then C *and* d *define a selfstarting overimplicit almost optimal method for which*

$$Q(z) = \det (I - zC) = \sum_{i=0}^{k} a_i z^i \tag{4.15}$$

130

and

$$P_s(z) = \sum_{=0}^{k} \left(\sum_{i=0}^{j} \frac{s^{j-i}}{(j-i)!} \, a_i \right) z^j \quad (s = 1, ..., k) \tag{4.16}$$

With this lemma, the proof of Theorem 4.1 is now easy. It is sufficient to construct with help of this lemma such a selfstarting method that the corresponding polynomials $P_s(z)$ and $Q(z)$ have convenient properties. The simplest way is to take for the polynomial $Q(z)$ (by given s, $1 \leq s \leq k$) the polynomial $R(-sz)$, where $R(z)$ is such a polynomial that the ratio $R(z)/R(-z)$ is the Padé approximation of e^z. Then it is $P_s(z) = Q(-z)$ and Theorem 4.1 follows from the well-known fact that $R(z)$ has zeros in the left halfplane.

5. Applicability of *A*-stable Methods

In this last section of the paper we will investigate the applicability of A-stable selfstarting methods for the numerical solution of partial differential equations of parabolic type. Let us study firstly the simpler problem which will indicate the results which can be expected.

Let us solve by an A-stable selfstaring method the system of ordinary differential equations

$$\mathbf{y}' = \mathbf{A}\mathbf{y} \tag{5.1}$$

and let us suppose that $\mathbf{A}$ is a constant matrix with eigenvalues with negative real parts. The application of the method (4.1) on this system gives

$$\begin{bmatrix} \mathbf{y}_{n+1} \\ \vdots \\ \mathbf{y}_{n+k} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_n \\ \vdots \\ \mathbf{y}_n \end{bmatrix} + h(\mathbf{C} \otimes \mathbf{A}) \begin{bmatrix} \mathbf{y}_{n+1} \\ \vdots \\ \mathbf{y}_{n+k} \end{bmatrix} + h \begin{bmatrix} d_1 \mathbf{A} \\ \vdots \\ d_k \mathbf{A} \end{bmatrix} \mathbf{y}_n \tag{5.2}$$

where $\mathbf{C} \otimes \mathbf{A}$ is so-called tensor product of the matrices $\mathbf{C}$ and $\mathbf{A}$ defined by

$$\mathbf{C} \otimes \mathbf{A} = \begin{bmatrix} c_{11}\mathbf{A} \cdots\cdots\cdots c_{1k}\mathbf{A} \\ \vdots \\ c_{k1}\mathbf{A} \cdots\cdots\cdots c_{kk}\mathbf{A} \end{bmatrix} \tag{5.3}$$

It is known that the eigenvalues of $\mathbf{C} \otimes \mathbf{A}$ are the products of the eigenvalues of $\mathbf{C}$ and $\mathbf{A}$. Since the eigenvalues of $\mathbf{C}$ have positive real parts (as it follows from the A-stability of the method under consideration) and the eigenvalues of $\mathbf{A}$ have negative real parts the matrix $\mathbf{C} \otimes \mathbf{A}$ cannot have real positive eigenvalues and, consequently, $\mathbf{I} - h(\mathbf{C} \otimes \mathbf{A})$ is regular for any $h > 0$. From here it follows firstly that an A-stable selfstarting method applied on any differential equation of the form (5.1) has sense for any $h > 0$. Further, it can be easily shown that

$$\mathbf{y}_{(r+1)s} = Q^{-1}(h\mathbf{A}) \, P_s(h\mathbf{A}) \, \mathbf{y}_{rs} \tag{5.4}$$

where $P_s(z)$ and $Q(z)$ are polynomials defined in the preceeding section by (4.8).

If λ is now the eigenvalue of $\mathbf{A}$ then $P_s(h\lambda) / Q(h\lambda)$ is the eigenvalue of $Q^{-1}(h\mathbf{A}) P_s(h\mathbf{A})$ and, since our method is A-stable, this number is therefore smaller than 1 in magnitude. Thus, it holds

$$\| [Q^{-1}(h\mathbf{A}) P_s(h\mathbf{A})]^n \| \to 0 \quad \text{for} \quad n \to \infty \tag{5.5}$$

and, consequently,

$$\| \mathbf{y}_{rs} \| \to 0 \quad \text{for} \quad r \to \infty . \tag{5.6}$$

From (5.6) it follows that any A-stable selfstarting almost optimal method can be applied to the system (5.1) with constant coefficients.

This analysis suggests that the solution of partial differential equations of parabolic type with coefficients independent of time variable by selfstarting A-stable methods need not meet any substantial difficulties.

Thus, let there be given a differential equation

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left(p(x) \frac{\partial u}{\partial x} \right) - q(x)u + f(x) \tag{5.7}$$

with the initial condition

$$u(x, 0) = g(x) \tag{5.8}$$

and with the boundary conditions

$$\alpha_1 u(t, a) - \beta_1 p(a) \frac{\partial u(t, a)}{\partial x} = \gamma_1 ,$$

$$\alpha_2 u(t, b) + \beta_2 p(b) \frac{\partial u(t, b)}{\partial x} = \gamma_2$$

$$(p(x) \geqq p_0 > 0, \quad q(x) \geqq 0, \quad \alpha_i \geqq 0, \quad \beta_i \geqq 0, \quad \alpha_i + \beta_i > 0) . \tag{5.9}$$

Let us describe first of all what it is understood under the solution of (5.7) by a selfstarting method (4.1). Let $t_i = \tau i$ ($\tau > 0$ is a constant) be the points on the time axis and denote the approximate solution of (5.7) on the line $t = t_i$, $a < x < b$ by $u_i(x)$. The solution of (5.7) by a selfstarting method (4.1) then means the determination of $u_i(x)$ as a solution of the boundary value problem for the system of ordinary differential equations

$$\begin{bmatrix} u_{n+1}(x) \\ \vdots \\ u_{n+k}(x) \end{bmatrix} = \begin{bmatrix} u_n(x) \\ \vdots \\ u_n(x) \end{bmatrix} + \tau \mathbf{C} \begin{bmatrix} \mathscr{L}u_{n+1} + f \\ \vdots \\ \mathscr{L}u_{n+k} + f \end{bmatrix} + \tau \mathbf{d}(\mathscr{L}u_n + f) \tag{5.10}$$

where

$$\mathscr{L}u \equiv \frac{\partial}{\partial x} \left(p(x) \frac{\partial u}{\partial x} \right) - q(x) u . \tag{5.11}$$

Now, it can be really proved without any serious problems that

(i)   The system (5.10) has a solution for any $\tau > 0$.

(ii) $u_i(x) \to u(t_i, x)$ for $\tau \to 0$ and $i\tau = \text{const.}$

(iii)   The rate of convergence is (under the convenient smoothness conditions) at least τ^k.

In the case of time dependent coefficients, the situation is substantionally more complicated. The essence of the difficulties can be easy comprehended if we apply our methods to the solution of the system of ordinary differential equations (5.1) where the matrix $\boldsymbol{A}$ is variable. In the equation (5.2) we have then instead of $\boldsymbol{C} \otimes \boldsymbol{A}$ the matrix

$$\boldsymbol{\Phi}_n = \begin{bmatrix} c_{11}\boldsymbol{A}(x_{n+1}) & c_{12}\boldsymbol{A}(x_{n+2}) & \cdots & c_{1k}\boldsymbol{A}(x_{n+k}) \\ \vdots & \vdots & & \vdots \\ c_{k1}\boldsymbol{A}(x_{n+1}) & c_{k2}\boldsymbol{A}(x_{n+2}) & \cdots & c_{kk}\boldsymbol{A}(x_{n+k}) \end{bmatrix} . \tag{5.12}$$

In this case, $h > 0$ can exist for which $\boldsymbol{I} - h\boldsymbol{\Phi}_n$ is singular even if $\boldsymbol{C}$ corresponds to an A-stable method and $\boldsymbol{A}(x)$ has the eigenvalues with negative real parts for any x. Thus, the system corresponding to (5.2) need not have the solution for some h and so (5.6) cannot be satisfied. It can seem that this difficulties are inherent in the problem alone and not in the method used since in the case under investigation the system (5.1) need not be Ljapunov stable and the validity of (5.6) cannot be therefore expected. But in the case that $\boldsymbol{A}(x)$ is symetric (and has negative eigenvalues) the system is Ljapunov stable but nevertheless, it can be shown that the difficulties indicated above may occur.

From what it was said it seems that the A-stability alone is not sufficient to guarantee the convergence in the case of the parabolic equation with time dependent coefficients. About these problems we know very little; we are able only, on the basis of the overimplicit formulae of Adams type, to construct concrete formulae converging even in the case of time dependent problems.

References

[1] DAHLQUIST, G.: A Special Stability Problem for Linear Multistep Methods. BIT 3, (1963), 27–43.
[2] HENRICI, P.: Discrete Variable Methods in Ordinary Differential Equations. J. Wiley & Sons, Inc., New York–London (1962).
[3] PRÁGER, M., TAUFER, J., VITÁSEK, E.: Overimplicit Multistep Methods. Apl. Mat. 18, (1973), 399–421.