Aplikace matematiky

Hassan Nasr Ahmed Ismail

Generalized periodic overimplicit multistep methods (GPOM methods)

Aplikace matematiky, Vol. 24 (1979), No. 4, 250–272

Persistent URL: http://dml.cz/dmlcz/103805

GENERALIZED PERIODIC OVERIMPLICIT MULTISTEP METHODS

(GPOM METHODS)

HASSAN NASR AHMED ISMAIL

(Received August 27, 1977)

1. INTRODUCTION

In the course of development of numerical analysis, a number of numerical methods for solving initial-value problems for ordinary differential equations have been proposed. The first methods for this purposes were individual methods as, e.g., Adams methods, Runge-Kutta methods etc. Later, the synthesis of properties of these individual methods gave the origin to the theory of general one step methods and linear k-step methods.

This paper represents an attempt of further synthesis of properties of individual methods in introducing a class of methods which contains all important methods known. The advantage of our approach consists on the one hand in methodical reasons – we are able to prove, e.g., the convergence of a number of individual methods in one single proof – on the other hand, the presence of free parameters in a method allows us to construct individual methods having convenient properties for solving particular problems, as for example stiff systems of differential equations or evolution problems.

The new method introduced in the paper will be called "Generalized Periodic Overimplicit Multistep (GPOM) Method" especially for the reason that the main idea of this method consists in the fact that in one step of the method one computes not only one unknown value of the approximate solution (as in classical methods) but a group of unknown values of the approximate solution from a (generally non-linear) system of equations. A similar idea is studied by Práger, Taufer and Vitásek [1], but here, moreover. the distribution of the points at which the solution is sought is allowed to be general. This fact enlarges on the one hand very substantially the class of the methods under investigation, on the other hand, it simplifies the convergence proof.

For simplicity we shall formulate our method only for one differential equation

(1) $$y' = f(x, y), \quad x \in [a, b], \quad y(a) = \eta$$

where the function $f(x, y)$ is defined, continuous and satisfying the Lipschitz condition with respect to y in $[a, b] \times (-\infty, \infty)$ so that the existence and uniqueness of the solution of (1) is guaranteed in the whole $[a, b]$.

The approximate solution will be studied at k points simultaneously supposing that it is known at l points. Let us suppose that $1 \leqq l \leqq k$, the assumption which, as we shall see, is not restrictive.

To describe the structure of that k and l points we let the mesh size to be a positive number h and if m is a positive integer we define the basic point x_{jk}, $j = 0, 1, \ldots$ by

$$x_{jk} = a + jmh \quad j = 0, 1, \ldots$$

(the distance between two consecutive basic points is mh).

Also we define the intermediate points x_{jk+i}, $i = 1, \ldots, k - 1$ by

$$x_{jk+i} = x_{jk} + \mu_i h$$

where μ_i are real numbers.

The values $i = 0$, $i = k$ with $\mu_0 = 0$, $\mu_k = m$ show that these intermediate points will correspond to the basic points.

It is easy to prove the periodicity property:

$$x_{s+k} = x_s + mh \quad \text{for any integer } s\,.$$

Let us notice that

(i) for some $i < j$ it may happen that $\mu_i \geqq \mu_j$;
(ii) there may exist i for which $\mu_i > m$ and $\mu_i < 0$;
(iii) there may exist i not necessarily 0 or k for which $\mu_i = 0$ or m.

Let us introduce now some notation which will simplify the definition of one step of our method. Let $\boldsymbol{x}_j$ and $\tilde{\boldsymbol{y}}(\boldsymbol{x}_j)$, $j = 0, 1, \ldots$ be l-dimensional vectors defined by:

(2) $$\boldsymbol{x}_j = [x_{jk}, x_{jk+1}, \ldots, x_{jk+l-1}]^{\mathsf{T}} \quad \text{for} \quad j = 0, 1, \ldots$$

(3) $$\tilde{\boldsymbol{y}}(\boldsymbol{x}_j) = [\tilde{y}(x_{jk}), \ldots, \tilde{y}(x_{jk+l-1})]^{\mathsf{T}}$$

where $\tilde{y}(x_{jk+i})$ denote the approximate solution at the points x_{jk+i} for $j = 0, 1, \ldots$; $i = 0, \ldots, k - 1$. The last vector will serve as the input data for one step of our method.

Analogously, let us define the k-dimensional vector $\boldsymbol{z}_j$ by

(4) $$\boldsymbol{z}_j = [x_{jk+l}, x_{jk+l+1}, \ldots, x_{(j+1)k+l-1}]^{\mathsf{T}}$$

and the corresponding k-dimensional vector $\tilde{\boldsymbol{y}}(\boldsymbol{z}_j)$ of the approximate solution by

(5) $$\tilde{\boldsymbol{y}}(\boldsymbol{z}_j) = [\tilde{y}(x_{jk+l}), \ldots, \tilde{y}(x_{(j+1)k+l-1})]^{\mathsf{T}}\,.$$

This vector will serve as the output data for one step of our method. Further, let us define the l-dimensional and k-dimensional vector-valued functions $\boldsymbol{f}(\boldsymbol{x}_j, \tilde{\boldsymbol{y}}(\boldsymbol{x}_j))$ and $\boldsymbol{f}(\boldsymbol{z}_j, \tilde{\boldsymbol{y}}(\boldsymbol{z}_j))$, respectively, by

$$(6) \qquad \boldsymbol{f}(\boldsymbol{x}_j, \tilde{\boldsymbol{y}}(\boldsymbol{x}_j)) = [f(x_{jk}, \tilde{y}(x_{jk})), \ldots, f(x_{jk+l-1}, \tilde{y}(x_{jk+l-1}))]^{\mathrm{T}},$$

$$(7) \qquad \boldsymbol{f}(\boldsymbol{z}_j, \tilde{\boldsymbol{y}}(\boldsymbol{z}_j)) = [f(x_{jk+l}, \tilde{y}(x_{jk+l})), \ldots, f(x_{(j+1)k+l1}, \tilde{y}(x_{(j+1)k+l-1}))]^{\mathrm{T}}.$$

These two vectors represent the values of the right hand term of the differential equation (1) at the points (2), (3) and (4), (5), respectively.

Now we have prepared all to be able to define one step of our method: "Let a distribution of basic and intermediate points be given (i.e., the mesh size h, the integers k, l satisfying $l \leqq k$, the integer m and the constants μ_i). Further, let $k \times l$ matrices $\boldsymbol{B} = \{b_{ij}\}$, $\boldsymbol{D} = \{d_{ij}\}$ and a $k \times k$ matrix $\boldsymbol{C} = \{c_{ij}\}$ be given. Then the system

$$(8) \qquad \tilde{\boldsymbol{y}}(\boldsymbol{z}_j) = \boldsymbol{B}\,\tilde{\boldsymbol{y}}(\boldsymbol{x}_j) + h\boldsymbol{C}\boldsymbol{f}(\boldsymbol{z}_j, \tilde{\boldsymbol{y}}(\boldsymbol{z}_j)) + h\boldsymbol{D}\boldsymbol{f}(\boldsymbol{x}_j, \tilde{\boldsymbol{y}}(\boldsymbol{x}_j)), \quad j = 0, 1, \ldots$$

will be called the k-stage generalized periodic overimplicit multistep method, or, briefly, GPOM method".

It is necessary to add some remarks to this definition. The equation (8) is understood in such a way that it has to define the vector of the approximate solution $\tilde{\boldsymbol{y}}(\boldsymbol{z}_j)$ provided the vector $\tilde{\boldsymbol{y}}(\boldsymbol{x}_j)$ is known. Thus, we must first show that the vector $\tilde{\boldsymbol{y}}(\boldsymbol{z}_j)$ is really defined by this equation. Further, for practical computation, the formula (8) will be used repeatedly for $j = 0, 1, \ldots$. This process is obviously well-defined since we suppose that $l \leqq k$ so that the group of values $\tilde{y}(x_{jk}), \ldots, \tilde{y}(x_{jk+l-1})$ can be always selected from the just computed group. We must naturally suppose that l initial values are given at the beginning of the computation.

The Lipschitz property of the right hand term of the given differential equation implies the following theorem which justifies the definition just introduced.

Theorem 1. *Let the right hand term of the given differential equation* (1) *satisfy the Lipschitz condition with respect to y and let h be sufficiently small. Then there exists one and only one solution of* (8).

3. GPOM METHODS AS A GENERALIZATION OF CLASSICAL METHODS

In this section we show that the class of methods just introduced represents the natural generalization of classical methods.

3.1 Dahlquist's method

Let the approximate solution $\tilde{y}(t_j)$ (where $t_j = a + jh$) of (1) be computed from the equation

$$(9) \qquad \sum_{\nu=0}^{k} \alpha_\nu\, \tilde{y}(t_{n+\nu}) = h \sum_{\nu=0}^{k} \beta_\nu f(t_{n+\nu}, \tilde{y}(t_{n+\nu})) \quad n = 0, 1, \ldots$$

where α_ν, β_ν are constants, $\alpha_k \neq 0$, $|\alpha_0| + |\beta_0| > 0$ and the initial values $\tilde{y}(t_\mu)$, $\mu = 0, \dots, k - 1$ are known.

The connection between Dahlquist's method and GPOM method is described in the following theorem.

Theorem 2. *Let a k-stage GPOM method be given with $l = k$, $m = 1$, $\mu_i = i$ for $i = 1, \dots, k - 1$ and with matrices $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{D}$:*

$$\boldsymbol{B} = \begin{bmatrix} 0, & 1, & 0, & \dots, & 0, & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0, & 0, & 0, & \dots, & 1, & 0 \\ 0, & 0, & 0, & \dots, & 0, & 1 \\ -\dfrac{\alpha_0}{\alpha_k}, & -\dfrac{\alpha_1}{\alpha_k}, & -\dfrac{\alpha_2}{\alpha_k}, & \dots, & -\dfrac{\alpha_{k-2}}{\alpha_k}, & -\dfrac{\alpha_{k-1}}{\alpha_k} \end{bmatrix}, \quad \boldsymbol{C} = \frac{1}{\alpha_k}\begin{bmatrix} 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & \beta_k \end{bmatrix},$$

$$\boldsymbol{D} = \frac{1}{\alpha_k}\begin{bmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \\ \beta_0 & \dots & \beta_{k-1} \end{bmatrix}.$$

Then this method is equivalent to Dahlquist's method (9).

Proof. Since $\boldsymbol{z}_j = \boldsymbol{x}_{j+1}$ then (8) will be

$$\tilde{\boldsymbol{y}}(\boldsymbol{x}_{j+1}) = \boldsymbol{B}\,\tilde{\boldsymbol{y}}(\boldsymbol{x}_j) + h\boldsymbol{C}\boldsymbol{f}(\boldsymbol{x}_{j+1}, \tilde{\boldsymbol{y}}(\boldsymbol{x}_{j+1})) + h\boldsymbol{D}\boldsymbol{f}(\boldsymbol{x}_j, \tilde{\boldsymbol{y}}(\boldsymbol{x}_j))\,. \tag{10}$$

Here the first $k - 1$ equations are trivial in virtue of the special form of $\boldsymbol{B}$, $\boldsymbol{C}$ and $\boldsymbol{D}$ while the last equation of (10) can be written in the form

$$\tilde{y}(t_{j+k}) = -\frac{1}{\alpha_k}\sum_{\nu=0}^{k-1}\alpha_\nu\,\tilde{y}(t_{j+\nu}) + \frac{h\beta_k}{\alpha_k}f(t_{j+k}, \tilde{y}(t_{j+k})) + \frac{h}{\alpha_k}\sum_{\nu=0}^{k-1}\beta_\nu f(t_{j+\nu}, \tilde{y}(t_{j+\nu}))$$

since from the periodicity we have $x_{jk+\nu} = x_{jk} + \nu h = a + (j + \nu)\,h = t_{j+\nu}$. The theorem is proved.

3.2 Predictor-Corrector Methods

The approximate value $\tilde{y}^*(t_{j+k})$ is computed by the explicit formula:

$$\alpha_k^*\,\tilde{y}^*(t_{j+k}) + \sum_{\nu=0}^{k-1}\alpha_\nu^*\,\tilde{y}(t_{j+\nu}) = h\sum_{\nu=0}^{k-1}\beta_\nu^* f(t_{j+\nu}, \tilde{y}(t_{j+\nu})) \tag{11}$$

(assuming $\tilde{y}(t_j), \dots, \tilde{y}(t_{j+k-1})$ are known) where

$$t_j = \alpha + jh\,; \quad j = 0, 1, \dots\,.$$

The approximate solution is corrected by the implicit formula

$$(12) \qquad \sum_{\nu=0} \alpha_\nu \, \tilde{y}(t_{j+\nu}) = h\beta_k f(t_{j+k}, \tilde{y}^*(t_{j+k})) + h \sum_{\nu=0}^{k-1} \beta_\nu f(t_{j+\nu}, \tilde{y}(t_{j+\nu}))\,.$$

The connection between the predictor-corrector methods and GPOM method is described in the following theorem.

Theorem 3. *Let a $2k$-stage GPOM method be given with $l = k$, $m = 1$, $\mu_i = i$ for $i = 0, \ldots, k-1$; $\mu_i = i + 1 - k$ for $i = k, \ldots, 2k-1$ and with matrices*

$$\boldsymbol{B} = \begin{bmatrix} 0, & 0, & \ldots, & 0, & 0, & -\dfrac{\alpha_0^*}{\alpha_k^*}, & 0, & 0, & \ldots, & 0, & -\dfrac{\alpha_0}{\alpha_k} \\ 1, & 0, & \ldots, & 0, & 0, & -\dfrac{\alpha_1^*}{\alpha_k^*}, & 1, & 0, & \ldots, & 0, & -\dfrac{\alpha_1}{\alpha_k} \\ 0, & 1, & \ldots, & 0, & 0, & -\dfrac{\alpha_2^*}{\alpha_k^*}, & 0, & 1, & \ldots, & 0, & -\dfrac{\alpha_2}{\alpha_k} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0, & 0, & \ldots, & 1, & 0, & -\dfrac{\alpha_{k-2}^*}{\alpha_k^*}, & 0, & 0, & \ldots, & 0, & -\dfrac{\alpha_{k-2}}{\alpha_k} \\ 0, & 0, & \ldots, & 0, & 1, & -\dfrac{\alpha_{k-1}^*}{\alpha_k^*}, & 0, & 0, & \ldots, & 1, & -\dfrac{\alpha_{k-1}}{\alpha_k} \end{bmatrix}^{\mathrm{T}},$$

$$\boldsymbol{D} = \begin{bmatrix} 0 & \ldots & 0 & \dfrac{\beta_0^*}{\alpha_k^*} & 0 & \ldots & 0 & \dfrac{\beta_0}{\alpha_k} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & \ldots & 0 & \dfrac{\beta_{k-1}^*}{\alpha_k^*} & 0 & \ldots & 0 & \dfrac{\beta_{k-1}}{\alpha_k} \end{bmatrix}^{\mathrm{T}}$$

and $\boldsymbol{C} = \begin{bmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{C}_1 & \boldsymbol{O} \end{bmatrix}$

where $\boldsymbol{C}_1$ is the $k \times k$ matrix given by $\boldsymbol{C}_1 = \begin{bmatrix} 0 & \ldots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \ldots & 0 & 0 \\ 0 & \ldots & 0 & \dfrac{\beta_k}{\alpha_k} \end{bmatrix}$

Then this method is equivalent to the predictor-corrector method (11), (12).

Proof. The equation (8) can be written in the form

$$\begin{bmatrix} \tilde{y}(x_{2jk+k}) \\ \vdots \\ \tilde{y}(x_{2jk+2k-1}) \\ \tilde{y}(x_{2(j+1)k}) \\ \vdots \\ \tilde{y}(x_{2(j+1)k+k-1}) \end{bmatrix} = \boldsymbol{B} \begin{bmatrix} \tilde{y}(x_{2jk}) \\ \vdots \\ \tilde{y}(x_{2jk+k-1}) \end{bmatrix} +$$

$$+ h\boldsymbol{C} \begin{bmatrix} f(x_{2jk+k}, \tilde{y}(x_{2jk+k})) \\ \vdots \\ f(x_{2jk+2k-1}, \tilde{y}(x_{2jk+2k-1})) \\ f(x_{2(j+1)k}, \tilde{y}(x_{2(j+1)k})) \\ \vdots \\ f(x_{2(j+1)k+k-1}, \tilde{y}(x_{2(j+1)k+k-1})) \end{bmatrix} + h\boldsymbol{D} \begin{bmatrix} f(x_{2jk}, \tilde{y}(x_{2jk})) \\ \vdots \\ f(x_{2jk+k-1}, \tilde{y}(x_{2jk+k-1})) \end{bmatrix}$$

In virtue of

$$\begin{bmatrix} x_{2jk+k} \\ \vdots \\ x_{2jk+2k-1} \end{bmatrix} = \begin{bmatrix} x_{2(j+1)k} \\ \vdots \\ x_{2(j+1)k+k-1} \end{bmatrix} = \begin{bmatrix} t_{j+1} \\ \vdots \\ t_{j+k} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} x_{2jk} \\ \vdots \\ x_{2jk+k-1} \end{bmatrix} = \begin{bmatrix} t_j \\ \vdots \\ t_{j+k-1} \end{bmatrix}$$

the assertion of the theorem follows immediately.

3.3 **Runge-Kutta methods**

3.3.1 **Explicit Runge-Kutta Formulae**

In this method the approximation $\tilde{y}(t_{n+1})$ of the exact solution at t_{n+1} is computed from

$$\tilde{y}(t_{n+1}) = \tilde{y}(t_n) + h \sum_{v=1}^{k} w_v K_v \tag{13}$$

supposing $\tilde{y}(t_n)$ is known, and K_i are given by

$$\begin{cases} K_1 = f(t_n, \tilde{y}(t_n)) , \\ K_v = f(t_n + \alpha_v h, \tilde{y}(t_n) + h \sum_{s=1}^{v-1} \beta_{vs} K_s) , \quad v = 2, \ldots, k . \end{cases} \tag{14}$$

The connection between the Runge-Kutta methods and GPOM method is described in the following theorem.

Theorem 4. *Let a k-stage GPOM method be given with $l = 1$, $m = 1$, $\mu_\nu = \alpha_{\nu+1}$, $\nu = 1, \ldots, k - 1$ and with matrices*

$$B = \begin{bmatrix} 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} 0, & 0, & \ldots, & 0, & 0, & 0 \\ \beta_{32}, & 0, & \ldots, & 0, & 0, & 0 \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \beta_{k,2}, & \beta_{k,3}, & \ldots, & b_{k,k-1}, & 0, & 0 \\ w_2, & w_3, & \ldots, & w_{k-1}, & w_k, & 0 \end{bmatrix}, \quad D = \begin{bmatrix} \beta_{21} \\ \vdots \\ \beta_{k1} \\ w_1 \end{bmatrix}.$$

Then this method is equivalent to the Runge-Kutta method (13), (14).

Proof. Our GPOM for $\nu = 1, \ldots, k - 1$ can be written in the form:

$$\tilde{y}(x_{jk+\nu}) = \tilde{y}(x_{jk}) + h \sum_{s=1}^{\nu} \beta_{\nu+1,s} f(x_{jk+s-1}, \tilde{y}(x_{jk+s-1})),$$

and further we have

$$\tilde{y}(x_{(j+1)k}) = \tilde{y}(x_{jk}) + h \sum_{s=1}^{k} w_s f(x_{jk+s-1}, \tilde{y}(x_{jk+s-1})).$$

Putting $\tilde{K}_\nu = f(x_{jk+\nu-1}, \tilde{y}(x_{jk+\nu-1}))$ for $\nu = 1, \ldots, k$ we see that

$$\tilde{y}(x_{(j+1)k}) = \tilde{y}(x_{jk}) + h \sum_{s=1}^{k} w_s \tilde{K}_s \tag{15}$$

where

$$\tilde{K}_1 = f(x_{jk}, \tilde{y}(x_{jk})),$$

$$\tilde{K}_\nu = f(x_{jk+\nu-1}, \tilde{y}(x_{jk}) + h \sum_{s=1}^{\nu-1} \beta_{\nu s} \tilde{K}_s) \quad \nu = 2, \ldots, k. \tag{16}$$

Noticing that $x_{jk} = a + jh = t_j$, $x_{jk+\nu-1} = x_{jk} + \mu_{\nu-1}h = t_j + \alpha_\nu h$, $\nu = 2, \ldots, k$ we see that the equations (15) and (16) are exactly the same as the equations (13), (14).

3.3.2 **Implicit Runge-Kutta Formulae**

In this method (see, e.g., [3]) the approximation $\tilde{y}(t_{n+1})$ of the exact solution at t_{n+1} is computed from

$$\tilde{y}(t_{n+1}) = \tilde{y}(t_n) + h \sum_{\nu=1}^{k} w_\nu K_\nu \tag{17}$$

supposing that $\tilde{y}(t_n)$ is known, and K_ν are given by

$$K_\nu = f(t_n + \alpha_\nu h, \tilde{y}(t_n) + h \sum_{s=1}^{k} \beta_{\nu s} K_s), \quad \nu = 1, \ldots, k. \tag{18}$$

Theorem 5. *Let a $(k+1)$-stage GPOM method be given with $l = m = 1$, $\mu_\nu = \alpha_\nu$ for $\nu = 1, \ldots, k$ and with*

$$B = \begin{bmatrix} 1 \\ \cdot \\ \cdot \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} \beta_{11} & \ldots & \beta_{1k} & 0 \\ \vdots & & \vdots & \vdots \\ \beta_{k1} & \ldots & \beta_{kk} & 0 \\ w_1 & \ldots & w_k & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}.$$

Then this method is equivalent to the implicit Runge-Kutta method (17), (18).

Proof. Analogously to the proof of Theorem 4 we can write (8) in the form:

$$(19) \qquad \tilde{y}(x_{j(k+1)+\nu}) = \tilde{y}(x_{j(k+1)}) + h \sum_{s=1}^{k} \beta_{\nu s} f(x_{j(k+1)+s}, \tilde{y}(x_{j(k+1)+s})), \quad \nu = 1, \ldots, k;$$

$$(20) \qquad \tilde{y}(x_{(j+1)(k+1)}) = \tilde{y}(x_{j(k+1)}) + h \sum_{s=1}^{k} w_s f(x_{j(k+1)+s}, \tilde{y}(x_{j(k+1)+s})).$$

Let us introduce for $\nu = 1, \ldots, k$ the quantities

$$\tilde{K}_\nu = f(x_{j(k+1)+\nu}, \tilde{y}(x_{j(k+1)+\nu})).$$

From (19) it is clear that $\tilde{K}_\nu$ satisfy the system

$$\tilde{K}_\nu = f(x_{j(k+1)+\nu}, \tilde{y}_{j(k+1)} + h \sum_{s=1}^{k} \beta_{\nu s} \tilde{K}_s), \quad \nu = 1, \ldots, k.$$

The assertion of the theorem now follows from the fact that

$$x_{j(k+1)} = t_j, \; x_{j(k+1)+\nu} = t_j + \alpha_\nu h \quad \text{for} \quad \nu = 1, \ldots, k.$$

3.4 Overimplicit Multistep Methods

(see Práger, Taufer and Vitásek [1])

The approximate solution is sought at equidistant points t_n. In one step of the method we compute – as in our method – the approximate values $\tilde{y}(t_{n+1}), \ldots, \tilde{y}(t_{n+k})$ simultaneously from the system

$$(21) \qquad \begin{bmatrix} \tilde{y}(t_{n+1}) \\ \vdots \\ \tilde{y}(t_{n+k}) \end{bmatrix} = B \begin{bmatrix} \tilde{y}(t_{n-l+1}) \\ \vdots \\ \tilde{y}(t_n) \end{bmatrix} + hC \begin{bmatrix} f(t_{n+1}, \tilde{y}(t_{n+1})) \\ \vdots \\ f(t_{n+k}, \tilde{y}(t_{n+k})) \end{bmatrix} + hD \begin{bmatrix} f(t_{n-l+1}, \tilde{y}(t_{n-l+1})) \\ \vdots \\ f(t_n, \tilde{y}(t_n)) \end{bmatrix}$$

supposing $\tilde{y}(t_{n-l+1}), \ldots, \tilde{y}(t_n)$ are known. Given an integer s, $1 \leqq s \leqq k$ the next step of the method starts with the values $\tilde{y}(t_{n+s-l+1}), \ldots, \tilde{y}(t_{n+s})$ so that (21) is understood to be used for $n = js + l - 1, j = 0, 1, \ldots$. Let us underline that *no* relation is supposed between l and s so that among the numbers $\tilde{y}(t_{n+s-l+1}), \ldots, \tilde{y}(t_{n+s})$ there may occur not only approximate values computed in the preceding step of the method but also values computed in former steps. For this reason we will investigate the cases $s \geqq l$ and $s < l$ separately.

3.4.1 Case $s \geqq l$

Theorem 6. *Let the overimplicit multistep method (21) with $s \geqq l$ be given and let us construct the k-stage GPOM method using l starting points with $m = s$, $\mu_\nu = \nu$ for $\nu = 1, \ldots, s - 1$; $\mu_\nu = \nu + l$ for $\nu = s, \ldots, k - 1$ and with $\boldsymbol{B}_1 = \boldsymbol{PB}$, $\boldsymbol{C}_1 = \boldsymbol{PCP}^{-1}$, $\boldsymbol{D}_1 = \boldsymbol{PD}$*

$$\textit{where} \quad \boldsymbol{P} = \begin{bmatrix} \boldsymbol{I}_{s-l} & \boldsymbol{O}_{s-l,l} & \boldsymbol{O}_{s-l,k-s} \\ \boldsymbol{O}_{k-s,s-l} & \boldsymbol{O}_{k-s,l} & \boldsymbol{I}_{k-s} \\ \boldsymbol{O}_{l,s-l} & \boldsymbol{I}_l & \boldsymbol{O}_{l,k-s} \end{bmatrix}.$$

Then this method is equivalent to the given overimplicit multistep method.

Proof. Let us mention first that from $s \geqq l$ and $s \leqq k$ we have $l \leqq k$ so that the GPOM method from the theorem is really well defined. The formula (8) written in detail is

$$\begin{bmatrix} \tilde{y}(x_{jk+l}) \\ \vdots \\ \tilde{y}(x_{jk+s-1}) \\ \tilde{y}(x_{jk+s}) \\ \vdots \\ \tilde{y}(x_{jk+k-1}) \\ \tilde{y}(x_{(j+1)k}) \\ \vdots \\ \tilde{y}(x_{(j+1)k+l-1}) \end{bmatrix} = \boldsymbol{PB} \begin{bmatrix} \tilde{y}(x_{jk}) \\ \vdots \\ \tilde{y}(x_{jk+l-1}) \end{bmatrix} +$$

$$+ h\boldsymbol{PCP}^{-1} \begin{bmatrix} f(x_{jk+l}, y(x_{jk+l})) \\ \vdots \\ f(x_{jk+s-1}, \tilde{y}(x_{jk+s-1})) \\ f(x_{jk+s}, \tilde{y}(x_{jk+s})) \\ \vdots \\ f(x_{jk+k-1}, \tilde{y}(x_{jk+k-1})) \\ f(x_{(j+1)k}, \tilde{y}(x_{(j+1)k})) \\ \vdots \\ f(x_{(j+1)k+l-1}, \tilde{y}(x_{(j+1)k+l-1})) \end{bmatrix} + h\boldsymbol{PD} \begin{bmatrix} f(x_{jk}, \tilde{y}(x_{jk})) \\ \vdots \\ f(x_{jk+l-1}, \tilde{y}(x_{jk+l-1})) \end{bmatrix}.$$

Premultiplying this equation by $\boldsymbol{P}^{-1}$ and taking into account that

$$\boldsymbol{P}^{-1} = \begin{bmatrix} \boldsymbol{I}_{s-l} & \boldsymbol{O}_{s-l,k-s} & \boldsymbol{O}_{s-l,l} \\ \boldsymbol{O}_{l,s-l} & \boldsymbol{O}_{l,k-s} & \boldsymbol{I}_l \\ \boldsymbol{O}_{k-s,s-l} & \boldsymbol{I}_{k-s} & \boldsymbol{O}_{k-s,l} \end{bmatrix}$$

we have

$$\begin{bmatrix} \tilde{y}(x_{jk+l}) \\ \vdots \\ \tilde{y}(x_{jk+s-1}) \\ \tilde{y}(x_{(j+1)k}) \\ \vdots \\ \tilde{y}(x_{(j+1)k+l-1}) \\ \tilde{y}(x_{jk+s}) \\ \vdots \\ \tilde{y}(x_{jk+k-1}) \end{bmatrix} = \boldsymbol{B} \begin{bmatrix} \tilde{y}(x_{jk}) \\ \vdots \\ \tilde{y}(x_{jk+l-1}) \end{bmatrix} + \tag{22}$$

$$+ h\boldsymbol{C} \begin{bmatrix} f(x_{jk+l}, \tilde{y}(x_{jk+l})) \\ \vdots \\ f(x_{jk+s-1}, \tilde{y}(x_{jk+s-1})) \\ f(x_{(j+1)k}, \tilde{y}(x_{(j+1)k})) \\ \vdots \\ f(x_{(j+1)k+l-1}, \tilde{y}(x_{(j+1)k+l-1})) \\ f(x_{jk+s}, \tilde{y}(x_{jk+s})) \\ \vdots \\ f(x_{jk+k-1}, \tilde{y}(x_{jk+k-1})) \end{bmatrix} + h\boldsymbol{D} \begin{bmatrix} f(x_{jk}, \tilde{y}(x_{jk})) \\ \vdots \\ f(x_{jk+l-1}, \tilde{y}(x_{jk+l-1})) \end{bmatrix}.$$

Using the identities

$$x_{jk+\nu} = x_{jk} + \mu_\nu h = a + jsh + \mu_\nu h = t_{js+\nu}\,, \quad \nu = 0, \ldots, s - 1$$
$$x_{jk+\nu} = t_{js+l+\nu}\,, \quad \nu = s, \ldots, k - 1$$

which follow directly from the definition of basic and intermediate points we can rewrite (22) in the form

$$\begin{bmatrix} \tilde{y}(t_{js+l}) \\ \vdots \\ \tilde{y}(t_{js+s-1}) \\ \tilde{y}(t_{js+s}) \\ \vdots \\ \tilde{y}(t_{js+s+l-1}) \\ \tilde{y}(t_{js+s+l}) \\ \vdots \\ \tilde{y}(t_{js+l-1+k}) \end{bmatrix} = \boldsymbol{B} \begin{bmatrix} \tilde{y}(t_{js}) \\ \vdots \\ \tilde{y}(t_{js+l-1}) \end{bmatrix} +$$

$$
+ hC \begin{bmatrix} f(t_{js+l}, \tilde{y}(t_{js+l})) \\ \vdots \\ f(t_{js+s-1}, \tilde{y}(t_{js+s-1})) \\ f(t_{js+s}, \tilde{y}(t_{js+s})) \\ \vdots \\ f(t_{js+s+l-1}, \tilde{y}(t_{js+s+l-1})) \\ f(t_{js+s+l}, \tilde{y}(t_{js+s+l})) \\ \vdots \\ f(t_{js+l+k-1}, \tilde{y}(t_{js+l+k-1})) \end{bmatrix} + hD \begin{bmatrix} f(t_{js}, \tilde{y}(t_{js})) \\ \vdots \\ f(t_{js+l-1}, \tilde{y}(t_{js+l-1})) \end{bmatrix},
$$

which is exactly the equation (21) with $n = js + l - 1$. The theorem is proved.

3.4.2 Case $s < l$

Theorem 7. *Let the overimplicit multistep method* (21) *with $s < l$ be given and let us construct $(k + l - s)$-stage GPOM method using l starting points with $m = s$, $\mu_\nu = \nu$ for $\nu = 1, \ldots, l - 1$, $\mu_\nu = \nu + s$ for $\nu = l, \ldots, k + l - s - 1$ and with $\boldsymbol{B}_1 = \boldsymbol{P}\boldsymbol{B}_0$, $\boldsymbol{C}_1 = \boldsymbol{P}\boldsymbol{C}_0\boldsymbol{P}^{-1}$, $\boldsymbol{D}_1 = \boldsymbol{P}\boldsymbol{D}_0$*

$$
\textit{where} \quad \boldsymbol{P} = \begin{bmatrix} \boldsymbol{O}_{k-s,l} & \boldsymbol{I}_{k-s} \\ \boldsymbol{I}_l & \boldsymbol{O}_{l,k-s} \end{bmatrix}, \quad \boldsymbol{B}_0 = \begin{bmatrix} \boldsymbol{O}_{l-s,s} & \boldsymbol{I}_{l-s} \\ & \boldsymbol{B} \end{bmatrix},
$$

$$
\boldsymbol{C}_0 = \begin{bmatrix} \boldsymbol{O}_{l-s} & \boldsymbol{O}_{l-s,k} \\ \boldsymbol{O}_{k,l-s} & \boldsymbol{C} \end{bmatrix}, \quad \boldsymbol{D}_0 = \begin{bmatrix} \boldsymbol{O}_{l-s,l} \\ \boldsymbol{D} \end{bmatrix}.
$$

Then this method is equivalent to the given overimplicit multistep method.

Proof. The equation (8) can be written in more detail as

$$
\begin{bmatrix} \tilde{y}(x_{j(k+l-s)+l}) \\ \vdots \\ \tilde{y}(x_{j(k+l-s)+k+l-s-1}) \\ \tilde{y}(x_{(j+1)(k+l-s)}) \\ \vdots \\ \tilde{y}(x_{(j+1)(k+l-s)+l-1}) \end{bmatrix} = \boldsymbol{P}\boldsymbol{B}_0 \begin{bmatrix} \tilde{y}(x_{j(k+l-s)}) \\ \vdots \\ \tilde{y}(x_{j(k+l-s)+l-1}) \end{bmatrix} +
$$

$$
+ h\boldsymbol{P}\boldsymbol{C}_0\boldsymbol{P}^{-1} \begin{bmatrix} f(x_{j(k+l-s)+l}, \tilde{y}(x_{j(k+l-s)+l})) \\ \vdots \\ f(x_{j(k+l-s)+k+l-s-1}, \tilde{y}(x_{j(k+l-s)+k+l-s-1})) \\ f(x_{(j+1)(k+l-s)}, \tilde{y}(x_{(j+1)(k+l-s)})) \\ \vdots \\ f(x_{(j+1)(k+l-s)+l-1}, \tilde{y}(x_{(j+1)(k+l-s)+l-1})) \end{bmatrix}
$$

$$
+ h\boldsymbol{P}\boldsymbol{D}_0 \begin{bmatrix} f(x_{j(k+l-s)}, \tilde{y}(x_{j(k+l-s)})) \\ \vdots \\ f(x_{j(k+l-s)+l-1}, \tilde{y}(x_{j(k+l-s)+l-1})) \end{bmatrix}
$$

or, after multiplication by $\boldsymbol{P}^{-1}$ as

$$
\begin{bmatrix}
\tilde{y}(x_{(j+1)(k+l-s)}) \\
\vdots \\
\tilde{y}(x_{(j+1)(k+l-s)+l-s-1}) \\
\tilde{y}(x_{(j+1)(k+l-s)+l-s}) \\
\vdots \\
\tilde{y}(x_{(j+1)(k+l-s)+l-1}) \\
\tilde{y}(x_{j(k+l-s)+l}) \\
\vdots \\
\tilde{y}(x_{j(k+l-s)+k+l-s-1})
\end{bmatrix}
= \boldsymbol{B}_0
\begin{bmatrix}
\tilde{y}(x_{j(k+l-s)}) \\
\vdots \\
\tilde{y}(x_{j(k+l-s)+l-1})
\end{bmatrix}
+
$$

$$
+ h\boldsymbol{C}_0
\begin{bmatrix}
f(x_{(j+1)(k+l-s)}, \tilde{y}(x_{(j+1)(k+l-s)})) \\
\vdots \\
f(x_{(j+1)(k+l-s)+l-s-1}, \tilde{y}(x_{(j+1)(k+1-s)+l-s-1})) \\
f(x_{(j+1)(k+l-s)+l-s}, \tilde{y}(x_{(j+1)(k+l-s)+l-s})) \\
\vdots \\
f(x_{(j+1)(k+l-s)+l-1}, \tilde{y}(x_{(j+1)(k+l-s)+l-1})) \\
f(x_{j(k+l-s)+l}, \tilde{y}(x_{j(k+l-s)+l})) \\
\vdots \\
f(x_{j(k+l-s)+k+l-s-1}, \tilde{y}(x_{j(k+l-s)+k+l-s-1}))
\end{bmatrix}
$$

$$
+ h\boldsymbol{D}_0
\begin{bmatrix}
f(x_{j(k+l-s)}, \tilde{y}(x_{j(k+l-s)})) \\
\vdots \\
f(x_{j(k+l-s)+l-1}, \tilde{y}(x_{j(k+l-s)+l-1}))
\end{bmatrix}. \tag{23}
$$

By the definition of basic and intermediate points we get the following identities:

$$
\begin{aligned}
& x_{(j+1)(k+l-s)} = a + (j+1)\,sh = t_{js+s}\,, \\
& \vdots \\
& x_{(j+1)(k+l-s)+l-s-1} = t_{js+s} + \mu_{l-s-1}h = t_{js+s} + (l-s-1)\,h = t_{js+l-1}\,, \\
& x_{(j+1)(k+l-s)+l-s} = t_{js+s} + \mu_{l-s}h = t_{js+s} + (l-s)\,h = t_{js+l}\,, \\
& \vdots \\
& x_{(j+1)(k+l-s)+l-1} = t_{js+s} + \mu_{l-1}h = t_{js+s} + (l-1)\,h = t_{js+s+l-1}\,, \\
& x_{j(k+l-s)+l} = a + jsh + \mu_l h = t_{js} + (l+s)\,h = t_{js+s+l}\,, \\
& \vdots \\
& x_{j(k+l-s)+k+l-s-1} = t_{js} + \mu_{k+l-s-1}h = t_{js+k+l-1}
\end{aligned}
$$

and

$$
\begin{aligned}
& x_{j(k+l-s)} = t_{js}, \\
& \vdots \\
& x_{j(k+l-s)+s-1} = t_{js} + \mu_{s-1}h = t_{js+s-1}, \\
& x_{j(k+l-s)+s} = t_{js} + \mu_s h = t_{js+s}, \\
& \vdots \\
& x_{j(k+l-s)+l-1} = t_{js} + \mu_{l-1}h = t_{js+l-1}.
\end{aligned}
$$

Thus the first $l - s$ equations in (23) are identities in virtue of the special form of the matrices $\boldsymbol{B}_0$, $\boldsymbol{C}_0$ and $\boldsymbol{D}_0$ and the last equations can be written in the form

$$(24)\qquad \begin{bmatrix} \tilde{y}(t_{js+l}) \\ \vdots \\ \tilde{y}(t_{js+k+l-1}) \end{bmatrix} = \boldsymbol{B} \begin{bmatrix} \tilde{y}(t_{js}) \\ \vdots \\ \tilde{y}(t_{js+l-1}) \end{bmatrix} + h\boldsymbol{C} \begin{bmatrix} f(t_{js+l}, \tilde{y}(t_{js+l})) \\ \vdots \\ f(t_{js+l+k-1}, \tilde{y}(t_{js+l+k-1})) \end{bmatrix}$$
$$+ h\boldsymbol{D} \begin{bmatrix} f(t_{js}, \tilde{y}(t_{js})) \\ \vdots \\ f(t_{js+l-1}, \tilde{y}(t_{js+l-1})) \end{bmatrix}$$

which is exactly the equation (21) with $n = js + l - 1$. The theorem is proved.

Let us note that Theorem 7 proves once more Theorem 2 since Dahlquist's method is obviously a special case of the overimplicit method.

4. CONVERGENCE AND RATE OF CONVERGENCE OF GPOM METHOD

Two features will play an essential role in the convergence proof: First, it must be possible to make an error in one step of the method small in a convenient sense and, secondly, the method must be stable in a suitable sense, since one deals with multistep method. Before formulating the convergence theorems we must, first of all, formulate the above mentioned concepts.

4.1 Order and Stability of GPOM Method

Let us first of all define the local truncation error. Let the GPOM method and a function $y \in C^1$ be given. The k-dimensional vector

$$(25)\qquad \boldsymbol{L}(y(x); h) = \begin{bmatrix} y(x + \mu_l h) \\ \vdots \\ y(x + \mu_{k-1} h) \\ y(x + (m + \mu_0) h) \\ \vdots \\ y(x + (m + \mu_{l-1}) h) \end{bmatrix} - h\boldsymbol{B} \begin{bmatrix} y(x + \mu_0 h \\ \vdots \\ y(x + \mu_{l-1} h) \end{bmatrix} -$$
$$- h\boldsymbol{C} \begin{bmatrix} y'(x + \mu_l h) \\ \vdots \\ y'(x + \mu_{k-1} h) \\ y'(x + (m + \mu_0) h) \\ \vdots \\ y'(x + (m + \mu_{l-1}) h \end{bmatrix} - h\boldsymbol{D} \begin{bmatrix} y'(x + \mu_0 h) \\ \vdots \\ y'(x + \mu_{l-1} h) \end{bmatrix}$$

will be called the local truncation error of the method.

Remark 1. Putting

$$\begin{aligned} w_i &= \mu_{l-1+i} && \text{for} \quad i = 1, \ldots, k - l ; \\ &= m + \mu_{l-k+i-1} && \text{for} \quad i = k - l + 1, \ldots, k , \end{aligned}$$

the local truncation error will assume the form

$$\text{(26)} \qquad \boldsymbol{L}(y(x); h) = \begin{bmatrix} y(x + w_1 h) \\ \vdots \\ y(x + w_k h) \end{bmatrix} - \boldsymbol{B} \begin{bmatrix} y(x + (w_{k-l+1} - m) h) \\ \vdots \\ y(x + (w_k - m) h \end{bmatrix} -$$

$$- h\boldsymbol{C} \begin{bmatrix} y'(x + w_1 h) \\ \vdots \\ y'(x + w_k h) \end{bmatrix} - h\boldsymbol{D} \begin{bmatrix} y'(x + (w_{k-l+1} - m) h) \\ \vdots \\ y'(x + (w_k - m) h) \end{bmatrix},$$

which may be sometimes useful.

Supposing $y(x)$ is sufficiently smooth we can expand any component of $\boldsymbol{L}$ in the Taylor expansion. After rearranging the terms according to the powers of h we get

$$\text{(27)} \qquad L_i(y(x); h) = (1 - \sum_{j=1}^{l} b_{ij}) \, y(x) +$$

$$+ [w_i - \sum_{j=1}^{l} b_{ij}(w_{k-l+j} - m) - \sum_{j=1}^{k} c_{ij} - \sum_{j=1}^{l} d_{ij}] \, y'(x) \, h + \ldots$$

$$\ldots + \left\{ \frac{1}{\nu!} [w_i^\nu - \sum_{j=1}^{l} b_{ij}(w_{k-l+j} - m)^\nu] - \frac{1}{(\nu - 1)!} [\sum_{j=1}^{k} c_{ij} w_j^{\nu-1} + \right.$$

$$\left. + \sum_{j=1}^{l} d_{ij}(w_{k-l+j} - m)^{\nu-1}] \right\} y^{(\nu)}(x) \, h^\nu + \ldots .$$

The reader will observe later in the convergence proof that the components of this vector which correspond to the l values used as the initial values for the next step of the method have a bigger influence on the total error than the components of the error which correspond to the remaining $k - l$ values computed in one step of the method. It seems natural to define the order of the method as follows. We say that the method has order $p(p \geqq 1)$ with respect to l if the following $kp + l$ equations are satisfied:

$$\text{(28)} \qquad \sum_{j=1}^{l} b_{ij} = 1 \quad \text{for} \quad i = 1, \ldots, k ;$$

$$\text{(29)} \qquad \mu_{l-1+i}^\nu - \sum_{j=1}^{l} b_{ij} \mu_{j-1}^\nu =$$

$$= \nu [\sum_{j=1}^{k-l} c_{ij} \mu_{l-1+j}^{\nu-1} + \sum_{j=k-l+1}^{k} c_{ij}(m + \mu_{l-k-1+j})^{\nu-1} + \sum_{j=1}^{l} d_{ij} \mu_{j-1}^{\nu-1}]$$

$$\text{for} \quad i = 1, \ldots, k - l ; \quad \nu = 1, \ldots, p - 1 ;$$

$$(30) \qquad (\mu_{l-k-1+i} + m)^{\nu} - \sum_{j=1}^{l} b_{ij}\mu_{j-1}^{\nu} =$$

$$= \nu\Big[\sum_{j=1}^{k-l} c_{ij}\mu_{l-1+j}^{\nu-1} + \sum_{j=k-l+1}^{k} c_{ij}(m + \mu_{l-k-1+j})^{\nu-1} + \sum_{j=1}^{l} d_{ij}\mu_{j-1}^{\nu-1}\Big]$$

$$\text{for} \quad i = k - l + 1, \ldots, k\,; \quad \nu = 1, \ldots, p\,.$$

This definition can be, obviously, expressed in another way: the method has order p with respect to l if — for any sufficiently smooth function $y(x)$ — the first $k - l$ components of the local truncation error defined by (25) are of order h^p and the remaining l components are of order h^{p+1}.

We say that the GPOM method is consistent with respect to l if its order with respect to l is at least one, i.e., if it holds

$$(31) \qquad \sum_{j=1}^{l} b_{ij} = 1 \quad \text{for} \quad i = 1, \ldots, k\,,$$

$$(32) \qquad (\mu_{l-k-1+i} + m) - \sum_{j=1}^{l} b_{ij}\mu_{j-1} = \sum_{j=1}^{k} c_{ij} + \sum_{j=1}^{l} d_{ij}$$

$$\text{for} \quad i = k - l + 1, \ldots, k\,.$$

Introducing the l-dimensional vector $\boldsymbol{\mu} = (\mu_0, \ldots, \mu_{l-1})^{\mathrm{T}}$ we can write the consistency with respect to l in a more concise matrix form:

$$(33) \qquad \boldsymbol{B}\boldsymbol{i}^{(l)} = \boldsymbol{i}^{(k)}\,,$$

$$(34) \qquad m\boldsymbol{i}^{(l)} - (\boldsymbol{E} - \boldsymbol{I})\,\boldsymbol{\mu} = \boldsymbol{R}(\boldsymbol{C}\boldsymbol{i}^{(k)} + \boldsymbol{D}\boldsymbol{i}^{(l)})$$

where $\boldsymbol{i}^{(l)}, \boldsymbol{i}^{(k)}$ are l, k-dimensional vectors with all components equal to unities respectively; and $\boldsymbol{E} = \boldsymbol{R}\boldsymbol{B}$ where $\boldsymbol{R} = (\boldsymbol{O}_{l,k-l}, \boldsymbol{I}_l)$.

Before discussing sufficient conditions for convergence we must explain what is meant by stability: The GPOM method is stable if there exists a constant Γ such that

$$\|\boldsymbol{E}^n\| \leqq \Gamma$$

for any positive integer n.

Remark 2. Since there exists a regular matrix $\boldsymbol{T}$ such that $\boldsymbol{T}^{-1}\boldsymbol{E}\boldsymbol{T} = \boldsymbol{J}$ is in the Jordan canonical form and since $\boldsymbol{E}^n = \boldsymbol{T}\boldsymbol{J}^n\boldsymbol{T}^{-1}$ the definition of the stability can be expressed in a way that all eigenvalues λ_i of $\boldsymbol{E}$ must satisfy the inequalities $|\lambda_i| \leqq 1$ and the elementary divisors corresponding to those λ_i for which $|\lambda_i| = 1$ must be linear.

4.2 Sufficient Conditions for the Convergence of the GPOM Method

In this section we shall prove that the stability and the consistency are sufficient conditions for the convergence of our method. Before formulating the corresponding theorem we introduce the convergence concept and two important lemmas which are easy to prove.

Convergence of GPOM Method

The GPOM method will be said to be convergent if it holds

$$\lim_{\substack{h \to 0 \\ x_{jk} = x}} \tilde{y}(x_{jk+i}) = y(x) \quad \text{for} \quad i = 0, \ldots, k - 1 .$$

Here $y(x)$ is the exact solution of a differential equation of the form (1) with the right-hand term satisfying the Lipschitz condition determined by the initial condition (1) and $\tilde{y}(x)$ is any solution of the corresponding equation (8) determined by the initial conditions $\tilde{y}(x_\mu)$, $\mu = 0, \ldots, l - 1$ satisfying

$$\lim_{h \to 0} \tilde{y}(x_\mu) = \eta \quad \text{for} \quad \mu = 0, \ldots, l - 1 .$$

Lemma 1. *Let $\boldsymbol{V}$ be any square matrix for which $\|\boldsymbol{V}\| < 1$ holds where $\|\cdot\|$ is the matrix norm induced by any vector norm. Then the matrices $\boldsymbol{I} + \boldsymbol{V}$ and $\boldsymbol{I} - (\boldsymbol{I} + \boldsymbol{V})^{-1}$ are regular and it holds*

$$\|(\boldsymbol{I} + \boldsymbol{V})^{-1}\| \leqq \frac{1}{1 - \|\boldsymbol{V}\|}, \quad \|\boldsymbol{I} - (\boldsymbol{I} + \boldsymbol{V})^{-1}\| \leqq \frac{\|\boldsymbol{V}\|}{1 - \|\boldsymbol{V}\|} .$$

(See [4].)

Lemma 2. *Let $\phi(v)$, $\psi(v)$, $\chi(v)$ be defined for $v = 0, \ldots, n$ and let $\chi(v) \geqq 0$ for $v = 0, \ldots, n$. Further, let $\phi(v) \leqq \psi(v) + \sum_{\mu=0}^{v-1} \chi(\mu)\, \phi(\mu)$ for $v = 0, \ldots, n$. Then*

$$\phi(v) \leqq \psi(v) + \sum_{\mu=0}^{v-1} \chi(\mu)\, \psi(\mu) \prod_{s=\mu+1}^{v-1} (1 + \chi(s)) \quad \text{for} \quad v = 0, \ldots, n .$$

(See [5].)

Theorem 8. *The GPOM method which is stable and consistent with respect to l is convergent.*

Proof. The right-hand term of the given differential equation satisfies the Lipschitz condition. Consequently, its solution $y(x)$ has a continuous derivative and the k-dimensional local truncation error vector expression has sense. Then subtracting it from (8) and putting $e_s = \tilde{y}(x_s) - y(x_s)$, we get

$$\begin{bmatrix} e_{jk+l} \\ \vdots \\ e_{(j+1)k+l-1} \end{bmatrix} = \boldsymbol{B} \begin{bmatrix} e_{jk} \\ \vdots \\ e_{jk+l-1} \end{bmatrix} + h\boldsymbol{C\Phi}^{(k)}_{jk+l} \begin{bmatrix} e_{jk+l} \\ \vdots \\ e_{(j+1)k+l-1} \end{bmatrix} + h\boldsymbol{D\Phi}^{(l)}_{jk} \begin{bmatrix} e_{jk} \\ \vdots \\ e_{jk+l-1} \end{bmatrix} - \boldsymbol{L}(y(x_{jk}); h) \tag{35}$$

where

$$\boldsymbol{\Phi}_r^{(s)} = \begin{bmatrix} g_r \cdots 0 \\ 0 \quad g_{r+s-1} \end{bmatrix}$$

and

$$g_r = \frac{f(x_r, \tilde{y}(x_r)) - f(x_r, y(x_r))}{e_r} \quad \text{for} \quad e_r \neq 0\,,$$

$$g_r = 0 \quad \text{for} \quad e_r = 0\,.$$

Note that the fact that f satisfies the Lipschitz condition implies $\|\boldsymbol{\Phi}_r^{(s)}\| \leqq L$. Putting $\boldsymbol{e}_j = [e_{jk} \dots e_{jk+l-1}]^{\mathsf{T}}$ we can rewite (35) in the form

$$\begin{bmatrix} e_{jk+l} \\ \vdots \\ e_{(j+1)k+l-1} \end{bmatrix} - h\boldsymbol{C}\boldsymbol{\Phi}_{jk+l}^{(k)} \begin{bmatrix} e_{jk+l} \\ \vdots \\ e_{(j+1)k+l-1} \end{bmatrix} = \boldsymbol{B}\boldsymbol{e}_j + h\boldsymbol{D}\boldsymbol{\Phi}_{jk}^{(l)}\boldsymbol{e}_j - \boldsymbol{L}(y(x_{jk}); h)\,.$$

Thus we see that the error in the given step depends only on the components $e_{jk}, \dots$ $\dots, e_{jk+l-1}$ of the error in the preceding step. This could be expected since only the components $\tilde{y}(x_{jk}), \dots, \tilde{y}(x_{jk+l-1})$ of the vector $\tilde{\boldsymbol{y}}(\boldsymbol{z}_{j-1})$ were used for the computation of the vector $\tilde{\boldsymbol{y}}(\boldsymbol{z}_j)$. On the basis of Lemma 1 and taking into account that $\|\boldsymbol{\Phi}_r^{(s)}\| \leqq L$ we can assert that for $h \leqq h_0 < 1/(L\|\boldsymbol{C}\|)$ the matrix $(\boldsymbol{I} - h\boldsymbol{C}\boldsymbol{\Phi}_{jk+l}^{(k)})$ is regular.

Consequently, premultiplying both sides of the last equation by the matrix $(\boldsymbol{I} - h\boldsymbol{C}\boldsymbol{\Phi}_{jk+l}^{(k)})^{-1}$ which is denoted for simplicity by $\boldsymbol{\Lambda}$ we get

$$\begin{bmatrix} e_{jk+l} \\ \vdots \\ e_{(j+1)k+l-1} \end{bmatrix} = \boldsymbol{\Lambda}\boldsymbol{B}\boldsymbol{e}_j + h\boldsymbol{\Lambda}\boldsymbol{D}\boldsymbol{\Phi}_{jk}^{(l)}\boldsymbol{e}_j - \boldsymbol{\Lambda}\boldsymbol{L}(y(x_{jk}); h)\,.$$

Since we are interested only in the behaviour of $\boldsymbol{e}_j$ we premultiply, moreover, both sides of this equation by the matrix

$$\boldsymbol{R} = [\boldsymbol{O}_{l,k-l}\ \boldsymbol{I}_l]\,.$$

We get

$$\boldsymbol{e}_{j+1} = \boldsymbol{R}\boldsymbol{\Lambda}\boldsymbol{B}\boldsymbol{e}_j + h\boldsymbol{R}\boldsymbol{\Lambda}\boldsymbol{D}\boldsymbol{\Phi}_{jk}^{(l)}\boldsymbol{e}_j - \boldsymbol{R}\boldsymbol{\Lambda}\boldsymbol{L}(y(x_{jk}); h)$$

or, since $\boldsymbol{E} = \boldsymbol{R}\boldsymbol{B}$,

$$\boldsymbol{e}_{j+1} = \boldsymbol{E}\boldsymbol{e}_j - \boldsymbol{R}(\boldsymbol{I} - \boldsymbol{\Lambda})\,\boldsymbol{B}\boldsymbol{e}_j + h\boldsymbol{R}\boldsymbol{\Lambda}\boldsymbol{D}\boldsymbol{\Phi}_{jk}^{(l)}\boldsymbol{e}_j - \boldsymbol{R}\boldsymbol{\Lambda}\boldsymbol{L}(y(x_{jk}); h)\,.$$

Let

$$\boldsymbol{v}_j = -\boldsymbol{R}(\boldsymbol{I} - \boldsymbol{\Lambda})\,\boldsymbol{B}\boldsymbol{e}_j + h\boldsymbol{R}\boldsymbol{\Lambda}\boldsymbol{D}\boldsymbol{\Phi}_{jk}^{(l)}\boldsymbol{e}_j - \boldsymbol{R}\boldsymbol{\Lambda}\boldsymbol{L}(y(x_{jk}); h)\,. \tag{36}$$

Then

$$\boldsymbol{e}_{j+1} = \boldsymbol{E}\boldsymbol{e}_j + \boldsymbol{v}_j \quad \text{for} \quad j = 0, 1, \dots,$$

which is equivalent to

$$(37) \qquad e_j = E^j e_0 + \sum_{\nu=0}^{j-1} E^{j-1-\nu} v_\nu$$

as can be easily shown by induction. Thus, to be able to estimate e_j we must estimate v_ν. The first step to achieve it will be an estimation of L.

Since the derivative of the exact solution of our differential equation is continuous, we can define a function $\Omega(x)$ such that

$$(38) \qquad \Omega(\delta) = \max_{\substack{|x-x^*|\leqq\delta \\ x,x^*\in[a,b]}} |y'(x) - y'(x^*)|$$

and it will be

$$\lim_{\delta\to 0} \Omega(\delta) = 0\,.$$

Using this function we have

$$|y'(x_{sk} + w_i h) - y'(x_{sk})| \leqq \Omega(w_i h)\,, \quad i = 1, \ldots, k$$

and consequently, there exist constants $\Theta_i^{(1)}$ such that

$$|\Theta_i^{(1)}| \leqq 1$$

and

$$y'(x_{sk} + w_i h) = y'(x_{sk}) + \Theta_i^{(1)}\, \Omega(w_i h)\,.$$

Analogously, there exist constants $\Theta_i^{(2)}$ such that

$$|\Theta_i^{(2)}| \leqq 1$$

and

$$y'(x_{sk} + \mu_{i-1} h) = y'(x_{sk}) + \Theta_i^{(2)}\, \Omega(\mu_{i-1} h)\,.$$

Further, the mean value theorem yields

$$y(x_{sk} + w_i h) = y(x_{sk}) + w_i h y'(x_{sk} + \vartheta_i^{(1)} w_i h) \quad \text{with} \quad |\vartheta_i^{(1)}| \leqq 1\,,$$

$$y(x_{sk} + \mu_{i-1} h) = y(x_{sk}) + \mu_{i-1} h y'(x_{sk} + \vartheta_i^{(2)} \mu_{i-1} h) \quad \text{with} \quad |\vartheta_i^{(2)}| \leqq 1\,.$$

Therefore, these last two equations and the definition of the function Ω imply the existence of numbers $\Theta_i^{(3)}$ and $\Theta_i^{(4)}$ such that

$$|\Theta_i^{(3)}| \leqq 1\,, \quad |\Theta_i^{(4)}| \leqq 1\,,$$

and

$$y(x_{sk} + w_i h) = y(x_{sk}) + w_i h\, y'(x_{sk}) + w_i h \Theta_i^{(3)}\, \Omega(w_i h)\,,$$

$$y(x_{sk} + \mu_{i-1} h) = y(x_{sk}) + \mu_{i-1} h\, y'(x_{sk}) + \mu_{i-1} h \Theta_i^{(4)}\, \Omega(\mu_{i-1} h)\,.$$

Substituting the just obtained estimates into (26) we obtain

$$L_i(y(x_{sk}); h) = y(x_{sk}) + w_i h\, y'(x_{sk}) + w_i h \Theta_i^{(3)}\, \Omega(w_i h) -$$
$$- h \sum_{j=1}^{k} c_{ij}[y'(x_{sk}) + \Theta_j^{(1)}\, \Omega(w_j h)] -$$
$$- \sum_{j=1}^{l} b_{ij}[y(x_{sk}) + \mu_{j-1} h\, y(x_{sk}) + \mu_{j-1} h \Theta_j^{(4)}\, \Omega(\mu_{j-1} h)] -$$
$$- h \sum_{j=1}^{l} d_{ij}[y'(x_{sk}) + \Theta_j^{(2)}\, \Omega(\mu_{j-1} h)] =$$
$$= (1 - \sum_{j=1}^{l} b_{ij})\, y(x_{sk}) + (w_i - \sum_{j=1}^{l} b_{ij}\mu_{j-1} - \sum_{j=1}^{k} c_{ij} - \sum_{j=1}^{l} d_{ij})\, h y'(x_{sk}) +$$
$$+ w_i h \Theta_i^{(3)}\, \Omega(w_i h) - h \sum_{j=1}^{l} b_{ij}\mu_{j-1}\Theta_j^{(4)}\, \Omega(\mu_{j-1} h) -$$
$$- h \sum_{j=1}^{k} c_{ij}\Theta_j^{(1)}\, \Omega(w_j h) - h \sum_{j=1}^{l} d_{ij}\Theta_j^{(2)}\, \Omega(\mu_{j-1} h)\,.$$

Since our method is consistent with respect to l, the first two members in the right-hand term are for $i = k - l + 1, \dots, k$ equal to zero. Thus, if we put

$$M = \max_{i=1,\dots,k} \left(|w_i| + \sum_{j=1}^{l} |b_{ij}\mu_{j-1}| + \sum_{j=1}^{k} |c_{ij}| + \sum_{j=1}^{l} |d_{ij}|\right) \tag{39}$$

and

$$\gamma = \max\left(\max_{i=1,\dots,k} |w_i|,\ \max_{j=1,\dots,l} |\mu_{j-1}|\right)$$

we get

$$|L_i(y(x_{sk}); h)| \leqq hM\, \Omega(\gamma h) \quad \text{for} \quad i = k - l + 1, \dots, k\,; \quad s = 0, 1, \dots . \tag{40}$$

Let us now turn back to the estimation of $\boldsymbol{v}_\nu$. From Lemma 1 and for $h \leqq h_0 < 1/)L\|\boldsymbol{C}\|)$ we get

$$\|\boldsymbol{A}\| = \|(\boldsymbol{I} - h\boldsymbol{C}\boldsymbol{\Phi}_{jk+l}^{(k)})^{-1}\| \leqq \frac{1}{1 - h\, L\|\boldsymbol{C}\|} \leqq \frac{1}{1 - h_0\, L\|\boldsymbol{C}\|} = \frac{1}{1 - h_0\delta} = \beta$$

and

$$\|\boldsymbol{I} - \boldsymbol{A}\| = \|\boldsymbol{I} - (\boldsymbol{I} - h\boldsymbol{C}\boldsymbol{\Phi}_{jk+l}^{(k)})^{-1}\| \leqq \frac{h\, L\|\boldsymbol{C}\|}{1 - h\, L\|\boldsymbol{C}\|} \leqq \frac{h\, L\|\boldsymbol{C}\|}{1 - h_0\, L\|\boldsymbol{C}\|} = h\delta\beta$$

where

$$\delta = L\|\boldsymbol{C}\|\,, \quad \beta = \frac{1}{1 - h_0\, L\|\boldsymbol{C}\|}\,.$$

Using the obvious identity

$$(\boldsymbol{I} - h\boldsymbol{C}\boldsymbol{\Phi}_{jk+l}^{(k)})^{-1} = \boldsymbol{I} + h\boldsymbol{C}\boldsymbol{\Phi}_{jk+l}^{(k)}(\boldsymbol{I} - h\boldsymbol{C}\boldsymbol{\Phi}_{jk+l}^{(k)})^{-1}$$

or

$$A = I + hC\Phi^k_{jk+l}A$$

we can write

$$RAL(y(x_{sk}); h) = RL(y(x_{sk}); h) + hRC\Phi^k_{jk+l}AL(y(x_{sk}); h)\,. \tag{41}$$

Using the mean-value theorem we can write

$$\begin{aligned} L_i(y(x_{sk}); h) = (1 - \sum_{j=1}^{l} b_{ij})\, y(x_{sk}) + h[w_i y'(x_{sk} + \delta_i^{(1)} w_i h) - \\ - \sum_{j=1}^{l} b_{ij}\mu_{j-1} y'(x_{sk} + \delta_i^{(2)}\mu_{j-1}h) - \sum_{j=1}^{k} c_{ij} y'(x_{sk} + w_j h) - \\ - \sum_{j=1}^{l} d_{ij} y'(x_{sk} + \mu_{j-1}h)] \end{aligned} \tag{42}$$

for $i = 1, \ldots, k$ with $|\delta_i^{(1)}| \leqq 1$, $|\delta_i^{(2)}| \leqq 1$

but the first term on the right hand side of this equation is equal to zero since the method is consistent with respect to l (cf. (5)). Consequently we get

$$|L_i(y(x_{sk}); h)| \leqq hMY \quad \text{for} \quad i = 1, \ldots, k \tag{43}$$

where M is given by (39) and $Y = \max_{x \in [a,b]} |y'(x)|$.

In virtue of (40) and (43) the expression (41) can be estimated as

$$\|RAL(y(x_{sk}); h)\| \leqq hM\,\Omega(\gamma h) + h^2 \beta\delta MY\,. \tag{44}$$

Substituting this estimate into (36) we get

$$\|v_\nu\| \leqq \alpha h \|e_\nu\| + hM\,\Omega(\gamma h) + h^2\beta\delta MY$$

where

$$\alpha = \beta(\delta\|B\| + L\|D\|)\,.$$

Applying this estimate and the condition of stability we get

$$\|e_j\| \leqq \Gamma\|e_0\| + \Gamma \sum_{\nu=0}^{j-1} [\alpha h \|e_\nu\| + hM\,\Omega(\gamma h) + h^2\beta\delta MY]\,.$$

Using now Lemma 2 and the obvious identities

$$\sum_{\nu=0}^{j-1} (1 + \alpha\Gamma h)^{j-1-\nu} = \frac{(1 + \alpha\Gamma h)^j - 1}{(\alpha\Gamma h)}$$

and

$$\sum_{\nu=0}^{j-1} \nu(1 + \alpha\Gamma h)^{j-1-\nu} = \frac{(1 + \alpha\Gamma h)^j - 1}{(\alpha\Gamma h)^2} - \frac{j}{\alpha\Gamma h}$$

we have

$$\|e_j\| \leqq \Gamma(1 + \alpha\Gamma h)^j \|e_0\| + \Gamma(hM\Omega(\gamma h) + h^2\beta\delta MY)\frac{(1 + \alpha\Gamma h)^j - 1}{\alpha\Gamma h}.$$

Using the estimates

$$(1 + \alpha\Gamma h)^j < e^{\Gamma\alpha jh}$$

and

$$\frac{(1 + \alpha\Gamma h)^j - 1}{\alpha\Gamma h} \leqq j e^{\alpha\Gamma jh}$$

we obtain finally

$$\|e_j\| \leqq \Gamma e^{\alpha\Gamma(x_{jk} - a)/m}\left[\|e_0\| + M(\Omega(\gamma h) + h\beta\delta Y)\frac{1}{m}(x_{jk} - a)\right]$$

where m is the constant from the definition of basic points. This implies the convergence immediately.

The error estimate can be discussed after presenting the following clear lemma which can be proved by using Taylor expansion with the remainder in an integral form similarly to that in [2] p. 132.

Lemma 3. *Let the GPOM method of order $p(p \geqq 1)$ with respect to l be given and let $y \in C^{p+1}$. Then*

$$\|L(y(x); h)\| \leqq S_1 Y_1 h^p, \quad \|RL(y(x), h)\| \leqq S_2 Y_2 h^{p+1}$$

where

$$Y_1 = \max_{x \in [a,b]} |y^{(p)}(x)|; \quad Y_2 = \max_{x \in [a,b]} |y^{(p+1)}(x)|$$

and S_1 and S_2 are constants depending only on the parameters of the given method and independent of h.

Now we can formulate a theorem giving the error estimate of the GPOM method.

Theorem 9. *Let the solution $y(x)$ of (1) have $p + 1$ continuous derivatives in $[a, b]$. Further, let $\tilde{y}(x)$ be the approximate solution computed by the GPOM method of order $p \geqq 1$ with respect to l. Then it holds*

$$\|\tilde{y}(x_j) - y(x_j)\| \leqq$$

$$\leqq \Gamma e^{\alpha\Gamma(x_{jk} - a)/m}\left[\|\tilde{y}(x_0) - y(x_0)\| + \frac{1}{m}(x_{jk} - a)(S_2 Y_2 + \delta\beta S_1 Y_1) h^p\right].$$

Proof. The error e_j satisfies again the equation (37). Using the assumptions of the theorem and the lemma we can estimate v_ν:

$$\|v_\nu\| \leqq \alpha h\|e_\nu\| + S_2 Y_2 h^{p+1} + \delta\beta S_1 Y_1 h^{p+1}.$$

Continuing in the same way as in the proof of Theorem 8 we get the final result.

Remark 3. Theorem 9 shows that if the errors in the initial conditions are of orders h^p then the total error of the method under consideration is also of order h^p. Essperially, this interpretation of Theorem 9 is very important from the practical point of view.

Remark 4. Defining the order of the method we formulate in fact the assumptions on the behaviour of the local truncation error by putting stronger conditions only on those components of it which are used as starting values for the next step of the method. The assumptions concerning the remaining components might be weakened due to the identity (41) since they are multiplied by h. A natural question arises asking whether it is possible to continue this process. The answer is probably affirmative since it is possible to write

$$(\boldsymbol{I} - h\boldsymbol{C\Phi})^{-1} = \boldsymbol{I} + h\boldsymbol{C\Phi} + h^2(\boldsymbol{C\Phi})^2 + \dots h^{r-1}(\boldsymbol{C\Phi})^{r-1} + \\ + h^r(\boldsymbol{C\Phi})^r(\boldsymbol{I} - h\boldsymbol{C\Phi})^{-1}.$$

Consequently, it is sufficient to assume that

$$(45) \qquad \boldsymbol{RL} = O(h^{p+1}), \quad \boldsymbol{RC\Phi L} = O(h^p), \dots, \boldsymbol{R}(\boldsymbol{C\Phi})^{r-1}\boldsymbol{L} = O(h^{p-r+2}), \\ \boldsymbol{R}(\boldsymbol{C\Phi})^r(\boldsymbol{I} - h\boldsymbol{C\Phi})^{-1}\boldsymbol{L} = O(h^{p-r+1})$$

to guarantee the global truncation error to be of order h^p.

However, the development of algebraic relations for the parameters of the method from (45) is extremely complex (very similar to that in Runge-Kutta formulae) and so we have till now no definite results in this direction. We hope that we shall be able to solve these problems in some further paper.

References

[1] *Práger, Taufer* and *Vitásek:* Overimplicit Multistep Methods. Aplikace matematiky, 18 (1973), 399—421.

[2] *Henrici, P.:* Discrete Variable Methods in Ordinary Differential Equations. John Wiley and Sons, Inc. (1962).

[3] *Lambert J. D.:* Computational Methods in Ordinary Differential Equations. John Wiley and Sons Ltd. (1973).

[4] *Ralston A.:* A First Course in Numerical Analysis. McGraw-Hill, Inc. (1965).

[5] *Babuška, Práger* and *Vitásek:* Numerical Processes in Differential Equations. SNTL -Publishers of Technical Literature, Prague (1966).

Souhrn

ZOBECNĚNÉ PERIODICKÉ SILNĚ IMPLICITNÍ MNOHOKROKOVÉ METODY

HASSAN NASR AHMED ISMAIL

V článku se studují metody pro řešení úloh s počátečními podmínkami pro obyčejné diferenciální rovnice, jejichž podstata spočívá v tom, že v jednom kroku se počítají přibližné hodnoty hledaného řešení v několika bodech definičního intervalu najednou z jakési (obecně nelineární) soustavy rovnic. Studovaná třída metod je zformulována natolik obecně, že obsahuje všechny běžně známé třídy metod (lineární k-krokové metody, metody typu prediktor-korektor, explicitní i implicitní metody typu Runge-Kutta atd.). Jsou nalezeny postačující podmínky konvergence a podmínky pro to, aby daná metoda byla určitého řádu.

Author's address: Hassan Nasr Ahmed Imsail, Military Technical College, Cairo, EAR.