BOLLETTINO UNIONE MATEMATICA ITALIANA

LIDIA ACETO

Some applications of the Pascal matrix to the study of numerical methods for differential equations

Bollettino dell'Unione Matematica Italiana, Serie 8, Vol. 8-B (2005), n.3, p. 639–651.

Unione Matematica Italiana

<http://www.bdim.eu/item?id=BUMI_2005_8_8B_3_639_0>

L'utilizzo e la stampa di questo documento digitale è consentito liberamente per motivi di ricerca e studio. Non è consentito l'utilizzo dello stesso per motivi commerciali. Tutte le copie di questo documento devono riportare questo avvertimento.

> Articolo digitalizzato nel quadro del programma bdim (Biblioteca Digitale Italiana di Matematica) SIMAI & UMI http://www.bdim.eu/

Bollettino dell'Unione Matematica Italiana, Unione Matematica Italiana, 2005.

Some Applications of the Pascal Matrix to the Study of Numerical Methods for Differential Equations (*).

LIDIA ACETO (**)

Sunto. – In questo articolo analizziamo i legami tra la matrice di Pascal e una nuova classe di metodi numerici per equazioni differenziali ottenuti come generalizzazione dei metodi di Adams. In particolare, proveremo che i metodi in tale classe possono essere utilizzati per risolvere problemi di tipo stiff in quanto le regioni di assoluta stabilità ad essi associate contengono il semipiano negativo.

Summary. – In this paper we introduce and analyze some relations between the Pascal matrix and a new class of numerical methods for differential equations obtained generalizing the Adams methods. In particular, we shall prove that these methods are suitable for solving stiff problems since their absolute stability regions contain the negative half complex plane.

1. – Introduction.

The Adams-Moulton methods have been used in the past mainly for approximating the solution of non-stiff problems (see, for example, [7,11]). The reason is that, with the exception of the trapezoidal rule, they have bounded absolute stability regions whose sizes decrease as the order increases. Recently, the boundary value methods (BVMs) approach to deal with the numerical approximation of differential problems has evidenced that some generalization of Adams methods can be successfully used also for solving stiff problems. Such approach consists in replacing a given continuous initial value problem with a k-step linear multistep formula, associated with boundary conditions instead of initial ones, as traditionally done. The use of a discrete boundary value problem is justified by the fact that a k-step linear multistep formula ge-

^(*) Work supported by GNCS-INdAM.

^(**) Comunicazione presentata a Milano in occasione del XVII Congresso U.M.I.

nerates a *k*th order discrete problem, whose particular solution is obtained by fixing *k* independent conditions. Since only one of them is provided by the continuous problem, there is the freedom of choosing the k-1 remaining values. The BVMs approach splits such additional conditions at the beginning and at the end of the interval. This leads an improvement of the stability properties of the linear multistep formula (see [6] and references therein). As a matter of fact, some classes of BVMs obtained from a generalization of Adams methods, namely the Reverse Adams methods [4] (up to order four) and the GAMs [5], have stability regions containing the negative half complex plane. For this reason, such methods are suitable for approximating the solution of stiff problems. The code GAM [8, 9], one of the codes for solving stiff initial value problems, is in fact based on some methods in the class of GAMs.

The aim of this paper is to introduce and to analyze the properties of a new class of BVMs, obtained as a generalization of Adams methods as well, and having an odd number of steps. This will be the subject of Section 3. Some numerical experiments, comparing the performance of the presented new methods with other existing ones, will be presented in Section 4. Before that, in the next section the family of BVMs that generalizes the classical Adams methods is examined, along with the extension of the classical results about stability.

2. - The family of generalized Adams methods.

The family of generalized Adams methods (fGAMs) contains k-step methods defined by

(1)
$$y_{n+j} - y_{n+j-1} = h \sum_{i=0}^{k} \beta_i^{(j)} f_{n+i}, \quad j \in \{1, 2, \dots, k\},$$

where y_n is the approximation to the solution at the grid points $t_n = t_0 + nh$, n = 0, ..., N, $h = (T - t_0)/N$ and $f_n = f(t_n, y_n)$.

For a fixed j, the coefficients $\{\beta_i^{(j)}\}\$ are uniquely derived by imposing that the method is of maximum order k + 1. This requirement is equivalent to derive such parameters as solution of the linear system [6]

(2)
$$W(0) \boldsymbol{\beta}^{(j)} = \boldsymbol{v}$$

where

$$W(0) = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 2 & \cdots & k \\ 0 & 1 & 2^2 & \cdots & k^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 2^k & \cdots & k^k \end{pmatrix}, \quad v = \begin{pmatrix} 1 \\ \frac{j^2 - (j-1)^2}{2} \\ \frac{j^3 - (j-1)^3}{3} \\ \vdots \\ \frac{j^{k+1} - (j-1)^{k+1}}{k+1} \end{pmatrix}$$

and $\boldsymbol{\beta}^{(j)} = (\beta_0^{(j)}, \beta_1^{(j)}, \dots, \beta_k^{(j)})^T$. By introducing

(3)
$$\boldsymbol{\xi}(x) = (1, x, x^2, \dots, x^k)^T,$$

the vector v can be written as $v = \int_{j-1}^{j} \xi(t) dt$. Moreover, by considering the change of variable t = x + j - 1, it becomes

(4)
$$v = \int_{0}^{1} \xi(x+j-1) \, dx \, .$$

Such relation, when substituted in (2), leads to

(5)
$$\boldsymbol{\beta}^{(j)} = W(0)^{-1} \int_{0}^{1} \boldsymbol{\xi}(x+j-1) \, dx \, dx$$

In order to get an analytic expression of such solution, we need to use the Pascal matrix, whose entries are

$$P_{ij} = \left\{ egin{array}{cc} (i \ j) & ext{for } i \geqslant j \ 0 & ext{otherwise} \end{array}
ight. i, j = 0, 1, \dots, k ,$$

and some of its properties. It is an easy matter to check that

$$P\xi(x) = \xi(x+1).$$

Thus, the Vandermonde matrix based on the distinct values x, x + 1, ..., x + k, i.e.,

(6)
$$W(x) = (\xi(x) \ \xi(x+1) \ \xi(x+2) \ \dots \ \xi(x+k))$$

becomes

$$W(x) = P^{x}(\xi(0) \ \xi(1) \ \xi(2) \ \dots \ \xi(k))$$

or, equivalently,

$$W(x) = P^x W(0).$$

By observing that $W(x) e_{s-1} = \xi(x+s-1)$, where e_{s-1} denotes the *s*th standard unit basis vector in \mathbb{R}^{k+1} , one has that

(7)
$$\xi(x+s-1) = P^x W(0) e_{s-1}.$$

Moreover, by using such relation in (5) one obtains that

(8)
$$\boldsymbol{\beta}^{(j)} = W(0)^{-1} \left(\int_{0}^{1} P^{x} dx \right) W(0) \boldsymbol{e}_{j-1}, \quad j \in \{1, \dots, k\},$$

that is, the coefficients defining each method in the fGAMs are the columns entries of the matrix $W(0)^{-1} \left(\int_{0}^{1} P^{x} dx \right) W(0)$. In [1], by a long proof which we shall omit here, it has been shown that

(9)
$$\left(W(0)^{-1} \int_{0}^{1} P^{x} dx W(0) \right)_{ir} = \int_{0}^{1} \binom{x+r}{i} \binom{k-(x+r)}{k-i} dx ,$$

 $i, r = 0, \ldots, k$. Then, the following result holds true.

THEOREM 2.1. – For a fixed j, the entries of vector $\boldsymbol{\beta}^{(j)}$, defining each method in the fGAMs, are

(10)
$$\boldsymbol{\beta}_{i}^{(j)} = \int_{0}^{1} {\binom{x+j-1}{i} \binom{k-(x+j-1)}{k-i}} dx, \quad i = 0, \dots, k.$$

The methods obtained in correspondence of the choice j = k are the classical Adams-Moulton methods. They are used as initial value problems, i.e., by imposing in the discrete problem all the k - 1 additional conditions at the initial points. It is known that the only Adams-Moulton method having an unbounded absolute stability region is the trapezoidal rule, which corresponds to the case k = 1. For k > 1, the stability polynomial

(11)
$$\pi^{(k)}(z, q) = z^{k-1}(z-1) - q \sum_{i=0}^{k} \beta_i^{(k)} z^i,$$

obtained applying the formula to the test equation $y' = \lambda y$, Re $\lambda < 0$, and posing $q = h\lambda$, has roots inside the unit circle only if *h* is small enough. Such severe requirement on the location of the roots can be weakened by considering

642

the BVMs approach [6]. In fact, when using formula (1) associated with j initial conditions and k - j final ones on the test equation, if the roots of the characteristic polynomial

(12)
$$\pi^{(j)}(z, q) = z^{j-1}(z-1) - h\lambda \sum_{i=0}^{k} \beta_i^{(j)} z^i \equiv \varrho^{(j)}(z) - q\sigma^{(j)}(z)$$

are such that

(13)
$$|z_1| \leq |z_2| \leq \ldots \leq |z_{j-1}| < |z_j| < |z_{j+1}| \leq \ldots \leq |z_k|$$

 $|z_{j-1}| < 1 < |z_{j+1}|,$

for n and N-n large, the solution y_n behaves as

(14)
$$y_n \approx \gamma z_j^n$$
,

where γ only depends on the initial conditions. Moreover, when *h* is suitably small, the *generating root* z_j will be the approximation to a certain order *p* of e^q , i.e.,

(15)
$$z_i = e^q + O(h^{p+1}).$$

This leads to generalize the classic concepts of 0-stability, convergence and Astability. We consider here a short overview about such generalizations applied to the methods in the fGAMs. Before that, we need to introduce the following definitions generalizing the classical ones of Schur and Von Neumann polynomials.

DEFINITION 2.1. – A polynomial p(z) of degree k is an

- $S_{j, k-j}$ -polynomial if its roots are such that

$$|z_1| \leq \ldots \leq |z_j| < 1 < |z_{j+1}| \leq \ldots \leq |z_k|;$$

- $N_{j, k-j}$ -polynomial if

$$|z_1| \leq \ldots \leq |z_j| \leq 1 < |z_{j+1}| \leq \ldots \leq |z_k|,$$

being simple the roots of unit modulus.

DEFINITION 2.2. – For a fixed j, a BVM in the fGAMs is $0_{j,k-j}$ -stable if $\varrho^{(j)}(z)$ is an $N_{j,k-j}$ -polynomial.

Since $\varrho^{(j)}(z) = z^{j-1}(z-1)$ (see (12)), all the methods in such family are $0_{j, k-j}$ -stable by construction. Moreover, by ignoring the effects of round-off errors, it can be shown that they are also convergent [6].

LIDIA ACETO

DEFINITION 2.3. – The region $\mathcal{O}_{j, k-j}$ of the complex plane defined by

 $\mathcal{O}_{j,k-j} = \{ q \in \mathbb{C} : \pi^{(j)}(z,q) \text{ is an } S_{j,k-j} \text{-polynomial} \}$

is called the region of (j, k-j)-absolute stability.

DEFINITION 2.4. – For a fixed j, a BVM in the fGAMs is said to be $A_{j, k-j}$ -stable if $\mathbb{C}^- \subseteq \mathcal{O}_{j, k-j}$.

The boundary locus, defined by

(16)
$$\Gamma^{(j)} = \left\{ q \in \mathbb{C} : q = q^{(j)}(e^{i\theta}), \ 0 \le \theta < 2\pi \right\},$$

where (see (12))

$$q^{(j)}(e^{i heta}) = rac{arrho^{(j)}(e^{i heta})}{\sigma^{(j)}(e^{i heta})}\,,$$

assumes then special importance in discussing the stability of the methods. In fact, when this set is a regular Jordan curve, it coincides with the boundary of the (j, k - j)-absolute stability region of the method. In general, it is difficult to recognize when this happens. Nevertheless, it holds true if

(17)
$$Re(q^{(j)}(e^{i\theta})) = \frac{c(1-\cos\theta)^m}{\phi^{(j)}(\theta)}$$

(18)
$$Im(q^{(j)}(e^{i\theta})) = \frac{\sin\theta f(\theta)}{\phi^{(j)}(\theta)}, \quad \theta \in [0, 2\pi)$$

with $c \ge 0$, m a positive integer, $f(\theta) > 0$ and

$$\phi^{(j)}(\theta) = |\sigma^{(j)}(e^{i\theta})|^2$$

which decreases in $(0, \pi)$ and increases in $(\pi, 2\pi)$ (for more details, see [6]).

Numerical evidences suggest that the $A_{j, k-j}$ -stability is not satisfied for all values of $j \in \{1, ..., k\}$. However, it is known that by choosing in (1)

(19)
$$j = \nu \equiv \begin{cases} \frac{k}{2} & \text{for even } k \\ \frac{k+1}{2} & \text{for odd } k \end{cases},$$

for all $k \ge 1$ the obtained methods are $A_{\nu, k-\nu}$ -stable. In particular, when k is even these methods are the GAMs and, respectively, when k is odd they are the so called ETRs [3]. Nevertheless, in the fGAMs we will show that there

644

exists another class of methods, neglected so far, having good stability properties for fixed stepsize.

3. - Odd-GAMs.

The class of k-step methods in the fGAMs having $k \equiv 2\nu - 1$ and obtained by choosing $j = \nu - 1$ in (1), is now considered. Such methods, that we shall call *Odd-GAMs* (*OGAMs*), are defined by

(20)
$$y_{n+\nu-1} - y_{n+\nu-2} = h \sum_{i=0}^{k} \beta_{i}^{(\nu-1)} f_{n+i}.$$

The coefficients $\{\beta_i^{(\nu-1)}\}\$, whose explicit form can be derived from (10), are

$$\beta_i^{(\nu-1)} = \int_0^1 \binom{x+\nu-2}{i} \binom{k-(x+\nu-2)}{k-i} \, dx, \qquad i = 0, \, 1, \, \dots, \, k \, .$$

They make (20) of order k + 1. In Table 1, we report the coefficients of the OGAMs for k = 3, 5, 7. For convenience, we list the normalized coefficients $\hat{\beta}_i = \eta_k \beta_i^{(\nu-1)}, i = 0, ..., k$.

These methods are used with $\nu - 1$ initial conditions and ν final ones, i.e., as BVMs with $(\nu - 1, \nu)$ -boundary conditions. We skip here and in the following the problem related to the choice of the additional conditions, already discussed in [6].

By direct calculation we have verified, up to order 10, that the points on the boundary locus of OGAMs satisfy (17)-(18). In Table 2, the corresponding parameters c and m are reported.

k	$\nu - 1$	η_k	\widehat{eta}_0	$\widehat{\beta}_{1}$	\widehat{eta}_2	\widehat{eta}_3	\widehat{eta}_4	$\widehat{\beta}_{5}$	\widehat{eta}_{6}	$\widehat{\beta}_{7}$
$ \begin{array}{c} 3 \\ 5 \\ 7 \end{array} $	$\begin{array}{c}1\\2\\3\end{array}$	$24 \\ 1440 \\ 120960$	$9 \\ -27 \\ 351$	$19 \\ 637 \\ -4183$	$-5 \\ 1022 \\ 57627$	$1 \\ -258 \\ 81693$	77 - 20227	$-11 \\ 7227$	- 1719	191

TABLE 1. - Normalized coefficients of OGAMs.

TABLE 2. – Parameters c and m of formula (17) for OGAMs of order 4,6,8,10.

k	3	5	7	9
m	3	4	5	6
c	$\frac{1}{6}$	$\frac{11}{180}$	$\frac{191}{7560}$	$\frac{2497}{226800}$

LIDIA ACETO

EXAMPLE 3.1. – For k = 3, one obtains the following analytic expression of the boundary locus associated to the corresponding OGAM,

$$Re(q^{(1)}(e^{i\theta})) = \frac{\frac{1}{6}(1 - \cos\theta)^3}{\phi^{(1)}(\theta)},$$
$$Im(q^{(1)}(e^{i\theta})) = \frac{\frac{1}{6}\sin\theta(8 - 3\cos\theta + \cos^2\theta)}{\phi^{(1)}(\theta)},$$

where $\phi^{(1)}(\theta) = \frac{1}{72}(65 + 11\cos\theta - 13\cos^2\theta + 9\cos^3\theta).$

In Figure 1 the boundary loci of OGAMs are shown for k = 3, 5, ..., 29. The stability regions are the unbounded sets delimited by the corresponding curves. It is evident that \mathbb{C}^- is contained in the $(\nu - 1, \nu)$ -absolute stability region of each method. Therefore, all these formulae are $A_{\nu-1,\nu}$ -stable.

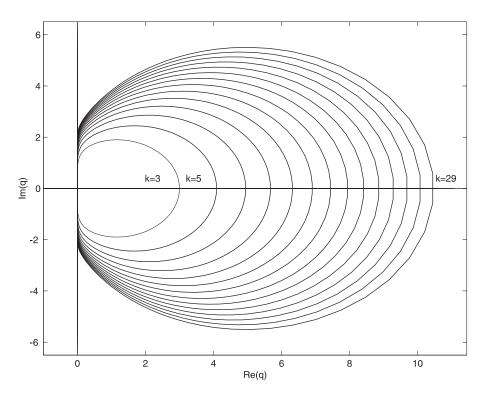


Fig. 1. – Boundary loci of the OGAMs of order k + 1, with k = 3, 5, ..., 29.

646

4. - Numerical examples.

This section deals with numerical examples of two types: those which confirm the properties of OGAMs previously derived and those which give a first comparison of their performances with other existing methods.

In order to verify the order of convergence for the OGAMs, we approximate the solutions of the following linear stiff problem

(21)
$$\mathbf{y}' = \begin{pmatrix} -21 & 19 & -20 \\ 19 & -21 & 20 \\ 40 & -40 & -40 \end{pmatrix} \mathbf{y}, \quad \mathbf{y}(0) = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

by using the OGAMs with k = 3, 5, 7, 9 steps over the time interval [0, 1]. In Table 3 the measured maximum absolute errors, for the specified constant stepsize h, and the estimated rate of convergence are reported. In all the cases, as the stepsize is decreased, the observed convergence rate is near the expected one, thus confirming the predicted order k + 1 for each formula.

We now compare the performances of OGAMs with those of GBDF (see [2] and [6]) and some implicit Runge-Kutta methods, respectively, when they are applied to the following stiff problem:

(22)
$$\mathbf{y}' = \begin{pmatrix} -2 & 1\\ 378 & -379 \end{pmatrix} \mathbf{y}, \quad \mathbf{y}(0) = \begin{pmatrix} -3\\ 2 \end{pmatrix}$$

First of all, we integrate (22) by using the OGAM, GBDF, Lobatto IIIA and Lobatto IIIC methods of order four, with constant stepsize h. In Figure 2 the relative error, evaluated as

(23)
$$\max_{1 \leq i \leq N} \frac{\|\boldsymbol{y}(t_i) - \boldsymbol{y}_i\|_{\infty}}{1 + \|\boldsymbol{y}(t_i)\|_{\infty}},$$

is plotted in [0,0.003] where the solution has its greatest variation. As one can see, OGAM works better than GBDF and Lobatto IIIC and worse than Lobatto IIIA.

Such behavior can be partially explained by looking at the principal part of

h	k = 3		k = 5		k=7		k = 9	
	error	rate	error	rate	error	rate	error	rate
2e-2 1e-2 5e-3 2.5e-3	9.544e-03 8.070e-04 6.926e-05 5.004e-06	$ 3.56 \\ 3.54 \\ 3.79 $	4.014e-03 1.031e-04 8.751e-07 1.640e-08	$5.28 \\ 6.88 \\ 5.74$	1.515e-03 7.952e-06 4.969e-08 1.860e-10	 7.57 7.32 8.06	3.188e-04 2.349e-06 2.693e-09 1.244e-12	7.089.7711.08

TABLE 3. - Results for OGAMs of order 4, 6, 8, 10 on problem (21).

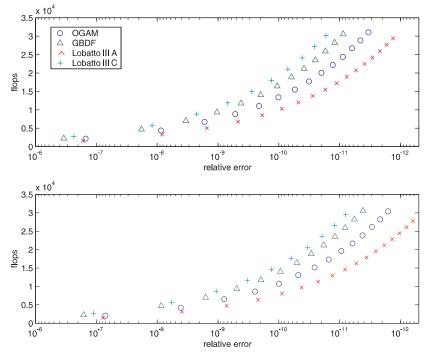


Fig. 2. - Problem (22) (above) and Problem (24) (below) in the transient phase.

the truncation error

$$\tau_n = c_{p+1} y^{(p+1)}(t_n) h^{p+1} + O(h^{p+2}),$$

where p is the order of the method and c_{p+1} a coefficient depending only on the method. As an example, one can see in Table 4 that the error constant for OGAM is smaller than that for GBDF.

Similar conclusions can be obtained by applying the same methods on the following linear problem:

(24)
$$\boldsymbol{y}' = \boldsymbol{d} \boldsymbol{y}, \quad \boldsymbol{y}(0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

TABLE 4. - Error constants of OGAMs and GBDF.

p	4	6		
method	OGAM GBDF	OGAM GBDF		
c_{p+1}	$-\frac{19}{5!6}$ $\frac{36}{5!6}$	$\frac{271}{7!12} - \frac{576}{7!12}$		

TABLE 5. - Results for the 3-step OGAM and the 3-stage Lobatto IIIA, IIIC methods on problem (25).

h	OGAM		Lobatto I	IIA	Lobatto IIIC		
	error	rate	error	rate	error	rate	
1e-1 5e-2 25e-3 125e-4	8.144e-12 4.683e-13 2.764e-14 1.988e-15		9.575e-11 2.391e-11 5.946e-12 1.457e-12	 2.00 2.01 2.03	8.315e-10 2.082e-10 5.206e-11 1.301e-11	 2.00 2.00 2.00	

where

$$\mathbf{d} = \begin{pmatrix} 0 & 1 \\ -10001 & -10000 \end{pmatrix},$$

with eigenvalues $\lambda_1 \approx -1$, $\lambda_2 \approx -10^4$ (see Figure 2 below).

Consider now the Prothero-Robinson problem [10]

(25)
$$y' = \lambda(y - \varphi(t)) + \dot{\varphi}(t), \qquad y(t_0) = \varphi(t_0), \qquad \operatorname{Re} \lambda \leq 0,$$

with $\varphi(t) = \sin t$, $\lambda = -10^6$, $t \in [0, 1]$.

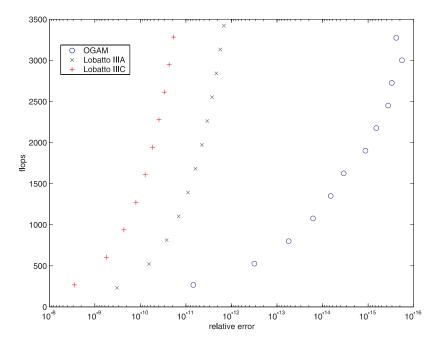


Fig. 3. - Prothero-Robinson problem.

LIDIA ACETO

We approximate such problem by using the 3-stage Lobatto IIIA, IIIC methods and the 3-step OGAM, with constant stepsize h. Under the assumption that simultaneously $h \rightarrow 0$ and $\lambda h \rightarrow \infty$, we observe from Table 5 that the order of convergence for the implicit Runge-Kutta methods (Lobatto IIIA and Lobatto IIIC methods) is considerably smaller than the classical order. On the contrary, it is evident that OGAM does not suffer of loss of accuracy. Moreover, from Figure 3 it turns out that, for a fixed cost, the relative error for OGAM is several order of magnitude less than the other methods.

5. - Conclusions.

Because of their good stability properties, OGAMs appear to be attractive for the numerical solution of stiff problems. The numerical examples confirm their potentiality. Therefore, an extension of code GAM [8], based also on OGAMs, could be considered.

Acknowledgements. The author wishes to thank Donato Trigiante for his helpful comments and constructive criticism during the preparation of this paper.

REFERENCES

- L. ACETO, On the Stability Problem arising in Numerical Methods for ODEs, Ph. D. Thesis, Genova, 2001.
- [2] L. ACETO D. TRIGIANTE, On the A-stable methods in the GBDF class, Nonlinear Analysis: Real World Applications, 3 (2002), 9-23.
- [3] P. AMODIO F. MAZZIA, A Boundary Value Approach to the Numerical Solution of Initial Value Problems by Multistep Methods, J. Difference Eq. Appl., 1 (1995), 353-367.
- [4] P. AMODIO F. MAZZIA, Boundary Value Methods based on Adams-type methods, Appl. Num. Math., 18 (1995), 23-35.
- [5] L. BRUGNANO D. TRIGIANTE, Boundary Value Method: the Third Way Between Linear Multistep and Runge-Kutta Methods, Comput. Math. Appl., 36 (1998), 269-284.
- [6] L. BRUGNANO D. TRIGIANTE, Solving Differential Problems by Multistep Initial and Boundary Value Methods, Gordon and Breach Science Publishers, Amsterdam, 1998.
- [7] E. HAIRER G. WANNER, Solving Ordinary Differential Equations II, Springer Series in Computational Mathematics, vol.14, Springer-Verlag, Berlin, 1991.
- [8] F. IAVERNARO F. MAZZIA, GAM, August 1997. Available via www at URL http://www.dm.uniba.it/~mazzia/ode/readme.html

- [9] F. IAVERNARO F. MAZZIA, Solving Ordinary Differential Equations by Generalized Adams Methods: properties and implementation techniques, Appl. Num. Math., 28 (1998), 107-126.
- [10] A. PROTHERO A. ROBINSON, On the Stability and Accuracy of One-Step Methods for Solving Stiff Systems of Ordinary Differential Equations, Math. of Comput., 28 (1974), 145-162.
- [11] L. F. SHAMPINE, Numerical Solution of Ordinary Differential Equations, Chapman & Hall, New York, 1994.

Dipartimento di Matematica Applicata «U. Dini», Università di Pisa Via Diotisalvi, 2 - 56126 Pisa (Italy). E-mail: l.aceto@dma.unipi.it

Pervenuta in Redazione il 28 giugno 2004