ELSEVIER

Applied Numerical Mathematics 18 (1995) 397–411

APPLIED NUMERICAL MATHEMATICS

Convergence aspects of step-parallel iteration of Runge–Kutta methods ☆

W.A. van der Veen *, J.J.B. de Swart, P.J. van der Houwen

CWI, P.O. Box 94079, 1090 GB Amsterdam, Netherlands

Received 19 September 1994; revised 6 December 1994; accepted 6 December 1994

Abstract

One of the most powerful methods for solving initial value problems for ordinary differential equations is an implicit Runge–Kutta method such as the Radau IIA methods. These methods are both highly accurate and highly stable. However, the iterative scheme needed for solving the implicit RK equations requires a lot of computational effort. The arrival of parallel computer systems has changed the situation in the sense that the effective computational effort can be reduced to a large extent. One option is the application of the iteration scheme concurrently at a number of step points on the t-axis. In this paper, we shall analyse the convergence of a special class of such step-parallel iteration methods.

Keywords: Numerical analysis; Runge–Kutta methods; Parallelism; Convergence factors

1. Introduction

We consider parallel methods for solving d-dimensional initial value problems (IVPs):

$$y'(t) = f(y(t)), \quad y(t_0) = y_0, \qquad y, f \in \mathbb{R}^d. \tag{1.1}$$

One of the most powerful methods for solving this IVP is an implicit Runge–Kutta (RK) method such as the Radau IIA methods. These methods are L-stable and have order $p = 2s - 1$, s being the number of stages. However, the iterative scheme needed for solving the implicit RK equations requires a lot of computational effort. Because of this, implicit RK methods have never been popular on sequential computers. Parallel computer systems have changed the situation, and various attempts have been made to develop parallel iteration

☆ The research reported in this paper was partly supported by the Technology Foundation (STW) in The Netherlands.

* Corresponding author.

SSDI 0168-9274(95)00063-1

schemes for solving the implicit RK equations. We mention the work of Jackson and Nørsett [6], Lie [7], Bellen et al. [1,2] and Chartier [4]. Also at CWI, parallel iteration schemes have been investigated. For stiff problems, we applied Newton-type iteration in which the *sd*-by-*sd* Jacobian matrix of the implicit equations was approximated by a block diagonal matrix with *d*-by-*d* blocks (cf. [10]). The sequential (or effective) costs per iteration of the resulting "simplified" Newton iteration method are reduced to solving s linear systems of dimension d in parallel. This iteration method was called the PDIRK iteration method (Parallel Diagonal-implicit Iterated RK method). Following the ideas of Bellen and co-workers, a further level of parallelism was introduced in [11–13] by applying the PDIRK iteration scheme concurrently at a number of step points on the t-axis. In this paper, we shall analyse the convergence of these step-parallel PDIRK methods.

2. The iteration scheme

Our starting point is the same corrector formula as in [12]. Using the General Linear Method notation of Butcher, the corrector formula reads (cf. [3,5])

$$Y_n = (E \otimes I)Y_{n-1} + h_n(A \otimes I)F(Y_n), \quad n = 1, \ldots, N. \tag{2.1a}$$

Here, h_n denotes the stepsize $t_n - t_{n-1}$, the s-by-s matrices A and E contain the method parameters, and $F(Y_n)$ contains the derivative values $(f(Y_{n,i}))$, where $Y_{n,i}$, $i = 1, 2, \ldots, s$, denote the d-dimensional components of the stage vector Y_n. In this paper we will assume that (2.1a) possesses s implicit stages and that the last stage corresponds to the step point t_n (e.g. Radau IIA-type methods). The s components $Y_{n,i}$ represent numerical approximations at the intermediate points $t_{n-1} + c_i h_n$, $i = 1, \ldots, s$, where $c = (c_i) = Ae$, e being the vector with unit entries. Furthermore, the matrix I is the d-by-d identity matrix, $\otimes$ denotes the Kronecker product, and we define $Y_0 = e \otimes y_0$. The dimensions of I and e may change, but will always be clear from the context.

Confining our considerations to RK methods, the matrix E in (2.1a) is of the form

$$E := \begin{pmatrix} 0 & \ldots & 0 & 1 \\ \vdots & & \vdots & \vdots \\ 0 & \ldots & 0 & 1 \end{pmatrix}. \tag{2.1b}$$

However, most of our analysis applies to the case of a General Linear Method where E is more general.

We approximate the solution Y_n of (2.1) by successive iterates $Y_n^{(j)}$ satisfying the iteration scheme

$$\begin{aligned} &Y_n^{(1)} \text{ to be defined by the predictor formula,} \\ &Y_n^{(j)} - h_n(B \otimes I)F\big(Y_n^{(j)}\big) = (E \otimes I)Y_{n-1}^{(j+j^*-1)} + h_n\big((A - B) \otimes I\big)F\big(Y_n^{(j-1)}\big), \\ &\quad j = 2, \ldots, m, \\ &Y_n^{(j)} = Y_n^{(m)}, \quad j > m, \end{aligned} \tag{2.2}$$

where $n = 1, 2, \ldots, N$, B is an s-by-s matrix, $Y_0^{(j)} = e \otimes y_0$ for all j, and j^* is an integer greater than or equal to 1. It will be assumed that the sequential costs of applying the predictor formula and the correction formula are comparable.

Irrespective the choice of the matrix B, the iteration scheme (2.2) possesses *parallelism across the steps*. For instance, if j^* is constant, then (2.2) shows that for a given $j \geqslant 1$, the iterates $\{Y_n^{(j)}, Y_{n-1}^{(j+j^*)}, Y_{n-2}^{(j+2j^*)}, \ldots, Y_1^{(j+nj^*-j^*)}\}$ can be computed concurrently. The sequential (or effective) costs consists of $N_{\text{seq}} := m + (N-1)j^*$ applications of the correction formula (if m depends on the step number n, then m is understood to be the number of iterations at the endpoint). Notice that for $j^* = m$, the iteration method (2.2) reduces to the conventional iteration strategy without step parallelism.

The matrix B defines the iteration method within a single step and therefore plays a crucial role in the degree of parallelism within the steps. There are several options for choosing the matrix B. For example, the case $B = O$ (fixed point iteration) was studied in [11] and the resulting method was called the PIRKAS method (Parallel Iterated RK Across the Steps). In addition to parallelism across the *steps*, PIRKAS methods also have parallelism across the *components of the iterates*, because all components of $F(Y_n^{(j-1)})$ can also be evaluated in parallel. Methods where B is a diagonal matrix D with positive diagonal entries minimizing the spectral radius of the matrix $I - D^{-1}A$ (such matrices can be found in [10]) were applied in [12] and were called PDIRKAS methods (Parallel Diagonal-implicitly Iterated RK Across the Steps). These methods are implicit because we have to solve nonlinear relations in each iteration. But the diagonal structure of B enables us to solve the components of $Y_n^{(j)}$ in parallel. Hence, we again have both parallelism across the *steps* and across the *components of the iterates*.

In an actual implementation, the number of iterations m performed at t_n and the parameter j^* are defined dynamically. The value of m is determined by the condition that for $j = m$ the iterates $Y_n^{(j)}$ satisfy the corrector equation (2.1) within a given tolerance. The value of j^* turns out to be decisive for the overall performance of the iteration process. It should be sufficiently large in order to have satisfactory convergence at t_n. Hence, both m and j^* may depend on t_n. In a theoretical analysis, however, it seems not feasible to allow the parameters m and j^* to be arbitrary functions of n, so that in deriving convergence results, m and j^* are assumed to be constant. In our first investigations of step-parallel iterations schemes in [11,12], we hoped that sufficient robustness could already be obtained for $j^* = 1$. We therefore analysed convergence only for $j^* = 1$. However, our numerical experiments have shown that j^* is at best 2 or 3. In this paper, we extend our earlier analysis to the case where j^* is allowed to be greater than 1.

3. Stability and convergence

Assuming that the corrector equation (2.1) is unconditionally stable and that the corrector equation is solved within a given tolerance, the method (2.2) will be stable whenever it is convergent. We shall discuss convergence for the familiar basic test equation $y'(t) = \lambda y(t)$, where λ is assumed to run through the spectrum of $\partial f/\partial y$. Furthermore, we assume h, m and j^* independent of n. When applied to the test equation, the iteration scheme assumes the form

$Y_n^{(1)}$ to be defined by the predictor formula,

$$Y_n^{(j)} = \begin{cases} KY_{n-1}^{(j+j^*-1)} + ZY_n^{(j-1)}, & j = 2, \ldots, m, \\ Y_n^{(m)}, & j > m, \end{cases} \tag{3.1}$$

where

$$K := (I - zB)^{-1}E, \quad Z := z(I - zB)^{-1}(A - B), \qquad z := \lambda h.$$

In [11,12] we discussed convergence of (3.1) for the case $j^* = 1$. In this paper, we shall allow j^* to be greater than 1. As already observed in [11], the convergence analysis of (2.2) cannot be restricted to a local analysis of the iteration errors at a fixed point t_n, but should be a global analysis where iteration errors at all preceding step points are involved. We shall distinguish two situations: (i) the predictor is based on iterates generated by the iteration scheme (3.1), and (ii) the iterates $Y_n^{(1)}$ are generated independently, that is, the predictor is completely independent of the iteration scheme. In the first situation, it is required that the predictor formula is explicitly given (to be referred to as the *given-predictor* case). In the second case, the predictor formula itself is not used in deriving the convergence conditions and may therefore have any form (the *independent-predictor* case). However, in the case of large integration intervals where n becomes large, the predictor formula should be sufficiently stable in order to generate useful first iterates. In fact, for large n, the region of convergence of (3.1) will be limited by the stability region of the predictor. In the given-predictor case, we confine our considerations to predictor formulas of the form

$$Y_n^{(1)} = P\big(Y_{n-1}^{(j^*)}\big). \tag{3.2}$$

For the test equation, the predictor formula (3.2) takes the form $Y_n^{(1)} = PY_{n-1}^{(j^*)}$, where $P = P(z)$ is an s-by-s matrix, to be called the *predictor matrix*. Thus, the step-parallel iteration method can be characterized by the matrices K, Z and P (if the predictor formula is explicitly specified).

In order to analyse convergence, we derive a relation between the vectors of iterates at t_n and t_{n-1}. Repeated application of the recursion (3.1) yields

$$Y_n^{(j)} = \begin{cases} \displaystyle\sum_{k=1}^{j-1} Z^{k-1}KY_{n-1}^{(j+j^*-k)} + Z^{j-1}Y_n^{(1)}, & j = 2,\dots,m, \\ Y_n^{(m)}, & j > m. \end{cases} \tag{3.3}$$

Let $\theta = 0$ and $\theta = 1$ respectively refer to the independent-predictor and given-predictor cases introduced above. Then the recursion (3.3) can be written in the compact form

$$V_n = SV_{n-1} + (1-\theta)CY_n^{(1)}, \quad V_n := \begin{pmatrix} Y_n^{(j^*)} \\ Y_n^{(j^*+1)} \\ Y_n^{(j^*+2)} \\ \vdots \\ Y_n^{(m+j^*-1)} \end{pmatrix}, \tag{3.4}$$

$$S := \begin{pmatrix} \theta Z^{j^*-1}P & Z^{j^*-2}K & . & \dots & K & O & . & \dots & O \\ \theta Z^{j^*}P & Z^{j^*-1}K & Z^{j^*-2}K & \dots & . & K & O & \dots & O \\ \vdots & \vdots & \vdots & & & & & & \vdots \\ \theta Z^{m-1}P & Z^{m-2}K & Z^{m-3}K & \dots & . & . & . & \dots & K \\ \vdots & \vdots & \vdots & & & & & & \vdots \\ \theta Z^{m-1}P & Z^{m-2}K & Z^{m-3}K & \dots & . & . & . & \dots & K \end{pmatrix}, \quad C := \begin{pmatrix} Z^{j^*-1} \\ Z^{j^*} \\ \vdots \\ Z^{m-1} \\ \vdots \\ Z^{m-1} \end{pmatrix}.$$

S and C are an m-by-m and an m-by-1 block matrix, respectively. In both matrices, the last j^* block rows are identical. If $j^* = 1$, then the first block row of S reduces to $(\theta P\ O \ldots O)$, so that S becomes a lower triangular block matrix.

If $\theta = 0$, then the predictor matrix P is ignored in (3.4). Instead, the predictor values $Y_n^{(1)}$, $n = 1, 2, \ldots$, are involved. These values may be any sequence of initial iterates. If $\theta = 1$, then the special form of the predictor formula $Y_n^{(1)} = PY_{n-1}^{(j^*)}$ is taken into account.

3.1. The iteration error

In the conventional iteration process where $j^* = m$, we can derive a relation between the iteration error $\varepsilon_n^{(j)} := Y_n^{(j)} - Y_n$ and $\varepsilon_n^{(j-1)}$. However, if $j^* < m$, then this is no longer possible. In [11,12] it was shown that if, and only if, $j^* = 1$, then there exists a relation between the set of iteration errors $\varepsilon^{(j)} := (\varepsilon_1^{(j)}, \varepsilon_2^{(j)}, \ldots, \varepsilon_n^{(j)})$ and $\varepsilon^{(j-1)}$. For $j^* > 1$, (3.4) allows us to express $\varepsilon_n^{(j)}$ in terms of the predictor errors introduced at the points $t_1, t_2, \ldots, t_n$. Thus, given the predictor errors, we can get insight into the effect of the parameters j^*, m and n on the iteration errors.

Let us introduce the ms-dimensional vector $U_n := e \otimes Y_n$, where Y_n denotes the solution of (2.1). Then, we may define the *stage vector iteration errors*

$$\varepsilon_n := V_n - U_n, \qquad \varepsilon_n^{(j)} := Y_n^{(j)} - Y_n, \quad j = j^*, \ldots, m + j^* - 1, \tag{3.5}$$

and the *predictor error* vector Δ_n

$$\Delta_n := \begin{pmatrix} \delta_n \\ \delta_{n-1} \\ \vdots \\ \delta_1 \end{pmatrix}, \qquad \delta_n := \theta PY_{n-1} + (1-\theta)Y_n^{(1)} - Y_n, \quad \theta = 0, 1. \tag{3.6}$$

Theorem 3.1. *Let the block rows of the matrix S^kC be denoted by $[S^kC]^{(j)}$, $j = j^*, j^* + 1, \ldots, m + j^* - 1$. Then, for any vector of predictor errors Δ_n, the iteration errors $\varepsilon_n^{(j)}$, $j = j^*, \ldots, m$, are given by*

$$\varepsilon_n^{(j)} = \Sigma_\theta^{(j)} \Delta_n, \quad \Sigma_\theta^{(j)} := \left([C]^{(j)} [SC]^{(j)} [S^2C]^{(j)} \ldots [S^{n-1}C]^{(j)} \right). \tag{3.7}$$

If $\theta = 1$, then the error equation can be written as

$$\varepsilon_n^{(j)} = \tilde{\Sigma}_1^{(j)} Y_0, \quad \tilde{\Sigma}_1^{(j)} := \sum_{k=0}^{n-1} [S^kC]^{(j)} \left(P - (I - zA)^{-1}E \right) \left((I - zA)^{-1}E \right)^{n-k-1}. \tag{3.7'}$$

Proof. On substitution of $Y_n^{(j)} = Y_n + \varepsilon_n^{(j)}$ into (3.4), we obtain

$$\begin{aligned}
\varepsilon_n^{(j)} &= [S]^{(j)} \varepsilon_{n-1} + [S]^{(j)} U_{n-1} + (1-\theta) Z^{j-1} Y_n^{(1)} - Y_n \\
&= [S]^{(j)} \varepsilon_{n-1} + \left[\theta Z^{j-1} P + \sum_{k=1}^{j-1} Z^{k-1} K \right] Y_{n-1} + (1-\theta) Z^{j-1} Y_n^{(1)} - Y_n \\
&= [S]^{(j)} \varepsilon_{n-1} + \theta Z^{j-1} PY_{n-1} + (I - Z^{j-1})(I - Z)^{-1} KY_{n-1} + (1-\theta) Z^{j-1} Y_n^{(1)} - Y_n \\
&= [S]^{(j)} \varepsilon_{n-1} + Z^{j-1} \left[\theta PY_{n-1} + (1-\theta) Y_n^{(1)} - Y_n \right] = [S]^{(j)} \varepsilon_{n-1} + Z^{j-1} \delta_n,
\end{aligned}$$

where $j = j^*, \ldots, m + j^* - 1$. Hence, $\varepsilon_n = S\varepsilon_{n-1} + C\delta_n$, and repeated application yields

$$\varepsilon_n = \sum_{k=0}^{n-1} S^k C \delta_{n-k}, \qquad \varepsilon_n^{(j)} = \sum_{k=0}^{n-1} \left[S^k C\right]^{(j)} \delta_{n-k}, \quad j = j^*, \ldots, m + j^* - 1. \tag{3.8}$$

This leads to the representation (3.7). If $\theta = 1$, then it follows from (3.6) and (2.1a) that

$$\delta_n := PY_{n-1} - Y_n = \left(P - (I - zA)^{-1}E\right)\left((I - zA)^{-1}E\right)^{n-1} Y_0, \tag{3.6'}$$

and substitution into (3.8) yields the result (3.7'). □

The s-by-s matrix $[S^k C]^{(j)}$ in (3.7) determines the amplification of the predictor error δ_{n-k} at the point t_{n-k}, and the accumulated amplification is determined by the matrix $\Sigma_\theta^{(j)}$. This amplification matrix, and therefore also $\varepsilon_n^{(j)}$, depends not only on j, n and on the variable z, but also on the parameters j^* and m. In general, $\varepsilon_n^{(j)}$ will decrease in magnitude as j^* and m increase. However, if for given j, n and j^*, the value of m becomes greater than $j + (n-1)(j^* - 1)$, then $\varepsilon_n^{(j)}$ does not depend on m anymore. The result (3.7') takes the predictor formula into account, but again the matrix $[S^k C]^{(j)}$ plays a crucial role in the amplification matrix $\tilde{\Sigma}_1^{(j)}$.

Let us assume that the predictor error vector Δ_n is bounded. Then, with respect to a norm $\|\cdot\|$ and for given values of j^* and n, the region of convergence associated with (3.7) is defined by the set

$$\mathbb{C}_\theta(n, j^*) := \left\{z : \|\Sigma_\theta^{(j)}\| \to 0 \text{ as } m = j \to \infty\right\}. \tag{3.9}$$

Similarly, the region of convergence associated with (3.7') is defined by

$$\tilde{\mathbb{C}}_1(n, j^*) := \left\{z : \|\tilde{\Sigma}_1^{(j)}\| \to 0 \text{ as } m = j \to \infty\right\}. \tag{3.9'}$$

Furthermore, adapting a definition given in [9, p. 88], we define for (3.7) and (3.7') the averaged *speed (or rate) of convergence* by

$$R_\theta(n, j^*, j, m, z) := -\frac{1}{j}\log\|\Sigma_\theta^{(j)}\|, \qquad \tilde{R}_1(n, j^*, j, m, z) := -\frac{1}{j}\log\|\tilde{\Sigma}_1^{(j)}\|. \tag{3.10}$$

3.2. A convergence theorem for $\theta = 0$

In order to determine the region of convergence and to get insight into the speed of convergence, we first derive an upper bound for $\|[S^k C]^{(j)}\|$. We confine our considerations to the case of arbitrary predictor formulas (i.e. $\theta = 0$). In deriving the upper bound for $\|[S^k C]^{(j)}\|$, an important tool is provided by the ε-pseudo-spectra of matrices which are defined as follows (see e.g. Reichel and Trefethen [8]):

Definition 3.2. Let $\|\cdot\|_2$ denote the 2-norm and let $\varepsilon > 0$. Then, (i) μ is an *ε-pseudo-eigenvalue* of the matrix M if $\|(\mu I - M)^{-1}\|_2 \geqslant \varepsilon^{-1}$, (ii) $\Lambda_\varepsilon(M)$ is the *ε-pseudo-spectrum* of M if it contains all ϵ-pseudo-eigenvalues of M, and (iii) $\rho_\varepsilon(M)$ is the *ε-pseudo-spectral radius* of M if it equals the maximal modulus of the points in $\Lambda_\varepsilon(M)$.

Obviously, for any positive ε, all eigenvalues of M are included in the ε-pseudo-spectrum of M. Furthermore, for a given matrix M and parameter ε, and sufficiently large values of $|\mu|$, we have that $\|(\mu I - M)^{-1}\|_2 = |\mu^{-1}|(1 + O(\mu^{-1})) < \varepsilon^{-1}$. Hence, the ε-pseudo-spectrum of M constitutes a finite set in the complex plane.

Theorem 3.3. *Let $\Lambda_\varepsilon(Z)$ and $\rho_\varepsilon(Z)$ respectively denote the ε-pseudo-spectrum and the ε-pseudo-spectral radius of Z, let $L_\varepsilon(Z)$ denote the length of the boundary $\partial\Lambda_\varepsilon(Z)$ of $\Lambda_\varepsilon(Z)$, and define*

$$\Gamma_\varepsilon(Z) := \frac{\varepsilon}{(\rho_\varepsilon(Z))^{j^*+1}} \max_{\partial\Lambda_\varepsilon(Z)} \|(\zeta I - Z)^{-1} Z^{j^*}\|_2,$$

$$\gamma_\varepsilon(Z) := \frac{1}{\varepsilon}\|K\|_2 (\rho_\varepsilon(Z))^{j^*}.$$

If $j \geqslant j^* + 1$, $m \to \infty$ and $\rho(Z) < 1$, then for any $\varepsilon > 0$ we have that

$$\|[S^k C]^{(j)}\|_2 \leqslant \frac{L_\varepsilon(Z)}{2\pi\varepsilon} \Gamma_\varepsilon(Z)(\rho_\varepsilon(Z))^j (\gamma_\varepsilon(Z))^k, \tag{3.11}$$

$$\|\Sigma_0^{(j)}\|_2 \leqslant \frac{L_\varepsilon(Z)}{2\pi\varepsilon} \Gamma_\varepsilon(Z)(\rho_\varepsilon(Z))^j \frac{1 - (\gamma_\varepsilon(Z))^n}{1 - \gamma_\varepsilon(Z)}. \tag{3.12}$$

Proof. The result (3.11) can be proved by means of convolution properties of the Fourier transform. Because the proof is rather lengthy, it is given in Appendix A to this paper. The estimate (3.12) directly follows from (3.11) by writing

$$\|\Sigma_0^{(j)}\|_2 \leqslant \sum_{k=0}^{n-1} \|[S^k C]^{(j)}\|_2 \leqslant \frac{L_\varepsilon(Z)}{2\pi\varepsilon} \Gamma_\varepsilon(Z)(\rho_\varepsilon(Z))^j \sum_{k=0}^{n-1} (\gamma_\varepsilon(Z))^k. \qquad \square$$

The ε-pseudo-spectral radius $\rho_\varepsilon(Z)$ is continuous in ε and monotonically decreasing to $\rho(Z)$ as $\varepsilon \to 0$. Since $L_\varepsilon(Z)/2\pi\varepsilon$ is bounded for all ε and because there is always an ε with $\rho_\varepsilon(Z) < 1$ (provided that $\rho(Z) < 1$), we conclude that, for fixed n, $\|\Sigma_0(j)\|_2$ converges to 0 as j increases. Thus, as a first corollary of Theorem 3.3 we have:

Corollary 3.4. *If the conditions of Theorem 3.3 are satisfied, if $\theta = 0$ and n is finite, then for all j^* the convergence region of the PDIRKAS method is given by $\mathbb{C}_0(n, j^*) := \{z: \rho(Z(z)) < 1\}$.*

Furthermore, it follows from Theorem 3.3 that, given the values of j^* and n, the speed of convergence is bounded below according to an inequality of the form

$$R_0(n, j^*, j, m, z) \geqslant -\log \rho_\varepsilon(Z(z)) - \frac{a_\varepsilon(n, j^*, z)}{j}, \tag{3.13}$$

where $a_\varepsilon(n, j^*, z)$ does not depend on j. This estimate illustrates the crucial role played by the ε-pseudo-spectral radius of the matrix $Z(z)$.

Table 1
Values of $\rho_\varepsilon(Z)$, $\Gamma_\varepsilon(Z)$ and $\gamma_\varepsilon(Z)$ for the four-stage Radau IIA method at $z = 10\mathrm{i}$

	j^*	$\varepsilon = 1.00$	$\varepsilon = 0.25$	$\varepsilon = 0.1$	$\varepsilon = 0.05$	$\varepsilon = 0.001$
$\rho_\varepsilon(Z)$		2.478	1.451	1.080	0.903	0.542
$\Gamma_\varepsilon(Z)$	2	0.177	0.503	0.823	1.063	1.910
$\Gamma_\varepsilon(Z)$	3	0.053	0.268	0.558	0.807	1.899
$\Gamma_\varepsilon(Z)$	4	0.014	0.114	0.305	0.516	1.852
$\gamma_\varepsilon(Z)$	2	06.04	08.28	11.47	16.04	289.0
$\gamma_\varepsilon(Z)$	3	14.97	12.02	12.39	14.49	156.60
$\gamma_\varepsilon(Z)$	4	37.09	17.44	13.39	13.08	084.89

In order to see the effect of the quantities $\rho_\varepsilon(Z)$, $\Gamma_\varepsilon(Z)$ and $\gamma_\varepsilon(Z)$ on the convergence, we have listed their values in Table 1, for the four-stage Radau IIA corrector with matrix $B = D$ as defined in [10] with Z evaluated at the point $z = 10\mathrm{i}$ (this point is in the neighbourhood where experimentally the convergence speed is minimal). These figures together with the estimate (3.12) indicate that for larger values of j and n, the convergence behaviour is largely determined by the factor $(\rho_\varepsilon(Z))^j(\gamma_\varepsilon(Z))^{n-1}$. Hence, given the value of j^*, roughly the same reduction factor is obtained if jn^{-1} is constant.

3.3. *Stiff and nonstiff convergence for* $\theta = 0$

In this section, we consider the convergence in the neighbourhood of the origin (*nonstiff* convergence) and at infinity (*stiff* convergence). For the nonstiff convergence, it is convenient to have an alternative representation for the inequality (3.12). As a second corollary of Theorem 3.3 we have:

Corollary 3.5. *Let the conditions of Theorem* 3.3 *be satisfied and define the matrix*

$$Z_0 := z^{-1}Z = (I - zB)^{-1}(A - B).$$

Then (3.12) *can be represented in the form*

$$\| \Sigma_0^{(j)} \|_2 \leqslant \frac{L_\varepsilon(Z_0)}{2\pi\varepsilon} |z|^{j-1}\Gamma_\varepsilon(Z_0)(\rho_\varepsilon(Z_0))^j \frac{1 - |z|^{(j^*-1)n}(\gamma_\varepsilon(Z_0))^n}{1 - |z|^{j^*-1}\gamma_\varepsilon(Z_0)}. \tag{3.12'}$$

Let j and j^* be fixed with $j^* > 1$. Then, the nonstiff convergence factor is uniformly bounded for all n.

Proof. From the definition of the ε-pseudo-spectral radius it is easily seen that for any matrix M and any constant α, the relation $\rho_\varepsilon(\alpha M) = |\alpha| \rho_\delta(M)$, where $\delta = |\alpha|^{-1}\varepsilon$. Since $Z = zZ_0$, we have

$$\rho_\varepsilon(Z) = |z| \rho_\delta(Z_0),\ \Gamma_\varepsilon(Z) = |z|^{-1}\Gamma_\delta(Z_0),$$

$$\gamma_\varepsilon(Z) = |z|^{j^*-1}\gamma_\delta(Z_0),\ \frac{L_\varepsilon(Z)}{2\pi\varepsilon} = \frac{L_\delta(Z_0)}{2\pi\delta},$$

where $\delta := |z|^{-1}\varepsilon$. Hence,

$$\|[S^k C]^{(j)}\|_2 \leqslant \frac{L_\delta(Z_0)}{2\pi\delta}|z|^{j-1}\Gamma_\delta(Z_0)(\rho_\delta(Z_0))^j\left(|z|^{j^*-1}\gamma_\delta(Z_0)\right)^k.$$

Because this inequality holds for any positive δ, we may replace δ with ε to obtain (3.12′).

For $j^* > 1$, (3.12′) implies

$$\|\Sigma_0^{(j)}\|_2 \leqslant \frac{L_\varepsilon(Z_0)}{2\pi\varepsilon}|z|^{j-1}\Gamma_\varepsilon(Z_0)(\rho_\varepsilon(Z_0))^j(1+O(z^{j^*-1})) \quad \text{as } z\to 0.$$

Thus, this bound on the nonstiff convergence factor does not anymore depend on n. □

We remark that for $j^* = 1$, $\gamma_\varepsilon(Z_0) = \varepsilon^{-1}\|K\|_2\rho_\varepsilon(Z_0)$. Hence, unless we can find an ε such that $\gamma_\varepsilon(Z_0) < 1$, the bound on $\|\Sigma_0^{(j)}\|_2$ will increase exponentially with n. Since $\|K\|_2 \to 1$ and $Z_0 \to A - B$ as $z \to 0$, we obtain the condition $\rho_\varepsilon(A-B) < \varepsilon$ which is usually not fulfilled.

For the stiff convergence we have:

Corollary 3.6. *Let the conditions of Theorem* 3.3 *be satisfied, and let j and j* be fixed. Then, the stiff convergence factor is uniformly bounded for all n.*

Proof. It follows from Theorem 3.3 and the observation $\|K(z)\|_2 = O(z^{-1})$ as $z \to \infty$, that for all j^*

$$\|\Sigma_0^{(j)}\|_2 \leqslant \frac{L_\varepsilon(Z)}{2\pi\varepsilon}\Gamma_\varepsilon(Z(\infty))(\rho_\varepsilon(Z(\infty)))^j + O(z^{-1}) \quad \text{as } z\to\infty,$$

so that the stiff convergence factor is uniformly bounded in n. □

3.4. Minimal speed of convergence for $\theta = 0$

In order to see the effect of the value of n, j^*, m and j on the true speed of convergence as defined by (3.7) and (3.10), we have computed the minimal value of $R_0(n, j^*, j, m, z)$ in the left-hand z-plane. This value will be denoted by $R_0^*(n, j^*, j, m)$. Of course, $R_0^*(n, j^*, j, m)$ refers to a "worst-case" situation, and restricting z to special subregions (e.g. the negative axis) would lead to larger speeds of convergence. However, the qualitative behaviour would not be changed.

In particular, we consider the PC pair consisting of an unconditionally stable predictor and the four-stage Radau IIA corrector with matrix $B = D$ as in Table 1. Using the infinity norm, Table 2 lists $R_0^*(n, j^*, j, m)$ for a few values of n and j^* with $j = m = 32$ (we recall that for $j^* = m$, (2.2) reduces to the conventional iteration strategy without step parallelism). These

Table 2
Minimal convergence speeds for $j = m = 32$ and $\theta = 0$

n	$j^* = 1$	$j^* = 2$	$j^* = 3$	$j^* = 4$	$j^* = 32$
2	0.144	0.153	0.163	0.172	0.222
4	0.025	0.047	0.069	0.091	0.222
8	−0.144	−0.113	−0.075	−0.031	0.222

Table 3
Minimal convergence speeds for $n = 8$ and $\theta = 0$

$j = m$	$j^* = 1$	$j^* = 2$	$j^* = 3$	$j^* = 4$	$j^* = \mathrm{m}$
8	−0.605	−0.675	−0.600	−0.463	0.050
32	0.025	0.047	0.069	0.091	0.222
64	0.017	0.041	0.066	0.091	0.250

figures show the dramatic effect of n on the amplification factors. It is also clear that the n-effect is less as j^* is larger.

Next, we computed $R_0^*(n, j^*, j, m)$ as a function of m and j^* for n fixed with $j = m$. Table 3 lists results for $n = 8$. As expected, the performance improves as m increases.

From a practical point of view, we have to take into account the sequential costs when discussing the performance of the iteration process. Recalling that the sequential costs of iteration across n steps are measured by the value of $N_{\mathrm{seq}} = (n-1)j^* + m$, we see that large values of m are less alarming than they would be in conventional iteration processes with $j^* = m$, where the sequential costs after n steps are given by $N_{\mathrm{seq}} = nm$. As long as j^* is less than the number of iterations required by conventional iteration, across-the-steps iteration will be more efficient. We illustrate this for the case where $n = 8$ and N_{seq} is constant for all j^*. Table 4 lists values of $R_0^*(n, j^*, j, m)$ for $N_{\mathrm{seq}} = 96$. After a rapid increase until $j^* = 9$, the convergence speed starts to decrease because m becomes too small. We also listed the value of $jR_0^*(n, j^*, j, m)$ that may be considered as a measure of the efficiency of the iteration process after j iterations. Surprisingly, for $j^* \leqslant 9$, the efficiency hardly depends on j^*. Apparently, the decrease of the number of iterations m per step is fully compensated by the increase of j^*, until m becomes too small at $j^* = 10$.

3.5. *Minimal speed of convergence for* $\theta = 1$

Finally, we study the effect of including the predictor formula into the convergence analysis (the given-predictor case with $\theta = 1$). Two special predictor formulas are considered, viz. the *modified correction* formula

$$\begin{aligned} &\boldsymbol{Y}_n^{(1)} - h_n(B \otimes I)\boldsymbol{F}\big(\boldsymbol{Y}_n^{(1)}\big) = (E \otimes I)\boldsymbol{Y}_{n-1}^{(j^*)} + h_n((A-B) \otimes I)\boldsymbol{F}\big((E^* \otimes I)\boldsymbol{Y}_{n-1}^{(j^*)}\big), \\ &E^* = WX^{-1}, \qquad W := \left(\frac{1}{i}\boldsymbol{c}^i\right), \quad X := \big((\boldsymbol{c}-\boldsymbol{e})^{i-1}\big), \quad i = 1, \ldots, s, \end{aligned} \tag{3.14}$$

Table 4
Minimal convergence speeds and efficiency for $j = m$, $n = 8$, $N_{\mathrm{seq}} = 96$ and $\theta = 0$

	$j^* = 1$	$j^* = 2$	$j^* = 3$	...	$j^* = 7$	$j^* = 8$	$j^* = 9$	$j^* = 10$	$j^* = 11$	$j^* = 12$
$j = m = 96 - 7j^*$	89	82	75	...	47	40	33	26	19	12
$R_0^*(n, j^*, j, m)$	0.075	0.082	0.091	...	0.149	0.175	0.200	0.200	0.168	0.125
$jR_0^*(n, j^*, j, m)$	6.7	6.7	6.8	...	7.0	7.0	6.6	5.2	3.2	1.5

Table 5
Minimal convergence speeds for $j = m = 32$ and $n = 4$

Predictor	Error formula	$j^* = 1$	$j^* = 2$	$j^* = 3$	$j^* = 4$	...	$j^* = m$
(3.14)	$\tilde{R}_1^*(n, j^*, j, m, z)$	0.016	0.034	0.056	0.081	...	0.212
(3.15)	$\tilde{R}_1^*(n, j^*, j, m, z)$	0.041	0.063	0.084	0.106	...	0.238

and the *backward Euler* formula

$$Y_n^{(1)} - h(D^* \otimes I)F(Y_n^{(1)}) = (E \otimes I)Y_{n-1}^{(j^*)}, \quad D^* := \operatorname{diag}(c). \tag{3.15}$$

For (3.14) and (3.15), the matrix P occurring in the error formula (3.7′) is defined by $P = (I - zB)^{-1}(E + z(A - B)E^*)$ and $P = (I - zD^*)^{-1}E$, respectively. Again using the infinity norm, Table 5 lists the minimal convergence speed $\tilde{R}_1^*(n, j^*, j, m)$ in the left-hand z-plane. This table shows that, in spite of its low order, the backward Euler predictor is more effective than the high-order modified correction predictor. This indicates that the *stiff* iteration error components play a crucial role in the iteration process. Note that both cases show roughly the same increase of the convergence speed as j^* increases.

Acknowledgement

The authors are grateful to Prof. L.N. Trefethen for the interest he showed in our attempts to prove convergence for the PDIRKAS method. The use of ε-pseudo-spectral radius in the final proof of Theorem 3.3 was suggested by him when he visited CWI.

Appendix A. Proof of Theorem 3.3

The proof of Theorem 3.3 in this paper was given by the first author.

First, we show that in Theorem 3.3, $m \to \infty$ can be replaced by $m = \infty$, without sacrificing any generality. For $m = \infty$ the iteration process (3.2) is well defined. Furthermore, all the theory developed in the paper is still valid. Only the expression for the sequential costs should be reformulated. It is apparent from (3.2) that if $m = \infty$ and $\theta = 0$, $Y_n^{(j)}$ depends only on $Y_{n-1}^{(j^*+1)}, \ldots, Y_{n-1}^{(j^*-1+j)}$. Applying this repeatedly, we see that $Y_n^{(j)}$ depends only on $Y_1^{(j^*+1)}, \ldots, Y_1^{((n-1)(j^*-1)+j)}$. Therefore $Y_n^{(j)}$ is independent of m, provided that $m \geqslant (n-1)(j^* - 1) + j$. Hence, $[S^k C]^{(j)}$ with $m = \infty$ equals $[S^k C]^{(j)}$ with $m \geqslant (k-1)(j^* - 1) + j$.

The proof consists of two parts. First, $[S^k C]^{(j)}$ will be written as a line integral along the unit circle (see Lemma A.1). Thereafter, we obtain an upper bound for this integral (see Lemma A.2). In the proofs of these lemmas we shall use matrix Fourier analysis and more specifically the convolution property and the inverse Fourier transform.

Lemma A.1. *If* $\theta = 0$, $m = \infty$ *and* $\rho(Z) < 1$ *then, for* $j \geqslant j^* + 1$,

$$[S^k C]^{(j)} = \frac{1}{2\pi i} \oint_{|\zeta|=1} [R(Z, \zeta)K]^k R(Z, \zeta)\zeta^{(k-1)j^*+j-1} \, d\zeta \, Z^{j^*}$$

with $R(Z, \zeta) = (\zeta I - Z)^{-1}$.

Proof. In the case $\theta = 0$ the iterates $Y_n^{(j^*+1)}$, $Y_n^{(j^*+2)}, \ldots$, do not depend on $Y_{n-1}^{(j^*)}$, see Eq. (3.3). Therefore we redefine the block vectors V_n and C and block matrix S as follows:

$$V_n = \begin{pmatrix} Y_n^{(j^*+1)} \\ Y_n^{(j^*+2)} \\ \vdots \end{pmatrix}, \qquad C = \begin{pmatrix} Z^{j^*} \\ Z^{j^*+1} \\ \vdots \end{pmatrix}, \tag{A.1}$$

$$S = \begin{pmatrix} Z^{j^*-1}K & Z^{j^*-2}K & \ldots & K & 0 & . & . & \ldots \\ Z^{j^*}K & Z^{j^*-1}K & \ldots & ZK & K & 0 & . & \ldots \\ Z^{j^*+1}K & Z^{j^*}K & \ldots & Z^2K & ZK & K & 0 & \ldots \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \end{pmatrix}.$$

With these changes $V_n = SV_{n-1} + CY_n^{(1)}$ still holds. It can easily be seen that also Eq. (3.7) is still valid. The new matrix S is a so-called Toeplitz matrix. This is a matrix that represents a convolution. For showing that S is a convolution operator and for doing Fourier analysis we shall use as general argument for the operator S any V of the form

$$V = \begin{pmatrix} V^{(j^*+1)} \\ V^{(j^*+2)} \\ \vdots \end{pmatrix},$$

where $V^{(j)}$, $j = j^* + 1, j^* + 2, \ldots$, are matrices of order s. For simplifying the Fourier analysis, we introduce the notation: $V(j) = V^{(j^*+1+j)}$, $j = 0, 1, \ldots$. In particular, this notation will be applied to C, SC, $S^2C, \ldots$. Furthermore, these block vectors are considered as being sequences of s-by-s matrices, that is with the block vector V corresponds the sequence $\{V(j)\}_{j=0}^{\infty}$.

The new matrix S is of the form

$$\begin{pmatrix} A_0 & A_1 & A_2 & . & . & \ldots \\ A_{-1} & A_0 & A_1 & A_2 & . & \ldots \\ A_{-2} & A_{-1} & A_0 & A_1 & A_2 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix}.$$

In our case $A_j = Z^{j^*-1-j}K$ if $j \leqslant j^* - 1$ and $A_j = 0_{s*s}$ if $j > j^* - 1$. Though S is a double infinite matrix, S^kV is well defined for any sequence V of s-by-s matrices. This is because for any k and r, the rth row of S^k contains a finite number of nonzero entries. In terms of the sequence $\{A_j\}_{j=-\infty}^{\infty}$ the product SV can be written as

$$[SV](l) = \sum_{j=0}^{\infty} A_{-l+j}V(j) = \sum_{j=0}^{\infty} A_{-(l-j)}V(j), \quad l = 0, 1, \ldots. \tag{A.2}$$

Eq. (A.2) shows that S is a convolution operator: $SV = \{A_{-l}\} * V$. To obtain a formula for $[S^kC]^{(j)}$ we shall employ Fourier transforms along with their convolution property. The Fourier transform of an infinite sequence of matrices $\{A_l\}_{l=-\infty}^{\infty}$ is defined by

$$\mathcal{F}(A)(\omega) = \sum_{m=-\infty}^{\infty} A_m e^{-i\omega m},$$

and the convolution property gives:

$$\mathcal{F}(SV) = \mathcal{F}(\{A_{-l}\} * V) = \mathcal{F}(\{A_{-l}\})\mathcal{F}(V). \tag{A.3}$$

Let $H(\omega)$ be defined by

$$\begin{aligned} H(\omega) = \mathcal{F}(\{A_{-l}\})(\omega) &= \sum_{p=-\infty}^{\infty} A_{-p} e^{-i\omega p} \\ &= \sum_{p=-\infty}^{\infty} A_p e^{i\omega p} = \sum_{p=-\infty}^{j^*-1} Z^{j^*-1-p} K e^{i\omega p} \\ &= \sum_{p=0}^{\infty} Z^p K e^{-i\omega p} e^{i\omega(j^*-1)} = (I - e^{-i\omega} Z)^{-1} K e^{i\omega(j^*-1)}. \end{aligned}$$

The series $\Sigma_{p=0}^{\infty} Z^p e^{-i\omega p}$ is convergent because $\rho(e^{-i\omega} Z) = \rho(Z) < 1$. Using the convolution property (A.3) repeatedly, we can calculate $\mathcal{F}(S^k C)$ as follows

$$\begin{aligned} \mathcal{F}(S^k C)(\omega) &= \mathcal{F}(\{A_{-l}\} * \cdots * \{A_{-l}\} * C) \\ &= H^k(\omega)\mathcal{F}(C)(\omega). \end{aligned}$$

For the block vector C (see (A.1)), we have that $C(j) = Z^{j^*+j}$, $j = 0, 1 \ldots$. The Fourier transform of C is

$$\mathcal{F}(C) = \sum_{p=0}^{\infty} Z^p e^{-i\omega p} Z^{j^*} = Z^{j^*} (I - e^{-i\omega} Z)^{-1}.$$

Finally the Fourier transform of $S^k C$ is

$$\mathcal{F}(S^k C) = \left[(I - e^{-i\omega} Z)^{-1} K\right]^k (I - e^{-i\omega} Z)^{-1} Z^{j^*} e^{i\omega k(j^*-1)}.$$

Since we have been able to obtain an explicit expression for the Fourier transform of $S^k C$, the inverse Fourier transform can be used to calculate $[S^k C]^{(j)}$ as follows:

$$\begin{aligned} [S^k C]^{(j)} = [S^k C](j - j^* - 1) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathcal{F}(S^k C)(\omega) e^{i\omega(j-j^*-1)} \, d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[(I - e^{-i\omega} Z)^{-1} K\right]^k (I - e^{-i\omega} Z)^{-1} e^{i\omega[k(j^*-1)+j-j^*-1]} \, d\omega \, Z^{j^*}. \end{aligned} \tag{A.4}$$

In this formula and in all formulas below, it is assumed that $j \geqslant j^* + 1$. Substituting $\zeta = e^{i\omega}$, (A.4) can be written as

$$\begin{aligned} [S^k C]^{(j)} &= \frac{1}{2\pi i} \oint_{|\zeta|=1} \left[(I - \zeta^{-1} Z)^{-1} K\right]^k (I - \zeta^{-1} Z)^{-1} \zeta^{(k-1)j^*+j-k-2} \, d\zeta \, Z^{j^*} \\ &= \frac{1}{2\pi i} \oint_{|\zeta|=1} \left[(\zeta I - Z)^{-1} K\right]^k (\zeta I - Z)^{-1} \zeta^{(k-1)j^*+j-1} \, d\zeta \, Z^{j^*}, \end{aligned} \tag{A.5}$$

which proves Lemma A.1. □

Remark. The integral in (A.5) can be calculated by using the calculus of residues. This gives:

$$[S^k C]^{(j)} = \sum_{i_1 + \cdots + i_{k+1} = (k-1)j^* + j - k - 1} Z^{i_1} K \cdots Z^{i_k} K Z^{i_{k+1} + j^*}, \tag{A.6}$$

where the indices $i_1, \ldots, i_{k+1}$ assume all positive values as long as

$$i_1 + \ldots + i_{k+1} = (k-1)j^* + j - k - 1$$

is satisfied. It is difficult to find a sharp upper bound based on (A.6).

A suitable bound based on the previous lemma can be obtained by using the concept of the ε-pseudo-spectrum of a matrix, which was defined in Section 3.2. We shall prove:

Lemma A.2. *Let* $\Lambda_\varepsilon(Z)$, $\rho_\varepsilon(Z)$, $\Gamma_\varepsilon(Z)$ *and* $\gamma_\varepsilon(Z)$ *be defined as in Theorem* 3.3. *If* $\theta = 0$, $m = \infty$ *and* $\rho(Z) < 1$, *then for any* $j \geqslant j^* + 1$ *and* $\varepsilon > 0$

$$\| [S^k C]^{(j)} \|_2 \leqslant \frac{L_\varepsilon(Z)}{2\pi\varepsilon} \Gamma_\varepsilon(Z) (\rho_\varepsilon(Z))^j [\gamma_\varepsilon(Z)]^k.$$

Proof. The integrand of the integral in (A.5) is analytic outside $\Lambda_\varepsilon(Z)$ for any $\varepsilon > 0$. Therefore

$$[S^k C]^{(j)} = \frac{1}{2\pi i} \oint_{\partial\Lambda_\varepsilon(Z)} \left[(\zeta I - Z)^{-1} K \right]^k (\zeta I - Z)^{-1} \zeta^{(k-1)j^* + j - 1} \, d\zeta \, Z^{j^*},$$

with ε any positive real number. Now the integral can be bounded in the following way

$$\| [S^k C]^{(j)} \|_2 \leqslant \frac{1}{2\pi} \| K \|_2^k \oint_{\partial\Lambda_\varepsilon(Z)} \| (\zeta I - Z)^{-1} Z^{j^*} \|_2 \| (\zeta I - Z)^{-1} \|_2^k | \zeta |^{(k-1)j^* + j - 1} | d\zeta |$$

$$\leqslant \frac{L_\varepsilon(Z)}{2\pi} \| K \|_2^k \left(\frac{1}{\varepsilon} \right)^{k+1} \Gamma_\varepsilon(Z) [\rho_\varepsilon(Z)]^{kj^* + j}.$$

This proves Lemma A.2. □

The assertion (3.11) in Theorem 3.3 is identical with the assertion of Lemma A.2.

References

[1] A. Bellen, *Parallelism across the Steps for Difference and Differential Equations*, Lecture Notes in Mathematics 1386 (Springer-Verlag, Berlin, 1987) 22–35.

[2] A. Bellen, R. Vermiglio and M. Zennaro, Parallel ODE-solvers with stepsize control, *J. Comput. Appl. Math.* 31 (1990) 277–293.

[3] J.C. Butcher, *The Numerical Analysis of Ordinary Differential Equations, Runge–Kutta and General Linear Methods* (Wiley, Chichester, 1987).

[4] P. Chartier, Parallelism in the numerical solution of initial value problems for ODEs and DAEs, Thesis, Université de Rennes I, France (1993).

[5] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations, II: Stiff and Differential-Algebraic Problems* (Springer-Verlag, Berlin, 1991).

[6] K.R. Jackson and S.P. Nørsett, The potential for parallelism in Runge–Kutta methods, *SIAM J. Numer. Anal.* 32 (1995) 49–82.
[7] I. Lie, Some aspects of parallel Runge–Kutta methods, Report 3/87, Department of Mathematics, University of Trondheim, Norway (1987).
[8] L. Reichel and L.N. Trefethen, Eigenvalues and pseudo-eigenvalues of Toeplitz matrices, *Linear Algebra Appl.* 162–164 (1992) 153–185.
[9] D.M. Young, *Iterative Solution of Large Linear Systems* (Academic Press, New York, 1971).
[10] P.J. van der Houwen and B.P. Sommeijer, Iterated Runge-Kutta methods on parallel computers, *SIAM J. Sci. Statist. Comput.* 12 (1991) 1000–1028.
[11] P.J. van der Houwen, B.P. Sommeijer and W.A. van der Veen, Parallel iteration across the steps of high order Runge–Kutta methods for nonstiff initial value problems, *J. Comput. Appl. Math.* (to appear).
[12] P.J. van der Houwen, B.P. Sommeijer and W.A. van der Veen, Parallelism across the steps in iterated Runge–Kutta methods for stiff initial value problems, *Numer. Algorithms* 8 (1994) 293–312.
[13] W.A. van der Veen, Step-parallel algorithms for stiff initial value problems, *Comput. Math. Appl.* (to appear).