Parallel predictor-corrector methods

P.J. van der Houwen, B.P. Sommeijer, J.J.B. de Swart

Department of Numerical Mathematics

Report NM-R9408 March 1994

CWI is the National Research Institute for Mathematics and Computer Science. CWI is part of the Stichting Mathematisch Centrum (SMC), the Dutch foundation for promotion of mathematics and computer science and their applications.
SMC is sponsored by the Netherlands Organization for Scientific Research (NWO). CWI is a member of ERCIM, the European Research Consortium for Informatics and Mathematics.

P.O. Box 94079, 1090 GB Amsterdam (NL)
Kruislaan 413, 1098 SJ Amsterdam (NL)
Telephone +31 20 592 9333
Telefax +31 20 592 4199

Parallel Predictor-Corrector Methods

P.J. van der Houwen, B.P. Sommeijer & J.J.B. de Swart
CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

Abstract

In this paper we construct predictor-corrector methods using block Runge-Kutta methods as correctors. Like conventional Runge-Kutta methods, these correctors compute stage values at nonuniformly distributed, intermediate points. The predictor-corrector nature of the methods make them suitable for implementation on parallel computers. Comparisons of an 8th-order, 5-processor predictor-corrector method using Radau II points with the celebrated 8(7) Runge-Kutta method of Prince and Dormand shows speed-up factors of about 2.7.

CR Subject Classification (1991): G.1.7
Keywords and Phrases: numerical analysis, predictor-corrector iteration, Runge-Kutta methods, parallelism.

Note: The research reported in this paper was partly supported by STW (Netherlands Foundation for the Technical Sciences).

1. Introduction

We shall consider predictor-corrector methods (PC methods) for solving the (nonstiff) initial value problem (IVP)

$$\mathbf{y}'(t) = \mathbf{f}(\mathbf{y}(t)), \quad \mathbf{y}(t_0) = \mathbf{y}_0, \quad \mathbf{y}, \mathbf{f} \in \mathbb{R}^d \tag{1.1}$$

on parallel computers. On one-processor computers, PC methods based on linear multistep methods of Adams-type are most widely used. However, the use of multi-processor computers enables us to apply PC methods with a much more powerful corrector than in the conventional one-processor PC methods. This has already been observed and tried out in a number of papers, e.g. in [2], [9], [8], [7], and [5]. Moreover, parallel computers also allow us to use local Richardson extrapolation for automatic stepsize control without additional *sequential* costs, because the "reference" solution used in the error estimate can be computed in parallel.

So far, the investigations of parallel PC methods have mainly been concerned with Runge-Kutta (RK) correctors. By their one-step nature, these correctors allow an easy and highly efficient stepsize strategy (provided that the predictor formula is also of one-step type). However, for reaching the order of the RK corrector, the number of righthand side evaluations per step required by such one-step PC methods equals the order of the corrector. In [5] experiments are reported showing that the sequential costs of one-step PC methods based on the Gauss-Legendre corrector of order 10 are about half the sequential costs of the DOPRI8 code. The DOPRI8 code of Hairer-Nørsett-Wanner [3] is based on the 13-stage, 8th-order embedded RK method of Prince and Dormand [10], and is generally

Report NM-R9408
ISSN 0169-0388
CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

considered as one of the most efficient sequential codes. By sacrifycing the one-step nature of the predictor-corrector pair, the efficiency of parallel PC methods can be improved drastically, of course at the cost of a less easy implementation and stepsize strategy. A few first experiments were reported in [8].

In this paper, we shall try to find efficient predictor-corrector pairs by looking in the large class of general linear methods introduced by Butcher in 1966 (a detailed treatment of these methods can be found in Butcher [1, p. 335 ff.]). In particular, we shall be concerned with the family of block Runge-Kutta (BRK) correctors studied in [6], where a few results for BRK-based PC methods can be found. Here, we shall pursue these investigations. In particular, we pay attention to the stability of the corrector, because the weak point of most block methods is their small stability region.

In Section 2, we specify a family of two-stage BRK correctors and we discuss the order of accuracy and their stability. Section 3 analyses PC iteration of these BRK correctors and defines the convergence factors associated with the iteration process. The main results of this paper can be found in Section 4 where a number of BRK correctors are constructed that combine high order of accuracy, fast convergence and sufficiently large stability boundaries. Finally, in Section 5, PC methods based on BRK pairs are compared with DOPRI8 indicating the superiority of the BRK methods.

2. Block Runge-Kutta methods

For the definition and analysis of the block Runge-Kutta (BRK) methods, it is convenient to introduce some notations. Firstly, we shall frequently use the componentwise notation for functions of vectors. For example, $\mathbf{v}^2$ is understood to be the vector whose entries are the squares of the entries of $\mathbf{v}$. Furthermore, $\mathbf{e}$ denotes the vector with unit entries, $\mathbf{e}_i$ the ith unit vector whose entries vanish except for the ith entry which equals 1, I_{dd} is the d-by-d identity matrix, O_{mn} is the m-by-n zero matrix, and E_{mn} is the m-by-n matrix whose entries are zero except for its nth column which equals $\mathbf{e}$. The dimension of $\mathbf{e}$ and $\mathbf{e}_i$ may change, but will always be clear from the context.

Our starting point is a method of the form

$$\mathbf{Y} = (A\otimes I_{dd})\mathbf{Y}_{n-1} + h(B\otimes I_{dd})\mathbf{F}(\mathbf{Y}_{n-1}) + h(C\otimes I_{dd})\mathbf{F}(\mathbf{Y}), \tag{2.1a}$$

$$\mathbf{Y}_n = (A^*\otimes I_{dd})\mathbf{Y}_{n-1} + h(B^*\otimes I_{dd})\mathbf{F}(\mathbf{Y}_{n-1}) + h(C^*\otimes I_{dd})\mathbf{F}(\mathbf{Y}), \tag{2.1b}$$

$$\mathbf{y}_n = \mathbf{y}_{n-1} + h(\mathbf{b}_s^T\otimes I_{dd})\mathbf{F}(\mathbf{Y}_{n-1}) + h(\mathbf{c}_s^T\otimes I_{dd})\mathbf{F}(\mathbf{Y}_n), \qquad n = 1, \ldots, N. \tag{2.1c}$$

Here, the s-by-s matrices A, B, C, A^*, B^*, C^* and the s-dimensional vectors $\mathbf{b}_s$ and $\mathbf{c}_s$ contain the method parameters, h denotes the stepsize $t_n - t_{n-1}$, and $\otimes$ denotes the Kronecker product. Furthermore, $\mathbf{Y}$ and $\mathbf{Y}_n$ represent numerical approximations to the exact solution vectors $\mathbf{y}(\mathbf{e}t_{n-1} + \mathbf{a}h)$ where $\mathbf{a}$ denotes the abscissa vector, and where for any vector $\mathbf{V} = (\mathbf{V}_i)$, $\mathbf{F}(\mathbf{V})$ contains the derivative values $(\mathbf{f}(\mathbf{V}_i))$. It is assumed that the components of $\mathbf{a}$ are distinct.

The formulas (2.1a), (2.1b) and (2.1c) are respectively called the stage vector equation with s *internal* stages, the output formula with s *external* stages, and the step point formula. The *external* stages are all explicit, while the internal stages can be implicit or explicit. For example, if C has q zero rows and r := s-q rows with nonzero entries, then there are q explicit stages and r fully implicit stages. The quantities $\mathbf{Y}$, $\mathbf{Y}_n$ and $\mathbf{y}_n$ are respectively called the stage vector, the output vector and the step point value.

With respect to parallel implementation, methods of this type have been studied in [6], and were called block Runge-Kutta (BRK) methods, because they can be obtained from conventional RK methods by replacing the scalar RK parameters by matrices and the stage values by blocks of stage values. Like RK methods, the stage values correspond to nonuniformly spaced points at the t-axis. In terms of the array notation used in [6], the method (2.1) can be represented as a two-stage BRK method with one explicit and one implicit (block) stage:

$$
(2.1') \qquad \begin{array}{c|cc} I & O & O \\ A & B & C \\ \hline A^* & B^* & C^* \end{array} .
$$

In the determination of the parameter matrices in the stage vector equation (2.1a), the order conditions (see Section 2.1) will play an important role, together with the requirement that the iteration method used for solving the stage vector equation is rapidly converging. In this paper, we will solve the stage vector equation by predictor-corrector (PC) type iteration. The convergence of PC iteration is largely controlled by the "magnitude" of the matrix C, that is, convergence is better as C is smaller in some sense. For example, its spectral radius is often a first indicator of the potential convergence speed (see Section 3). In this connection, we should remark that strictly triangular matrices C leads to a zero spectral radius (in fact, the method is an *explicit* method, so that no iteration process is needed). However, such explicit methods approximate the components of $\mathbf{Y}$ by *extrapolation* formulas which are considerably less accurate than the *interpolation* formulas associated with *implicit* methods. Fortunately, it turns out that high accuracy and fast convergence often go together. Hence, if the stage vector $\mathbf{Y}$ has sufficient accuracy and stability, then the output formula (2.1b) can be dropped (i.e. $A = A^*$, $B = B^*$, $C = C^*$, so that $\mathbf{Y}_n = \mathbf{Y}$).

In most of the BRK correctors constructed in this paper, we do not use an output formula. However, we shall show in Section 4.3 that output formulas can be used for stabilizing the corrector.

The step point formula can be used to increase the order at the step points (superconvergence). This can be achieved by setting $\mathbf{b}_s = \mathbf{0}$ and by identifying the components of $\mathbf{c}_s$ and the abscissa vector $\mathbf{a}$ with the quadrature weights and quadrature points of Gaussian quadrature formulas. Since in the methods considered in this paper, the order of the output vector will be at most s+1 or s+2, there is, as far as order of accuracy is concerned, no need for basing the abscissa vector on quadrature formulas of the highest possible order (Gauss-Legendre formulas). This leads us to use abscissa vectors with $a_1 \neq 0$ and $a_s = 1$ which often simplifies the implementation (e.g. the Radau II points fit into this group).

Finally, we remark that (2.1) reduces to an RK method by setting $A = A^* = E_{ss}$, $B = B^* = O$, $C = C^*$, $\mathbf{b}_s = \mathbf{0}$ and $\mathbf{c}_s^T = \mathbf{e}_s^T C$ with C denoting the RK matrix of the collocation method defined by the abscissa vector $\mathbf{a}$. We shall call this collocation method the *RK method associated with the abscissa vector* $\mathbf{a}$. By identifying the BRK method (2.1) in the first step with such an RK method, we can avoid the problem of computing starting values, because we only need the initial value y_0, and not the whole starting vector $\mathbf{Y}_0$.

2.1. Accuracy

Given the abscissa vector $\mathbf{a}$, the conditions for pth-order consistency of the stage vector equation (2.1a) are given by (see, e.g. [6])

$$A\mathbf{e} = \mathbf{e}, \quad A(\mathbf{a} - \mathbf{e})^j + jB(\mathbf{a} - \mathbf{e})^{j-1} + jC\mathbf{a}^{j-1} = \mathbf{a}^j, \quad j = 1, \ldots, p.$$

We may write these order conditions in the form

$$
(2.2a) \qquad \begin{aligned} &A\mathbf{e} = \mathbf{e}, \quad AX_{sp} + BW_{sp} + CV_{sp} = U_{sp}, \\ &U_{sp} := \left(\tfrac{1}{j}\mathbf{a}^j\right), \; V_{sp} := \left(\mathbf{a}^{j-1}\right), \; W_{sp} := \left((\mathbf{a} - \mathbf{e})^{j-1}\right), \; X_{sp} := \left(\tfrac{1}{j}(\mathbf{a} - \mathbf{e})^j\right), \; j = 1, \ldots, p, \end{aligned}
$$

where the lower indices again refer to the number of rows and columns of the matrix. If (2.2a) is satisfied, then p will be called the *internal stage order*. The vector of principal error constants associated with the stage vector equation is given by

(2.3a) $$\mathbf{E}_{p+1} := \frac{1}{(p+1)!}\left\{\mathbf{a}^{p+1} - A(\mathbf{a}-\mathbf{e})^{p+1} - (p+1)\begin{pmatrix}B & C\end{pmatrix}\begin{pmatrix}(\mathbf{a}-\mathbf{e})^p \\ \mathbf{a}^p\end{pmatrix}\right\}.$$

Note that for p = s, A = E_{ss} and B = O_{ss}, the stage vector equation reduces to the stage vector equation of the RK method with RK matrix $C = U_{ss}V_{ss}^{-1}$.

For the output formula (2.1b) we proceed as follows. Imposing the localizing assumption, that is, assuming that the components $\mathbf{Y}_{n-1,i}$ are on the locally exact solution curve through the point (t_{n-1}, y_{n-1}), we may set $\mathbf{Y}_{n-1} = \mathbf{y}(\mathbf{e}t_{n-2} + \mathbf{a}h)$ and, by virtue of (2.2a), $\mathbf{Y} = \mathbf{y}(\mathbf{e}t_{n-1} + \mathbf{a}h) + O(h^{p+1})$. Hence,

$$\begin{aligned}\mathbf{Y}_n &= (A^*\otimes I_{dd})\mathbf{Y}_{n-1} + h(B^*\otimes I_{dd})\mathbf{F}(\mathbf{Y}_{n-1}) + h(C^*\otimes I_{dd})\mathbf{F}(\mathbf{Y}) \\ &= (A^*\otimes I_{dd})\mathbf{y}(\mathbf{e}t_{n-2}+\mathbf{a}h) + h(B^*\otimes I_{dd})\mathbf{y}'(\mathbf{e}t_{n-2}+\mathbf{a}h) \\ &\quad + h(C^*\otimes I_{dd})\mathbf{y}'(\mathbf{e}t_{n-1}+\mathbf{a}h)) + O(h^{p+2}).\end{aligned}$$

By Taylor expansion it can be shown that

$$\mathbf{Y}_n = \mathbf{y}(\mathbf{e}t_{n-1} + \mathbf{a}h) + O(h^{p+2}) + O(h^{p^*+1}),$$

provided that

(2.2b) $$A^*\mathbf{e} = \mathbf{e}, \quad A^*X_{sp^*} + B^*W_{sp^*} + C^*V_{sp^*} = U_{sp^*},$$

where the matrices X_{sp^*}, W_{sp^*}, V_{sp^*}, and U_{sp^*} are defined as in (2.2a) with p replaced by p*. Thus, the output vector $\mathbf{Y}_n$ has order min {p+1,p*}. This order will be called the *external stage order*, or briefly *the* stage order.

There are two error vectors associated with the output formula, viz.

(2.3b)
$$\mathbf{E}^*_{1,p+2} := C^*\,\mathbf{E}_{p+1},$$
$$\mathbf{E}^*_{2,p^*+1} := \frac{1}{(p^*+1)!}\left\{\mathbf{a}^{p^*+1} - A^*(\mathbf{a}-\mathbf{e})^{p^*+1} - (p^*+1)\begin{pmatrix}B^* & C^*\end{pmatrix}\begin{pmatrix}(\mathbf{a}-\mathbf{e})^{p^*} \\ \mathbf{a}^{p^*}\end{pmatrix}\right\},$$

where $\mathbf{E}_{p+1}$ is defined by (2.3a).

Finally, we consider the step point formula (2.1c). It is possible to achieve order of consistency 2s for this formula, so that the order at the step points becomes min {2s,p+2,p*+1}. However, this may lead to rather large entries in $\mathbf{b}_s$ and $\mathbf{c}_s$. Alternatively, we may use a zero $\mathbf{b}_s$ vector and identify $\mathbf{c}_s^T$ with the last row vector of the RK matrix associated with **a**, that is, $\mathbf{c}_s^T = \mathbf{e}_s^T U_{ss}V_{ss}^{-1}$. If p_{RK} denotes the order of the RK method, then we have order of accuracy min $\{p_{RK},p+2,p^*+1\}$ at the step points. This will be called the *step point order*. We summarize the preceding discussion in the following theorem:

Theorem 2.1. If (2.2a) and (2.2b) are satisfied, then the BRK method (2.1) has stage order min {p+1,p*} and output vector errors given by (2.3b). If, in addition, $\mathbf{b}_s = \mathbf{0}$, $\mathbf{c}_s^T = \mathbf{e}_s^T U_{ss}V_{ss}^{-1}$, and if the abscissa vector **a** defines an RK method of order p_{RK}, then the step point order is bounded below by min$\{p_{RK},p+2,p^*+1\}$. []

2.2. Stability

In order to ensure stability for h = 0 (zero-stability), we shall require that A^* has s-1 eigenvalues inside the unit circle (since (2.2b) prescibes that $A^*\mathbf{e} = \mathbf{e}$, A^* necessarily has one eigenvalue 1). Such matrices will be referred to as *zero-stable matrices*.

For h > 0, stability also depends on the other parameter matrices and on the abscissa vector **a**. With respect to the scalar test equation $y' = \lambda y$, where λ runs through the spectrum of the Jacobian matrix $\partial \mathbf{f}(y_n)/\partial y$, we obtain the recursion

$$\mathbf{Y}_n = M(z)\mathbf{Y}_{n-1}, \quad M(z) := A^* + zB^* + zC^*(I - zC)^{-1}(A + zB), \quad z := \lambda h. \tag{2.4}$$

Assuming that the stability matrix M(z) has s distinct eigenvalues, we have that (cf. Varga [11])

$$\|\mathbf{Y}_n\|_2 \le \|M^n(z)\|_2 \, \|\mathbf{Y}_0\|_2, \quad \|M^n(z)\|_2 \approx \nu(z)\,[\rho(M(z))]^n \text{ as } n \to \infty, \tag{2.5}$$

where $\nu(z)$ is bounded by the condition number of M(z). This estimate suggests defining the stability region, and the real and imaginary stability intervals according to

$$\begin{aligned} &\mathbb{S} := \{z: \rho(M(z)) < 1\}, \\ &(-\beta_{real},0) := \{z: z \in \mathbb{S}, z < 0\}, \quad (-\beta_{imag},\beta_{imag}) := \{z: z \in \mathbb{S}, \mathrm{Re}\,(z) = 0, z \neq 0\}, \end{aligned} \tag{2.6}$$

where $\rho(.)$ denotes the spectral radius function. The quantities β_{real} and β_{imag} are respectively called the *real* and the *imaginary stability boundary* of the BRK method. By (2.6) stability conditions of the type $h < \beta/\rho(\partial \mathbf{f}/\partial y)$ are implied, so that we should require the method to have sufficiently large stability boundaries, say not less than 1. In addition, we should impose the condition that $\nu(z)$ is of moderate size, particularly for z = 0, because zero-stability implies $\rho(M(0)) = \rho(A^*) = 1$, so that

$$\|\mathbf{Y}_n\|_2 \le \nu(0)\,\|\mathbf{Y}_0\|_2 \text{ as } n \to \infty. \tag{2.5'}$$

If A^* is singular, then estimating an upper bound for $\nu(0)$ by means of the condition number of A^* is not possible. For example, this happens in the important case where $A^* = E_{ss}$ (in [6] BRK methods of this form were called BRK methods of *Adams type*). However, for such methods, $M^n(0) = [A^*]^n = E_{ss}$, hence $\|M^n(0)\|_2 = \|E_{ss}\|_2 = \sqrt{s}$, so that for z = 0, Adams-type BRK methods satisfy (2.5) with $\nu(0) = \sqrt{s}$.

3. The iteration scheme

We approximate the solution **Y** of (2.1a) by successive iterates $\mathbf{Y}^{(j)}$ satisfying the PC scheme (or fixed point iteration scheme)

$$\mathbf{Y}^{(j)} = (A \otimes I_{dd})\mathbf{Y}_{n-1} + h(B \otimes I_{dd})\mathbf{F}(\mathbf{Y}_{n-1}) + h(C \otimes I_{dd})\mathbf{F}\left(\mathbf{Y}^{(j-1)}\right), \quad j = 1, \ldots, m; \quad n \ge 1. \tag{3.1}$$

Evidently, if the iterates $\mathbf{Y}^{(j)}$ satisfying (3.1) converge to a fixed vector **V** as $j \to \infty$, then **V** = **Y**. In actual computation, the number of iterations m is dynamically determined by requiring that the corrector equation is solved within a given tolerance (cf. Section 5). This iteration scheme has a high degree of parallelism, because the *sequential* costs of each iteration are independent of the number of stages (provided that sufficiently many processors are available).

For the predictor formula providing $\mathbf{Y}^{(0)}$, we may take the explicit BRK method

(3.2) $$\mathbf{Y}^{(0)} = (A_0 \otimes I_{dd})\mathbf{Y}_{n-1} + h(B_0 \otimes I_{dd})\mathbf{F}(\mathbf{Y}_{n-1}).$$

One option defines A_0 and B_0 according to

(3.3a) $$\begin{pmatrix} A_0 & B_0 \end{pmatrix} = \begin{pmatrix} U_{s,2s-1} & \mathbf{e} \end{pmatrix} \begin{pmatrix} X_{s,2s-1} & \mathbf{e} \\ W_{s,2s-1} & \mathbf{0} \end{pmatrix}^{-1},$$

where $X_{s,2s-1}$, $W_{s,2s-1}$ and $U_{s,2s-1}$ are defined as in (2.2a). It is easily seen that (3.3a) satisfies (2.2a) for $A = A_0$, $B = B_0$, $C = O$ and $p = 2s-1$, so that (3.3a) generates predictor values of order 2s-1 (provided that the stage order of the BRK method (2.1) is at least 2s-1). In fact, (3.3a) is a Hermite integration formula generated by the abscissa vector **a**. In addition to the high orders of Hermite formulas, the error constants $\|\mathbf{E}_{p+1}\|_\infty$ as defined in (2.3a) are extremely small. In Table 3.1, this is illustrated for the Radau II abscissas. However, inspite of their high orders and relatively small error constants, Hermite predictor formulas have the drawback of extremely large coefficients in the parameter matrices A_0 and B_0, especially for larger values of s. This may cause considerable round-off errors unless sufficiently high arithmetic is used (we remark that to some extent, round-off can be suppressed by using shifted iterates $\mathbf{X}^{(j)} := \mathbf{Y}^{(j)} - \mathbf{e}\otimes y_{n-1}$ in an actual implementation (cf. [4, p.128])).

An alternative to the Hermite predictor formula is offered by the Adams-Bashforth-type formula defined by

(3.3b) $$A_0 = E_{ss}, \quad B_0 = U_{ss}W_{ss}^{-1},$$

which satisfies (2.2a) for $A = E_{ss}$, $B = B_0$, $C = O$ and $p = s$. Its error constants associated with the Radau II abscissas can be found in Table 3.1. From these figures it is clear that if arithmetic allows, we should use the Hermite predictors.

Table 3.1. Error constants $\|\mathbf{E}_{p+1}\|_\infty$ associated with Radau II abscissas for Hermite and Adams-Bashforth predictor formulas.

Predictor	s = 2	s = 3	s = 4	s = 5
Hermite	0.12	0.0087	0.00034	0.0000081
Adams-Bashforth	0.33	0.13	0.041	0.010

The method {(3.1),(3.2)} will be called a PIBRK method (Parallel Iterated BRK method). The sequential costs of PIBRK methods depend on the structure of the parameter matrices. Therefore, we postpone a discussion of computational costs until the special cases developed in this paper have been specified.

For the convergence analysis of (3.1) we define the iteration error

$$\varepsilon^{(j)} := \mathbf{Y}^{(j)} - \mathbf{Y}.$$

On substitution in (3.1), we obtain

$$\varepsilon^{(j)} = h(C\otimes I_{dd})\left[\mathbf{F}\left(\mathbf{Y}^{(j-1)}\right) - \mathbf{F}(\mathbf{Y})\right].$$

This relation immediately leads to the estimate

$$\| \varepsilon^{(j)} \| \le h L \| C \| \| \varepsilon^{(j-1)} \|,$$

where L denotes a Lipschitz constant on the righthand side function **f**. Although this estimate has the advantage of being valid for the general IVP (1.1), it does not provide much information for selecting efficient corrector methods. Therefore, we resort to the familiar approach of approximating the IVP by a linear model. In this way, we obtain detailed information on the iteration process for the class of linear IVPs. Like the linear stability theory, this linear convergence theory turns out to be highly reliable for a large class of *nonlinear* problems.

Assuming the righthand side function **f** sufficiently smooth, we may write

$$\mathbf{F}(\mathbf{U} + \delta) - \mathbf{F}(\mathbf{U}) = J(\mathbf{U})\delta + O(\delta^2),$$

where $J(\mathbf{U})$ is an sd-by-sd block-diagonal matrix whose diagonal blocks consists of the Jacobian matrices $\partial \mathbf{f}(\mathbf{U}_i)/\partial \mathbf{y}$, $\mathbf{U}_i$ being the components of **U**. On substitution, we straightforwardly derive the error recursion

$$\varepsilon^{(j)} = Z\varepsilon^{(j-1)} + O(\varepsilon^{(j-1)})^2.$$

where the matrix

$$Z = Z(hJ(\mathbf{Y})) := h\,(C \otimes I)\, J(\mathbf{Y}) \tag{3.4}$$

controls the convergence of the iteration scheme. Assuming that higher-order terms can be neglected, the iteration error of the stage vector satisfies

$$\varepsilon^{(j)} = [Z(hJ(\mathbf{Y}))]^j \varepsilon^{(0)} = h^j [(C \otimes I_{dd})\, J(\mathbf{Y})]^j \varepsilon^{(0)}. \tag{3.5}$$

Thus, the iteration matrix C plays a crucial role in the convergence of the PC iteration process.

We shall define the (averaged) *convergence factor* for the scalar test equation $y' = \lambda y$. For this test equation, (3.5) reduces to

$$\varepsilon^{(j)} = z^j C^j \varepsilon^{(0)}, \quad z := \lambda h. \tag{3.6}$$

Hence,

$$\| \varepsilon^{(m)} \|_\infty \le |z|^m \| C^m \|_\infty \| \varepsilon^{(0)} \|_\infty,$$

so that, with respect to the maximum norm, the (averaged) convergence factor over m iterations is given by

$$\alpha(m,z) := |z| \sqrt[m]{\| C^m \|_\infty}\,. \tag{3.7}$$

The region of convergence in the complex z-plane is given by $\alpha(m,z) < 1$, that is, the open disk

$$\mathbb{C}_m := \{z: |z| < \gamma_m\}, \quad \gamma_m := \frac{1}{\sqrt[m]{\| C^m \|_\infty}}, \tag{3.8}$$

where γ_m may be considered as the *convergence boundary*. From (3.8) we deduce the stepsize condition $h < \gamma_m/\rho(\partial \mathbf{f}/\partial \mathbf{y})$. Thus, large convergence boundaries relax the convergence condition and improve convergence at the same time.

In actual computation, one should satisfy both the convergence condition associated with (3.8) and the stability condition associated with (2.6), that is, the spectrum of the matrix $h\partial \mathbf{f}/\partial \mathbf{y}$ should be contained in the intersection of the stability region $\mathbb{S}$ and the convergence region $\mathbb{C}_m$ (here, we assume that the IVP is itself stable, so that the spectrum of $\partial \mathbf{f}/\partial \mathbf{y}$ is located in the left halfplane). As a consequence, there is no point in trying to construct correctors whose stability region is much larger than their region of convergence. However, it may be feasible to have correctors whose convergence region is much larger than their region of stability, because, as we just saw, large regions of convergence also improve convergence speed. Notice that strictly lower (or upper) triangular matrices C have zero convergence factors for $m \geq s+2$. However, as already remarked, then the generating BRK corrector (2.1) is explicit and therefore has reduced accuracy.

4. Construction of BRK correctors

In all BRK correctors considered in this paper, the abscissa vector **a** is identified with the Radau II points. We do not claim that these points are most optimal, but it is likely that results based on Radau II points are indicative for other sets of abscissas.

In the construction of BRK correctors, we have to take into account: (i) the consistency conditions (2.2), (ii) the zero-stability and condition of the matrix A^*, (iii) the stability region, and (iv) the rate of convergence. These aspects will be characterized by the step point order, by the condition number $\kappa_\infty(A^*)$ of A^* (provided A^* is nonsingular), the stability boundaries defined by (2.6), the condition number $\kappa_\infty(C)$ and the convergence boundaries defined in (3.8).

4.1. Adams-type correctors without output formulas

We start with the class of methods without output formula, that is, the output formula coincides with the stage vector equation, so that $A = A^*$, $B = B^*$, and $C = C^*$ (that is, there are no external stages). Within this class, zero-stability is automatically achieved by choosing the subclass of Adams methods, i.e. $A = E_{ss}$. One option for choosing the remaining matrices B and C is such that a high order of consistency is obtained. In our preliminary analysis of this type of methods we found that in general the convergence factors associated with the matrix C improve as the order of consistency increases. In particular, we observed that the entries in the upper part and in the lower righthand corner of C are relatively small. This observation led us to consider BRK correctors of which the matrix C is of the form

$$C = \begin{pmatrix} O & O \\ \underline{C}_1 & \underline{C}_2 \end{pmatrix},$$

where $\underline{C}_1$ and $\underline{C}_2$ respectively are an r-by-q and an r-by-r matrix with $q+r = s$. Evidently, such correctors have r *implicit* stages and q *explicit* stages. In particular, the matrix $\underline{C}_2$ determines the convergence of the PC iteration process.

We shall define the first q rows of the matrix B completely by consistency conditions. From (2.2a) it follows that the first q stages are consistent of order s if they coincide with the first q rows of the matrix $U_{ss}W_{ss}^{-1}$. In fact, the resulting formulas are *Adams-Bashforth* formulas (cf. the Adams-Bashforth predictor formula (3.3b)). Although these formulas are based on pure *extrapolation*, the extrapolation errors are relatively small, because they correspond to the first (and therefore smaller) components of the abscissa vector **a**.

The class of methods indicated above is defined by

(4.1a) $A = E_{ss}$, $B = (U_{ss} - CV_{ss})W_{ss}^{-1}$, $C = \begin{pmatrix} O & O \\ \underline{C}_1 & \underline{C}_2 \end{pmatrix}$,

(4.1b) $A^* = A$, $B^* = B$, $C^* = C$,

where the r-by-s matrix $\underline{C} := (\underline{C}_1\ \underline{C}_2)$ is still free. This method is zero-stable (because A is zero-stable). Since (4.1) satisfies (2.2a) for p = s it follows from Theorem 2.1 that the stage order equals s. In the following subsections, a few options for choosing the matrix $\underline{C}$ will be discussed.

4.1.1. Adams-Bashforth-Moulton methods. The most simple option defines the implicit stages by imposing consistency conditions of highest possible order, that is, $\underline{C}$ is defined by the r-by-s matrix occurring in the lower righthand corner of the s-by-2s matrix

(4.2) $$U_{s,2s} \begin{pmatrix} W_{s,2s} \\ V_{s,2s} \end{pmatrix}^{-1}.$$

The resulting BRK corrector defines the first q components of $\mathbf{Y}_n = \mathbf{Y}$ by (explicit) Adams-Bashforth-type formulas (of order s) and the last r components of $\mathbf{Y}_n$ by implicit Adams-type formulas of highest possible order of consistency (i.e. order 2s). In this respect, these implicit formulas resemble the conventional Adams-Moulton formulas. Therefore, we shall call refer to these correctors as *Adams-Bashforth-Moulton* correctors (ABM correctors). The special methods arising for q = 0 and q = r will be called *Adams-Moulton* (AM) correctors and *Adams-Bashforth* (AB) correctors, respectively.

We recall that the stage order of ABM correctors equals s. However, for AM correctors where no explicit stages occur (q = 0), the stage order becomes 2s. For q > 0, Theorem 2.1 shows that the step point order can be raised to s+1 by choosing in the step point formula $\mathbf{b}_s = \mathbf{0}$ and $\mathbf{c}_s^T = \mathbf{e}_s^T U_{ss} V_{ss}^{-1}$. However, it turns out that for ABM correctors $\mathbf{e}_s^T B \approx \mathbf{0}$ and $\mathbf{e}_s^T C \approx \mathbf{e}_s^T U_{ss} V_{ss}^{-1}$. Hence, in practical applications, we achieve superconvergence at the step points by defining the step point formula simply by $\mathbf{y}_n = (\mathbf{e}_s^T \otimes I_{dd})\mathbf{Y}_n$.

For a large number of ABM correctors we *numerically* computed the convergence and stability characteristics. An usual approach to determine the stability boundaries β_{real} and β_{imag} is to run along both axis in the complex z-plane and to check up to what point the spectral radius of the amplification matrix M(z) is bounded by one. Obviously, in the neighbourhood of the origin, this spectral radius is close to one (especially for the higher-order methods), and, due to rounding errors, it is hard to decide whether such a point really belongs to the stability interval or not. It might happen that the true value of the spectral radius is slightly larger than one, whereas its numerical value is slightly less than one (e.g., by a small multiple of the machine precision). In such cases the corresponding z-point is wrongly accepted as being stable. However, this is inherent to a numerical verification using a finite precision. In practice, such situations are of course quite innocent since an "instability" of this type will not manifest itself. Evidently, we may have the converse situation that a z-point is wrongly qualified as being unstable.

It turns out that for many ABM methods this numerical approach indicates a very small, or even empty imaginary stability interval. To circumvent the numerical uncertainty, we also computed the "practical" imaginary stability interval defined by $(-\beta^*_{imag}, \beta^*_{imag}) := \{z: \rho(M(z)) < 1+\varepsilon, \mathrm{Re}\, z = 0, \varepsilon > 0\}$. For sufficiently small values of ε, β^*_{imag} can be safely used as the imaginary stability boundary in practical computations. For a given value of r, it turns out that the stability boundaries *decrease* and the convergence boundaries *increase* with q. For r-values running from 2 to 5, Table 4.1 presents the cases where the stability boundaries β_{real} and β^*_{imag} (with $\varepsilon = 10^{-3}$) are both sufficiently large (say at least $\approx$ 1), while the convergence boundaries are maximal. A more extensive list including cases with smaller convergence and stability boundaries can be found in the Appendix to this paper. Furthermore, in Table 4.1 (and in all subsequent tables presenting stability boundaries),

numerical β-values less than 0.1 are replaced by an *. Notice that a large condition number for $\underline{C}_2$ implies relatively small convergence boundaries in the first few iterations.

Next, we discuss the sequential costs of the PIBRK method based on ABM correctors. Let us consider the following implementation of the PIBRK method:

$$\underline{\mathbf{Y}}^{(0)} = (\underline{A}_0 \otimes I_{dd})\mathbf{Y}_{n-1} + h(\underline{B}_0 \otimes I_{dd})\mathbf{F}^*_{n-1},$$

$$\bar{\mathbf{Y}}^* = (\bar{A} \otimes I_{dd})\mathbf{Y}_{n-1} + h(\bar{B} \otimes I_{dd})\mathbf{F}^*_{n-1},$$

(4.4)
$$\underline{\mathbf{Y}}^{(j)} = (\underline{A} \otimes I_{dd})\mathbf{Y}_{n-1} + h(\underline{B} \otimes I_{dd})\mathbf{F}^*_{n-1} + h(\underline{C}_1 \otimes I_{dd})\mathbf{F}(\bar{\mathbf{Y}}^*) + h(\underline{C}_2 \otimes I_{dd})\mathbf{F}(\underline{\mathbf{Y}}^{(j-1)}),$$

$$\mathbf{Y}_n = \begin{pmatrix} \bar{\mathbf{Y}}^* \\ \underline{\mathbf{Y}}^{(m)} \end{pmatrix}, \quad \mathbf{y}_n = (\mathbf{e}_s^T \otimes I_{dd})\mathbf{Y}_n, \quad \mathbf{F}^*_n = \begin{pmatrix} \mathbf{F}(\bar{\mathbf{Y}}^*) \\ \mathbf{F}(\underline{\mathbf{Y}}^{(m-1)}) \end{pmatrix},$$

where j = 1, ... , m. Here, upper and lower bars refer to the first q and last r rows of a matrix, and where the underlying matrices A, B and C are defined by (4.1) and (4.2).

Assuming that (4.4) converges for $m \to \infty$, it is easily verified that $(\bar{\mathbf{Y}}^*, \underline{\mathbf{Y}}^{(m)})$ converges to $(\bar{\mathbf{Y}}, \underline{\mathbf{Y}})$, i.e., to the corrector solution. If $1 \le q \le r$, then the sequential costs on r processors consist of m+1 righthand side evaluations, that is, the evaluation of $\mathbf{F}(\bar{\mathbf{Y}}^*)$ plus the evaluation of the m righthand side functions $\mathbf{F}(\underline{\mathbf{Y}}^{(j-1)})$. The evaluation of $\mathbf{F}(\bar{\mathbf{Y}}^*)$ can be done in parallel with that of $\mathbf{F}(\underline{\mathbf{Y}}^{(0)})$, but this would require q additional processors. However, if we apply local Richardson extrapolation for stepsize control and if we use additional processors for computing the "reference" solution, then these processors can also be used for evaluating $\mathbf{F}(\bar{\mathbf{Y}}^*)$. Since the "reference" solution is computed with a double step, it is likely that there is some idle time, in spite of the fact that larger steps will require more iterations to solve the stage vector equation. Hence, in such a case, the total sequential costs per step are just m righthand side evaluations.

Table 4.1. Characteristics for selected ABM correctors.

s = q+r	order	β_{real}	β_{imag}	β^*_{imag}	$\kappa_\infty(\underline{C}_2)$	γ_2	γ_3	γ_4	γ_{10}	...	γ_∞
3 = 1+2	4	3.33	*	3.81	17	2.48	3.08	3.55	5.16	...	6.17
4 = 2+2	5	0.94	*	0.91	15	3.82	4.55	5.20	7.79	...	9.06
4 = 1+3	5	2.88	*	2.53	81	1.79	2.39	2.96	5.91	...	8.23
5 = 2+3	6	1.26	*	1.78	64	2.48	3.26	3.99	7.47	...	10.35
5 = 1+4	6	1.88	*	2.90	383	1.68	2.11	2.59	5.45	...	10.68
6 = 2+4	7	1.38	0.99	0.99	239	2.06	2.64	3.26	6.60	...	12.28
5 = 0+5	6	2.26	*	3.16	13853	1.33	1.93	2.26	4.50	...	10.80
6 = 1+5	7	0.92	1.13	1.13	2700	1.56	2.05	2.47	5.12	...	13.05

Table 4.2. Characteristics for selected ABR correctors.

s = q+r	order	β_{real}	β_{imag}	β^*_{imag}	$\kappa_\infty(C_2)$	γ_2	γ_3	γ_4	γ_{10} ...	γ_∞
2 = 0+2	3	∞	∞	∞	7	1.41	1.59	1.86	2.36 ...	2.45
3 = 1+2	4	8.30	4.32	4.32	9	2.15	2.48	2.87	3.66 ...	4.31
4 = 2+2	5	1.05	*	0.93	10	3.39	3.92	4.49	5.93 ...	7.11
3 = 0+3	5	∞	∞	∞	18	1.41	1.82	2.21	3.03 ...	3.64
4 = 1+3	5	17.18	*	9.02	24	1.81	2.32	2.62	4.08 ...	4.94
5 = 2+3	6	1.97	*	1.89	27	2.45	3.08	3.47	5.80 ...	7.03
4 = 0+4	7	∞	∞	∞	34	1.41	1.82	2.21	4.28 ...	5.04
5 = 1+4	6	30.16	*	15.74	44	1.65	2.11	2.55	4.84 ...	5.99
6 = 2+4	7	3.35	*	2.86	50	2.04	2.61	3.15	5.80 ...	7.74
5 = 0+5	9	∞	∞	∞	55	1.41	1.82	2.21	4.44 ...	6.29
6 = 1+5	7	47.80	*	24.92	69	1.57	2.02	2.44	4.70 ...	6.87
7 = 2+5	8	5.23	*	4.57	79	1.84	2.36	2.85	5.40 ...	8.39

4.1.2. Adams-Bashforth-Radau methods. A second option identifies the matrix $(\underline{C}_1\ \underline{C}_2)$ in (4.1a) with the last r rows of the Radau IIA matrix $U_{ss}V_{ss}^{-1}$. Then the matrix $\underline{B}$ vanishes, so that the r implicit stages are determined by Radau formulas. The stage order and the step point order are the same as for ABM correctors, i.e. s and s+1, respectively. We shall call this corrector an *Adams-Bashforth-Radau corrector* (ABR corrector) because the first q components of $\mathbf{Y}_n$ are defined by Adams-Bashforth formulas and the last r components by Radau IIA formulas. For r = 0 the corrector reduces to the AB corrector and r = s leads to the Radau IIA correctors. The PIBRK method generated by the ABR correctors can be defined according to (4.4), so that the sequential costs are the same.

The analogue of Table 4.1 is given in Table 4.2 where we included the case q = 0 defining the pure Radau IIA corrector. A comparison of these selected methods with the corresponding ABM correctors reveals that ABR correctors have smaller convergence boundaries (particularly for larger m), but possess considerably larger stability boundaries. One may argue that the stability boundaries of the selected ABM correctors are sufficiently large for integrating nonstiff problems, so that the ABM correctors seem to be the more attractive ones. However, if the stage vector equation is not solved to convergence (for example, if the tolerance parameter in the stopping criterion is not sufficiently small), then we are faced with the fact that the stability region of the ABM method is much smaller than that of the ABR method. Section 5 will show that ABR is more efficient than ABM because of its better stability characteristics if the numbers of iterations is small.

4.2. Adams-type correctors with Radau output formula

By adding an output formula (that is, introducing external stages), it is possible to improve the stability of the corrector method. We shall illustrate this by using output formulas of Radau-type:

$$A^* = E_{ss},\ B^* = O_{ss},\ C^* = U_{ss}V_{ss}^{-1}. \tag{4.5}$$

If the stage order of $\mathbf{Y}$ is p, then $\mathbf{Y}_n$ has stage order min {p+1,s} and step point order at least s+1.

The stability matrix M(z) associated with {(4.1a),(4,5)} is given by

$$M(z) = E_{ss} + zU_{ss}V_{ss}^{-1}M_{Adams}(z), \tag{4.6}$$

where $M_{Adams}(z)$ is the stability matrix of (4.1). The large entries in $M_{Adams}(z)$ are responsible for the possibly poor stability of (4.1). Since the entries of the Radau matrix $U_{ss}V_{ss}^{-1}$ are rather small, it is likely that the large entries in $M_{Adams}(z)$ are neutralized, so that the stability region of M(z) is improved. This is confirmed by the Tables 4.3 and 4.4.

In comparison with the Adams methods without output formula, the higher stability has to be paid for by the additional evaluation of $\mathbf{F}(\mathbf{Y}_{n-1})$. This can be concluded from the following implementation of the generated PIBRK method (cf. (4.4)):

$$\underline{\mathbf{Y}}^{(0)} = (\underline{A}_0 \otimes I_{dd})\mathbf{Y}_{n-1} + h(\underline{B}_0 \otimes I_{dd})\mathbf{F}(\mathbf{Y}_{n-1}),$$

$$\bar{\mathbf{Y}}^{*} = (\bar{A} \otimes I_{dd})\mathbf{Y}_{n-1} + h(\bar{B} \otimes I_{dd})\mathbf{F}(\mathbf{Y}_{n-1}),$$

(4.7)
$$\underline{\mathbf{Y}}^{(j)} = (\underline{A} \otimes I_{dd})\mathbf{Y}_{n-1} + h(\underline{B} \otimes I_{dd})\mathbf{F}(\mathbf{Y}_{n-1}) + h(\underline{C}_1 \otimes I_{dd})\mathbf{F}(\bar{\mathbf{Y}}^{*}) + h(\underline{C}_2 \otimes I_{dd})\mathbf{F}(\underline{\mathbf{Y}}^{(j-1)}),$$

$$\mathbf{Y}_n = (E_{ss} \otimes I_{dd})\mathbf{Y}_{n-1} + h(U_{ss}V_{ss}^{-1} \otimes I_{dd}) \begin{pmatrix} \mathbf{F}(\bar{\mathbf{Y}}^{*}) \\ \mathbf{F}(\underline{\mathbf{Y}}^{(m-1)}) \end{pmatrix}, \quad y_n = (e_s^T \otimes I_{dd})\mathbf{Y}_n.$$

where j = 1, ... , m. Again, if this process converges, then it converges to the corrector solution.Note that the evaluation of $\mathbf{F}(\mathbf{Y}_{n-1})$ cannot be replaced by $\mathbf{F}^*_{n-1}$ as in (4.4), because then the effect of the stabilizing output formula is not taken into account. However, in the ABR case (where the mth iteration is identical with the output formula), we may replace the last r components of $\mathbf{F}(\mathbf{Y}_{n-1})$ by those of $\mathbf{F}^*_{n-1}$, without changing the corrector solution. Thus, with respect to the method (4.4), the additional costs are one righthand side evaluation in the ABR case and two righthand side evaluations in the ABM case. As before, if we have q additional processors at our disposal, then the total sequential costs per step can be reduced by one righthand side evaluation (cf. the discussion of the method (4.4)).

The Tables 4.3 and 4.4 illustrate the stabilizing effect of adding a Radau output formula.

Table 4.3. ABM (+ Radau) correctors.

Corrector	s = q+r	order	β_{real}	β_{imag}	β^*_{imag}
ABM	6 = 4+2	7	*	*	*
ABM + R	6 = 4+2	7	1.51	*	1.40
ABM	6 = 3+3	7	0.17	*	0.18
ABM + R	6 = 3+3	7	1.98	1.79	1.79
ABM	7 = 4+3	8	*	*	*
ABM + R	7 = 4+3	8	1.01	*	0.95
ABM	7 = 3+4	8	0.19	0.22	0.22
ABM + R	7 = 3+4	8	2.18	1.83	1.83
ABM	7 = 2+5	8	0.47	0.36	0.36
ABM + R	7 = 2+5	8	2.31	*	2.78

Table 4.4. ABR (+ Radau) correctors.

Corrector	s = q+r	order	β_{real}	β_{imag}	β^*_{imag}
ABR	6 = 4+2	7	*	*	*
ABR + R	6 = 4+2	7	1.52	*	1.41
ABR	6 = 3+3	7	0.18	*	0.19
ABR + R	6 = 3+3	7	1.70	*	1.81
ABR	7 = 4+3	8	*	*	*
ABR + R	7 = 4+3	8	0.98	*	0.97
ABR	7 = 3+4	8	0.27	*	0.28
ABR + R	7 = 3+4	8	1.90	*	2.08
ABR	8 = 3+5	9	0.40	*	0.43
ABR + R	8 = 3+5	9	2.27	*	2.54

4.3. More general correctors.

In the ABM and ABR correctors of Section 4.1, the first q (explicit) stages have order s, so that the resulting stage order can never exceed s. The stage order can easily be increased by using a number of the zero entries occurring in the matrices $\underline{A}$, B and C defined in (4.1). For example, adding to the ABR corrector, the (s-1)st column of $\bar{A}$ and the last column of $\underline{B}$ for satisfying additional consistency conditions, we obtain a corrector of order s+1. This corrector may be considered as a "minimal" modification of the ABR corrector and will be referred to as the *modified* ABR corrector. A drawback of these modified correctors is the rather large magnitude of the entries in the matrix $\bar{A}$, even in the case of this minimal modification. Using more zero entries for a further increase of the stage order leads to dramatically large entries, so that it does not seem feasible to use this approach for constructing correctors with stage order $\geq$ s+2.

Since the matrix $\underline{C}$ of the modified ABR correctors is no longer defined by the Radau IIA formulas, the step point formula $\mathbf{y}_n = (\mathbf{e}_s^T \otimes I_{dd})\mathbf{Y}_n$ does not have superconvergence at the step points. Nevertheless, in practical applications we do observe step point order s+2, because, again it turns out that $\mathbf{e}_s^T B \approx \mathbf{0}$ and $\mathbf{e}_s^T C \approx \mathbf{e}_s^T U_{ss} V_{ss}^{-1}$ (cf. the discussion in Subsection 4.1.1). Hence, the modified ABR method can be implemented according to (4.4), so that the sequential costs per step are the same as for the ABM and ABR methods.

The characteristics of a few modified ABR correctors are summarized in Table 4.5. A comparison with the corresponding ABR correctors of Table 4.2 reveals that the modification leads to comparable convergence boundaries and smaller stability boundaries. However, the stage order and (effective) step point order is raised by one. A detailed investigation of this promising family of methods will be subject of future research.

Table 4.5. Characteristics for selected modified ABR correctors.

s = q+r	order	β_{real}	β_{imag}	β^*_{imag}	$\kappa_\infty(\underline{C}_2)$	γ_2	γ_3	γ_4	γ_{10}	...	γ_∞
3 = 1+2	5	4.15	*	1.44	12	2.33	2.72	3.12	4.14	...	4.98
4 = 1+3	6	7.16	*	2.41	34	1.83	2.37	2.86	4.85	...	5.97
6 = 2+4	8	0.99	*	1.17	64	2.05	2.63	3.20	6.04	...	8.73
7 = 2+5	9	1.42	*	1.74	107	1.85	2.38	2.89	5.66	...	9.55

5. Numerical experiments

Our numerical tests were performed using 15-digits arithmetic. The accuracies obtained are given by the number of correct digits Δ, defined by writing the maximum norm of the absolute error at the endpoint in the form $10^{-\Delta}$. The PIBRK method is implemented according to (4.4) with the PC pairs (AB, ABM) and (AB, ABR), where the correctors have orders 7 and 8, and are selected from the Tables 4.1 and Table 4.2. In view of the relatively high corrector orders and the 15-digits arithmetic, we did not use Hermite predictors.

We took two well-known test problems from [3], viz. the Fehlberg problem

$$\begin{aligned} y_1' &= 2\,t\,y_1 \log\left(\max\{y_2, 10^{-3}\}\right), \quad y_1(0) = 1, \\ y_2' &= -2\,t\,y_2 \log\left(\max\{y_1, 10^{-3}\}\right), \quad y_2(0) = e, \end{aligned} \qquad 0 \le t \le 5, \tag{5.1}$$

and the Euler problem

$$\begin{aligned} y_1' &= y_2 y_3, & y_1(0) &= 0, \\ y_2' &= -y_1 y_3, & y_2(0) &= 1, \\ y_3' &= -.51 y_1 y_2, & y_3(0) &= 1, \end{aligned} \qquad 0 \le t \le 20. \tag{5.2}$$

5.1. Comparison of ABM and ABR correctors

We applied the PC pairs (AB, ABM) and (AB, ABR) with s = 2+4. These correctors are both of order 7, require four processors, and are equally expensive. The Tables 5.1 and 5.2 present Δ-values for a few values of h and m (overflow is indicated by *). From these figures, we may conclude that the efficiency of the two methods is comparable in the case of convergence, but for larger stepsizes (AB, ABR) is more robust than (AB, ABM). This can be explained by the larger stability regions of the (AB, ABR) method.

Table 5.1. Δ-values for (5.1) obtained by (4.4) with (AB, ABM) and (AB, ABR) PC pairs.

PC pair	s = q+r	h^{-1}	m = 1	m = 2	m = 3	m = 4	...	m = ∞
(AB, ABM)	2+4	10	*	*	*	*	...	3.7
(AB, ABR)	2+4	10	*	*	2.6	3.7	...	4.2
(AB, ABM)	2+4	20	0.5	*	3.0	6.5	...	7.5
(AB, ABR)	2+4	20	0.5	4.5	5.9	6.5	...	6.9
(AB, ABM)	2+4	40	4.6	*	8.8	9.3	...	9.6
(AB, ABR)	2+4	40	4.6	7.2	9.0	9.2	...	9.3
(AB, ABM)	2+4	80	7.9	9.2	11.7	11.7	...	11.7
(AB, ABR)	2+4	80	7.9	9.4	12.0	11.5	...	11.5

Table 5.2. Δ-values for (5.2) obtained by (4.4) with (AB, ABM) and (AB, ABR) PC pairs.

PC pair	s = q+r	h^{-1}	m = 1	m = 2	m = 3	m = 4	...	m = ∞
(AB, ABM)	2+4	1	*	*	*	*	...	4.6
(AB, ABR)	2+4	1	*	*	3.2	3.9	...	4.9
(AB, ABM)	2+4	2	*	*	2.2	6.4	...	6.5
(AB, ABR)	2+4	2	*	4.6	5.4	6.6	...	6.4
(AB, ABM)	2+4	4	4.4	*	8.0	8.4	...	8.4
(AB, ABR)	2+4	5	4.4	7.7	8.1	8.3	...	8.3
(AB, ABM)	2+4	8	7.2	9.1	10.5	10.6	...	10.6
(AB, ABR)	2+4	10	7.1	10.2	10.4	10.4	...	10.4

5.2. Comparisons with DOPRI8, PIRK8 and PIRK 10

Since (AB, ABR) pairs are more stable and therefore more robust, we restrict our considerations to this family of PC methods. In particular, we tested the s = 2+5 method. This parallel, eighth-order (AB, ABR) method was compared with the 8(7) RK pair of Prince and Dormand [10] and the parallel PC methods based on Gauss-Legendre correctors of order 8 and 10. For the Dormand-Prince method we took the DOPRI8 implementation of Hairer, Nørsett and Wanner [3], and for the Gauss-Legendre methods we used the four and five-processor one-step codes PIRK8 and PIRK10 developed in [5].

The (AB, ABR) method was equipped with a dynamic iteration strategy based on the requirement that the step point component of the residue left on substitution of the jth iterate into the stage vector equation should be less than a tolerance parameter TOL, i.e.

$$\|(\mathbf{e}_s^T \otimes I_{dd})\mathbf{R}^{(j)}\| \le \mathrm{TOL},$$

where

$$\mathbf{R}^{(j)} := \mathbf{Y}^{(j)} - (E_{ss}\otimes I_{dd})\mathbf{Y}_{n-1} - h(B\otimes I_{dd})\mathbf{F}(\mathbf{Y}_{n-1}) - h(C\otimes I_{dd})\mathbf{F}(\mathbf{Y}^{(j)}).$$

According to (4.4) we may write

$$\mathbf{R}^{(j)} = \mathbf{Y}^{(j)} - \mathbf{Y}^{(j+1)} - h(B\otimes I_{dd})(\mathbf{F}(\mathbf{Y}_{n-1}) - \mathbf{F}^*_{n-1}).$$

Furthermore, we require that TOL is a factor δ less than the local error. In our experiments, we shall estimate the local error at t_{n-1} by $\|(\mathbf{e}_s^T\otimes I_{dd})(\mathbf{Y}_{n-1} - \mathbf{Y}_{n-1}^{(0)})\|$, where $\mathbf{Y}_{n-1}^{(0)}$ denotes the prediction in the preceding step. Using the maximum norm and observing that $(\mathbf{e}_s^T\otimes I_{dd})(\mathbf{F}(\mathbf{Y}_{n-1}) - \mathbf{F}^*_{n-1})$ vanishes in the case of ABR correctors, we are led to the stopping criterion

$$\|(\mathbf{e}_s^T\otimes I_{dd})(\mathbf{Y}^{(j)} - \mathbf{Y}^{(j+1)})\|_\infty \le \delta\, \|(\mathbf{e}_s^T\otimes I_{dd})(\mathbf{Y}_{n-1} - \mathbf{Y}_{n-1}^{(0)})\|_\infty. \tag{5.3}$$

Below, the resulting implementation will be referred to as the ABR8 code (we did not yet implement a stepsize strategy, so that the results produced by this code may be improved when a stepsize strategy is included).

The Tables 5.3 and 5.4 show results taken from [5] and results obtained by ABR8 for $\delta = 10^{-4}$ (the number of sequential righthand side evaluations needed by the codes are obtained by interpolation to arrive at integer values of Δ). The superiority of ABR8 is clear. For these two

problems, the averaged factor by which ABR8 beats the 8th-order code DOPRI8 (to be considered as one of the most efficient sequential codes), the four-processor PIRK8 code of order 8, and the five-processor 10th-order code PIRK10 are about 2.6, 2.0 and 1.4, respectively. Similar speed-up factors were obtained for many other test problems.

Table 5.3. Number of sequential righthand side evaluations for the Fehlberg problem (5.1).

Method	Δ=5	Δ=6	Δ=7	Δ=8	Δ=9	Δ=10	Δ=11
DOPRI8	595	759	963	1227	1574	1990	2503
PIRK8	379	495	623	786	978	1383	1874
PIRK10	327	388	490	704	884	977	1078
ABR8	240	335	430	532	689	846	1067

Table 5.4. Number of sequential righthand side evaluations for the Euler problem (5.2).

Method	D=6	D=7	D=8	D=9	D=10	D=11	D=12
DOPRI8	415	576	728	898	1133	1422	1817
PIRK8	294	381	534	728	961	1172	1746
PIRK10	252	297	357	426	580	730	920
ABR8	160	192	223	293	379	506	643

6. Concluding remarks

The search for efficient parallel PC methods reported in this paper has resulted in several fastly converging and sufficiently stable PC pairs. With respect to the fully automatic code DOPRI8, the averaged speed-up factor of the fixed stepsize, five-processor ABR8 code is about 2.6. The efficiency of this code can be improved by including a stepsize strategy. If the local error estimate is based on local Richardson extrapolation where the "reference" solution is computed in parallel on an additional set of processors, then these processors can also be used for saving one function call per step (see the discussion of the scheme (4.4)). In the numerical examples of this paper, this would increase the speed-up factor by about 20%.

Acknowledgement

The authors are grateful to Dr. K.J. in 't Hout for his interest and useful remarks during the preparation of this paper.

References

[1] Butcher, J.C. (1987): The numerical analysis of ordinary differential equations, Runge-Kutta and general linear methods, Wiley, New York.

[2] Chu, M.T. & Hamilton, H. (1987): Parallel solution of ODE's by multi-block methods, SIAM J. Stat. Comput. 8, 342-353.

[3] Hairer, E., Nørsett, S.P. & Wanner, G. (1987): Solving ordinary differential equations I. Nonstiff problems, Springer Series in Comp. Math., vol. 8, Springer-Verlag, Berlin.

[4] Hairer, E. & Wanner, G. (1991): Solving ordinary differential equations, II. Stiff and differential-algebraic problems, Springer Series in Comp. Math., vol. 14, Springer-Verlag, Berlin.

[5] Houwen, P.J. van der, & Sommeijer, B.P. (1990): Parallel iteration of high-order Runge-Kutta methods with stepsize control, J. Comput. Appl. Math. 29, 111-127.

[6] Houwen, P.J. van der & Sommeijer, B.P. (1992): Block Runge-Kutta methods on parallel computers, Z. angew. Math. Mech. 72, 3-18.

[7] Jackson, K.R. & Nørsett, S.P. (1990): The potential for parallelism in Runge-Kutta methods, Part I: RK formulas in standard form, Technical Report No. 239/90, Department of Computer Science, University of Toronto.

[8] Lie, I (1987): Some aspects of parallel Runge-Kutta methods, Report 3/87, Dept. of Mathematics, University of Trondheim

[9] Nørsett, S.P. & Simonsen, H.H. (1989): Aspects of parallel Runge-Kutta methods, in: A. Bellen, C.W. Gear and E Russo (Eds.): Numerical Methods for Ordinary Differential Equations, Proceedings L'Aquila 1987, LNM 1386, Springer-Verlag, Berlin.

[10] Prince, P.J. & Dormand, J.R. (1981): High order embedded Runge-Kutta formulae, J. Comput. Appl. Math. 7, 67-75.

[11] Varga, R.S. (1962): Matrix iterative analysis, Prentice Hall, Englewood Cliffs, N.J.

Appendix

Table A1: Extension of Table 4.1

s	=	$q+r$	order	β_{real}	β_{imag}	β^*_{imag}	$\kappa_\infty(\underline{C}_2)$	γ_2	γ_3	γ_4	γ_{10}	...	γ_∞
2	=	2 + 0	3	0.91	*	0.29	-	∞	∞	∞	∞	...	∞
3	=	3 + 0	4	0.59	0.60	0.61	-	∞	∞	∞	∞	...	∞
4	=	4 + 0	5	0.48	*	0.44	-	∞	∞	∞	∞	...	∞
5	=	5 + 0	6	0.44	*	0.44	-	∞	∞	∞	∞	...	∞
6	=	6 + 0	7	0.41	*	0.42	-	∞	∞	∞	∞	...	∞
7	=	7 + 0	8	0.40	*	0.40	-	∞	∞	∞	∞	...	∞
8	=	8 + 0	9	0.39	*	0.39	-	∞	∞	∞	∞	...	∞
2	=	1 + 1	3	2.60	1.64	1.65	1.00	4.44	4.44	4.44	4.44	...	4.44
3	=	2 + 1	4	0.54	0.65	0.65	1.00	9.18	9.18	9.18	9.18	...	9.18
4	=	3 + 1	5	0.11	*	0.12	1.00	16.06	16.06	16.06	16.06	...	16.06
5	=	4 + 1	6	*	*	*	1.00	25.02	25.02	25.02	25.02	...	25.02
6	=	5 + 1	7	*	*	*	1.00	36.00	36.00	36.00	36.00	...	36.00
7	=	6 + 1	8	*	*	*	1.00	49.00	49.00	49.00	49.00	...	49.00
8	=	7 + 1	9	*	*	*	1.00	64.00	64.00	64.00	64.00	...	64.00
2	=	0 + 2	3	7.50	*	1.02	27.02	1.59	2.09	2.48	4.09	...	4.85
3	=	1 + 2	4	3.33	*	3.81	17.17	2.48	3.08	3.55	5.16	...	6.17
4	=	2 + 2	5	0.94	*	0.91	15.25	3.82	4.55	5.20	7.79	...	9.06
5	=	3 + 2	6	0.13	*	0.13	14.38	5.53	6.50	7.40	11.30	...	12.89
6	=	4 + 2	7	*	*	*	13.86	7.59	8.88	10.10	15.58	...	17.58
7	=	5 + 2	8	*	*	*	13.51	10.03	11.71	13.29	20.60	...	23.12
8	=	6 + 2	9	*	*	*	13.25	12.84	14.96	16.97	26.33	...	29.48
3	=	0 + 3	4	5.53	*	3.67	157.29	1.30	1.78	2.24	4.78	...	6.36
4	=	1 + 3	5	2.88	*	2.53	81.27	1.79	2.39	2.96	5.91	...	8.23
5	=	2 + 3	6	1.26	*	1.78	63.66	2.48	3.26	3.99	7.47	...	10.36
6	=	3 + 3	7	0.17	*	0.18	55.63	3.32	4.34	5.24	9.53	...	13.28
7	=	4 + 3	8	*	*	*	51.05	4.33	5.61	6.67	12.03	...	16.83
8	=	5 + 3	9	*	*	*	48.08	5.49	7.08	8.32	14.95	...	20.92
4	=	0 + 4	5	3.91	*	6.31	1109.38	1.53	1.77	2.14	4.61	...	8.38
5	=	1 + 4	6	1.88	*	2.90	383.04	1.68	2.11	2.59	5.45	...	10.68
6	=	2 + 4	7	1.38	0.99	0.99	238.64	2.06	2.64	3.26	6.60	...	12.28
7	=	3 + 4	8	0.19	0.22	0.22	180.61	2.57	3.31	4.07	8.01	...	14.71
8	=	4 + 4	9	*	*	*	150.77	3.17	4.10	5.01	9.67	...	17.67
5	=	0 + 5	6	2.26	*	3.16	1.39 E4	1.33	1.93	2.26	4.50	...	10.80
6	=	1 + 5	7	0.92	1.13	1.13	2684.02	1.56	2.05	2.47	5.12	...	13.05
7	=	2 + 5	8	0.47	0.36	0.36	1131.09	1.85	2.39	2.90	6.03	...	14.32
8	=	3 + 5	9	*	*	0.10	664.12	2.20	2.84	3.46	7.12	...	16.35
6	=	0 + 6	7	1.00	*	1.14	2.06 E5	1.12	1.70	2.23	4.50	...	13.07
7	=	1 + 6	8	0.34	0.36	0.36	2.37 E5	1.45	1.94	2.41	4.92	...	15.17
8	=	2 + 6	9	0.11	*	0.11	6762.53	1.73	2.22	2.72	5.61	...	16.60
7	=	0 + 7	8	0.35	0.36	0.37	3.62 E6	0.79	1.87	2.12	4.55	...	15.17
8	=	1 + 7	9	0.10	*	0.11	3.76 E5	1.33	1.96	2.32	4.82	...	17.33
8	=	0 + 8	9	0.10	*	0.11	8.21 E7	0.50	1.66	2.21	4.55	...	17.32

Table A2: Extension of Table 4.2

s	=	$q+r$	order	β_{real}	β_{imag}	β^*_{imag}	$\kappa_\infty(\underline{C}_2)$	γ_2	γ_3	γ_4	γ_{10}	...	γ_∞
2	=	2 + 0	3	0.91	*	0.29	-	∞	∞	∞	∞	...	∞
3	=	3 + 0	4	0.59	0.60	0.61	-	∞	∞	∞	∞	...	∞
4	=	4 + 0	5	0.48	*	0.44	-	∞	∞	∞	∞	...	∞
5	=	5 + 0	6	0.44	*	0.44	-	∞	∞	∞	∞	...	∞
6	=	6 + 0	7	0.41	*	0.42	-	∞	∞	∞	∞	...	∞
7	=	7 + 0	8	0.40	*	0.40	-	∞	∞	∞	∞	...	∞
8	=	8 + 0	9	0.39	*	0.39	-	∞	∞	∞	∞	...	∞
2	=	1 + 1	3	3.00	1.49	1.51	1.00	4.00	4.00	4.00	4.00	...	4.00
3	=	2 + 1	4	0.54	0.64	0.65	1.00	9.00	9.00	9.00	9.00	...	9.00
4	=	3 + 1	5	0.11	*	0.12	1.00	16.00	16.00	16.00	16.00	...	16.00
5	=	4 + 1	6	*	*	*	1.00	25.00	25.00	25.00	25.00	...	25.00
6	=	5 + 1	7	*	*	*	1.00	36.00	36.00	36.00	36.00	...	36.00
7	=	6 + 1	8	*	*	*	1.00	49.00	49.00	49.00	49.00	...	49.00
8	=	7 + 1	9	*	*	*	1.00	64.00	64.00	64.00	64.00	...	64.00
2	=	0 + 2	3	∞	∞	∞	7.00	1.41	1.59	1.86	2.36	...	2.45
3	=	1 + 2	4	8.30	4.32	4.32	9.34	2.15	2.48	2.87	3.66	...	4.31
4	=	2 + 2	5	1.05	*	0.93	10.25	3.39	3.92	4.49	5.93	...	7.11
5	=	3 + 2	6	0.13	*	0.13	10.70	5.03	5.81	6.63	8.98	...	10.75
6	=	4 + 2	7	*	*	*	10.94	7.04	8.14	9.26	12.78	...	15.21
7	=	5 + 2	8	*	*	*	11.09	9.42	10.89	12.38	17.31	...	20.48
8	=	6 + 2	9	*	*	*	11.19	12.16	14.06	15.98	22.56	...	26.57
3	=	0 + 3	4	∞	∞	∞	18.06	1.41	1.82	2.21	3.03	...	3.64
4	=	1 + 3	5	17.18	*	9.02	23.82	1.81	2.32	2.62	4.08	...	4.94
5	=	2 + 3	6	1.97	*	1.89	26.93	2.45	3.08	3.47	5.80	...	7.03
6	=	3 + 3	7	0.18	*	0.19	28.75	3.28	4.05	4.60	8.00	...	9.68
7	=	4 + 3	8	*	*	*	29.89	4.26	5.24	5.95	10.62	...	12.87
8	=	5 + 3	9	*	*	*	30.65	5.41	6.62	7.52	13.63	...	16.56
4	=	0 + 4	5	∞	∞	∞	34.19	1.41	1.82	2.21	4.28	...	5.04
5	=	1 + 4	6	30.16	*	15.74	43.75	1.65	2.11	2.55	4.84	...	5.99
6	=	2 + 4	7	3.35	*	2.86	49.85	2.04	2.61	3.15	5.80	...	7.74
7	=	3 + 4	8	0.27	*	0.28	53.85	2.54	3.24	3.91	7.11	...	9.96
8	=	4 + 4	9	*	*	*	56.59	3.14	4.00	4.77	8.69	...	12.59
5	=	0 + 5	6	∞	∞	∞	55.38	1.41	1.82	2.21	4.44	...	6.29
6	=	1 + 5	7	47.80	*	24.92	68.93	1.57	2.02	2.44	4.70	...	6.87
7	=	2 + 5	8	5.23	*	4.57	78.48	1.84	2.36	2.85	5.40	...	8.39
8	=	3 + 5	9	0.40	*	0.43	85.26	2.19	2.80	3.38	6.35	...	10.33
6	=	0 + 6	7	∞	∞	∞	81.63	1.41	1.82	2.21	4.50	...	7.66
7	=	1 + 6	8	70.66	*	37.01	99.26	1.53	1.96	2.39	4.73	...	7.96
8	=	2 + 6	9	7.61	*	7.29	112.56	1.73	2.22	2.69	5.22	...	9.32
7	=	0 + 7	8	∞	∞	∞	112.94	1.41	1.82	2.21	4.52	...	8.94
8	=	1 + 7	9	99.27	*	52.43	134.71	1.50	1.93	2.34	4.71	...	8.89
8	=	0 + 8	9	∞	∞	∞	149.32	1.41	1.82	2.21	4.53	...	10.30