1991

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

Analysis of parallel diagonal-implicit iteration of Runge-Kutta methods

Department of Numerical Mathematics Report NM-R9114 August

CWI, nationaal instituut voor onderzoek op het gebied van wiskunde en informatica

CWI is the research institute of the Stichting Mathematisch Centrum, which was founded on February 11, 1946, as a non-profit institution aiming at the promotion of mathematics, computer science, and their applications. It is sponsored by the Dutch Government through the Netherlands organization for scientific research (NWO).

Analysis of Parallel Diagonal-Implicit Iteration of Runge-Kutta Methods

P.J. van der Houwen & B.P. Sommeijer

CWI
P.O. Box 4079, 1009 AB Amsterdam,
The Netherlands

In this paper, we analyse parallel, diagonally-implicit iteration of Runge-Kutta methods (PDIRK methods) for solving large systems of stiff equations on parallel computers. Like Newton-iterated backward differentiation formulas (BDFs), these PDIRK methods are such that in each step the (sequential) costs consists of solving a number of linear systems with the same matrix of coefficients and with the same dimension as the system of differential equations. Although for PDIRK methods the number of linear systems is usually higher than for Newton iteration of BDFs, the more computational intensive work of computing the matrix of coefficients and its LU decomposition is identical. The advantage of PDIRK methods over Newton-iterated BDFs is their unconditional stability (A-stability for Gauss-based methods and L-stability for Radau-based methods) for any order of accuracy.

1980 Mathematics Subject Classification: 65M10, 65M20.
1982 CR Categories: 5.17.
Key Words and Phrases: Diagonally implicit Runge-Kutta methods, parallelism.

1. Introduction

Consider the initial-value problem for systems of ordinary differential equations (ODEs) of dimension d

$$(1.1) \qquad y'(t) = f(t,y(t)), \quad y(t_0) = y_0, \quad y : \mathbb{R} \to \mathbb{R}^d, \quad f : \mathbb{R}x\mathbb{R}^d \to \mathbb{R}^d, \quad t_0 \le t \le t_{end}.$$

In this paper, we analyse integration methods based on iteration of implicit Runge-Kutta (RK) methods of collocation type. Such RK methods possess both a large step point order and a large stage order. Furthermore, by a suitable choice of the collocation parameters, these RK methods are unconditionally stable for any order of accuracy.

We shall employ the *diagonally-implicit* iteration-type methods proposed in [7, 8]. These methods are designed in such a way that a large number of the implicit systems to be solved can be processed in parallel, so that the number of systems that have to be solved *sequentially* is substantially reduced when implemented on multi-processor computers. As a reference method, we take the method based on the backward differentiation formulas (BDFs), which is considered as one of the best methods for sequential computers. The sequential computations (i.e., the computations that cannot be performed in parallel on a multi-processor system) of the parallel diagonal-implicitly iterated RK (PDIRK) methods are of the same nature as those of Newton-iterated BDFs, that is, in each step, both types of methods require the sequential solution of a number of linear systems with the same matrix of coefficients and with the same dimension as the system of differential equations. Although, this number of linear systems is usually higher for PDIRK methods than for Newton iteration of BDFs, the effort required for computing the Jacobian and the LU decomposition of the matrix of coefficients is identical. For large systems of equations, these computations are the more computational intensive work, so that the overall computation time is primarily determined by the number of Jacobian updates and LU decompositions. The advantage of PDIRK methods over Newton-iterated BDFs is their A-stability (Gauss and Lobatto IIIA correctors), strong A-stability (Lagrange correctors derived in [7]) or even L-stability (e.g. Radau IIA correctors) for high orders of accuracy. The property that unconditional stability can be combined with high orders reduces the number of integration steps (and therefore the number of Jacobian updates and LU decompositions) considerably.

2. PDIRK Methods

In this section we define PDIRK methods by specifying the RK corrector, the iteration scheme for solving the stage vector equation, the predictor formula, and the formula for the step-point values. The various families of PDIRK methods are determined by special choices of the iteration parameters occurring in the iteration scheme. In order to simplify the notations, the formulas are given for *scalar* ODEs. The extension to *systems* of ODEs is straightforward.

Report NM-R9114
ISSN 0169-0388
CWI
P.O. Box 4079, 1009 AB Amsterdam, The Netherlands

2.1. The corrector

We consider RK methods of the form

$$(2.1)\qquad Y - hAf(et_n + ch, Y) = ey_n + haf(t_n, y_n), \quad y_{n+1} = y_n + hb_0 f(t_n, y_n) + hb^T f(et_n + ch, Y), \quad c := a + Ae,$$

where b_0 is a scalar parameter, e is the vector with unit entries, $a = (a_i)$, $b = (b_i)$ and $c = (c_i)$ are k-dimensional vectors, and $A = (a_{ij})$ is a k-by-k matrix. In (2.1) we used the convention that for any given vectors $v=(v_j)$ and $t = (t_j)$, $f(t, v)$ denotes the vector with entries $f(t_j, v_j)$. We always assume that the matrix A is nonsingular. If the vector a or the parameter b_0 does not vanish, then (2.1) presents an (s = k+1)-stage RK method requiring k *implicit* stages and one explicit stage. If $a = 0$ and $b_0 = 0$, then (2.1) reduces to the general (s = k)-stage RK method with s implicit stages. For a discussion of the order of accuracy and the stage order of RK methods, we refer to e.g. [4] and [3]. In the sequel, the method (2.1) will be called the *corrector*.

2.2. The iteration scheme

The stage vector equation in (2.1) is solved by applying the diagonal iteration method studied in [8] and [7]. Let $Y^{(\mu)}$ denote the successive iterates, then we may define the (highly parallel) iteration process

$$(2.2)\qquad \begin{aligned} Y^{(1)} - hD\, f(et_n + ch, Y^{(1)}) &= ey_n + ahf(t_n, y_n) + hA\, f(et_n + c^*h, Y^{(0)}) - hD\, f(et_n + c^*h, Y^{(0)}), \\ Y^{(\mu)} - hD\, f(et_n + ch, Y^{(\mu)}) &= ey_n + ahf(t_n, y_n) + hA\, f(et_n + ch, Y^{(\mu-1)}) - hD\, f(et_n + ch, Y^{(\mu-1)}), \end{aligned}$$

where $\mu = 2, \ldots, m$, and where D is a diagonal matrix whose diagonal elements δ_i ($i = 1, \ldots, k$) are the iteration parameters which are assumed to be positive. The parameter vector c^* depends on the predictor formula used for computing $Y^{(0)}$ and serves to make the arguments of f consistent in the first iteration (see Section 2.4). The step point formula defining y_{n+1} and the predictor formula will be discussed in the Sections 2.3 and 2.4, respectively. Together, the predictor formula, the iteration scheme (2.2), and the step point formula determine the PDIRK method.

Each iteration in (2.2) requires the solution of k nonlinear systems which can be obtained by applying modified Newton iteration. We shall call this last iteration the *inner* iteration method and the iteration (2.2) the *outer* iteration method. Notice that in each outer iteration the k nonlinear systems can be solved in parallel, provided that k processors are available. Thus, the *sequential* costs per step consists of computing $Y^{(0)}$ and of solving m nonlinear systems of ODE dimension.

For particular choices of the predictor formula (e.g., explicit RK formulas) and for step point formulas as defined in Section 2.3, the PDIRK method as described above can be interpreted as a diagonally implicit RK (DIRK) method using mk diagonally implicit stages. Since the k stages in each outer iteration can be computed in parallel, we arrive at a DIRK method with m *sequential* diagonally implicit stages. These methods form a subclass of the much wider class of the PaRK methods investigated by Jackson and Nørsett [9, 10].

In [7] and [8] the performance of PDIRK methods was studied in the case where in each of the m outer iterations the inner iteration method was continued until convergence before starting the next outer iteration (this iteration strategy is also used in conventional DIRK methods). However, this strategy may be rather expensive if many iterations are needed to get the inner iteration converged. Moreover, it does not take into account the special structure of the method. The essential difference with conventional DIRK methods lies in the fact that the i-th component of each stage vector $Y^{(\mu)}$ is an approximation to the exact solution at the points $t_n + c_ih$. This implies that $Y^{(\mu-1)}$ furnishes an excellent initial approximation to the solution $Y^{(\mu)}$ to be obtained in the inner iteration process. As a consequence, each outer iteration needs only a few inner iterations. Furthermore, in first approximation, the convergence of the inner-outer iteration scheme and the stability of the PDIRK method does not depend on the number of inner iterations. This motivates our strategy to perform only one inner iteration per outer iteration, leading to the iteration process

$$(2.3a)\qquad \begin{aligned} [I - hDJ]\,(Y^{(0)} - Y^{(1)}) &= Y^{(0)} - [ey_n + ahf(t_n, y_n) + hA\, f(et_n + c^*h, Y^{(0)}) - hD\, f(et_n + c^*h, Y^{(0)})] \\ &\quad - hD\, f(t_n e + ch, Y^{(0)}), \\ [I - hDJ]\,(Y^{(\mu-1)} - Y^{(\mu)}) &= Y^{(\mu-1)} - [ey_n + ahf(t_n, y_n) + hA\, f(et_n + ch, Y^{(\mu-1)})], \quad \mu = 2, \ldots, m. \end{aligned}$$

Here, J denotes an approximation to the derivative of f at the point (t_ne, y_ne). Evidently, if (2.3a) converges, then $Y^{(\mu)}$ converges to Y. In fact, one may interpret (2.3a) as a modified Newton iteration scheme for solving Y from the stage vector equation in (2.1) employing a *diagonal approximation* to the Jacobian of $Y - hAf(et_n+ch, Y)$.

It may be useful to consider (2.3a) in the case of *systems* of ODEs. Then, the k components $Y_i^{(\mu)}$ of the stage vector iterate $Y^{(\mu)}$ have to satisfy the equations

$$\begin{aligned} [I - h\delta_iJ]\,(Y_i^{(0)} - Y_i^{(1)}) &= Y_i^{(0)} - [y_n + a_ihf(t_n, y_n) + h\sum_{j=1}^{k} a_{ij}\, f(t_n + c_j^*h, Y_j^{(0)}) - h\delta_i\, f(t_n + c_i^*h, Y_i^{(0)})] \\ &\quad - h\delta_i\, f(t_n + c_ih, Y_i^{(0)}), \end{aligned}$$

$$[I - h\delta_i J]\,(Y_i^{(\mu-1)} - Y_i^{(\mu)}) = Y_i^{(\mu-1)} - \Big[y_n + a_i h f(t_n,y_n) + h\sum_{j=1}^{k} a_{ij}\, f\big(t_n + c_j h, Y_j^{(\mu-1)}\big)\Big],\quad \mu = 2, \ldots, m,$$

where $i = 1, \ldots, k$ and where now J denotes an approximation to the Jacobian matrix of f at the point (t_n, y_n). Notice that this iteration scheme can be viewed as a modified Newton method for solving the stage vector equation employing a *block-diagonal approximation* to the Jacobian. Clearly, the k linear systems that are to be solved in each outer iteration step can be solved in parallel. Since each system has dimension equal to that of the system of ODEs, the computational complexity per step and per processor essentially consists of the computation of $Y_i^{(0)}$, the evaluation and LU-decomposition of the matrix $I - h\delta_i J$ (or its updating), m+2 evaluations of f, and m forward-backward substitutions. Of these costs, the evaluation and LU-decomposition of $I - h\delta_i J$ are the most time consuming, while the evaluations of f and the forward-backward substitutions are relatively cheap (notice that the iteration parameters δ_i are independent of μ in order to avoid repeated LU-decompositions of $I - h\delta_i J$ in the successive iterations). Thus, when basing a code on PDIRK methods, first of all the number of stepsize changes (which automatically requires new LU-decompositions) and the number of Jacobian updates should be minimized.

It is of interest to compare the sequential costs of PDIRK methods with the sequential costs of the celebrated BDF-based methods. If the BDFs are solved by using m modified Newton iterations, then the sequential costs in each step of the PDIRK methods and the Newton-iterated BDFs are almost identical. We expect that PDIRK methods need more iterations but, because of their higher order, less steps to produce some given accuracy. As explained above, evaluations of f and the forward-backward substitutions are relatively cheap, so that for modest values of m, the sequential costs per step of PDIRK methods are expected to be not much higher than those of the BDFs. The reduced number of steps required by the PDIRKs should make them superior to the BDFs.

2.3. The step-point values

Suppose that we adopt $Y^{(m)}$ as a sufficiently accurate approximation to the exact stage vector solution Y of the corrector (2.1). Then, the most natural way to approximate the step-point value y_{n+1} in (2.1) defines this value according to the formula

$$y_{n+1} = y_n + hb_0 f(t_n, y_n) + hb^T f(et_n + ch, Y^{(m)}). \tag{2.4}$$

However, the presence of the righthand side evaluations in this formula may give rise to loss of accuracy in the case of stiff problems (cf. [12]). This difficulty can be overcome by applying a similar approach as proposed in [6] for the implementation of implicit RK methods. Observing that the corrector (2.1) can be written in the form

$$y_{n+1} = y_n + b_0 hf(t_n,y_n) + b^T A^{-1}[Y - ey_n - ahf(t_n,y_n)],$$

provided that A is nonsingular, we can approximate the corrector solution y_{n+1} by the formula

$$y_{n+1} = y_n + b_0 hf(t_n,y_n) + b^T A^{-1}[Y^{(m)} - ey_n - ahf(t_n,y_n)], \tag{2.3b}$$

where $Y^{(m)}$ denotes the last computed approximation to Y. In many cases the corrector satisfies the relations of stiff accuracy, i.e., $c_k = 1$, $b_0 = a_k$ and $b^T A^{-1} = e_k^T$, so that (2.3b) reduces to $y_{n+1} = e_k^T Y^{(m)}$. In order to avoid confusion, we shall from now on denote the corrector solution and stage vector values obtained from y_n by u_{n+1} and U, respectively.

2.4. The predictor

In [8] we considered one-step predictors of the form

$$Y^{(0)} := ey_n + hE\, f(et_n, ey_n) + hB\, f\big(et_n + c^*h, Y^{(0)}\big), \tag{2.5a}$$

where B and E are k-by-k matrices. Of particular interest are the cases where E vanishes and where B is either the zero matrix yielding *last step-value predictors* (LSP) or B = D yielding *implicit Euler predictors* (IEP).

However, by using information from the preceding step, that is the values of y_n and the stage vector $Y^{(m)}$ computed in the last step, we can construct more accurate predictors. In order to indicate to which step a particular stage vector corresponds, we define $Y_n := Y^{(m)}$ if $Y^{(m)}$ corresponds to the step $[t_{n-1},t_n]$. Consider the two-step predictor

$$Y^{(0)} = VY_n + vy_n + hB\, f\big(et_n + c^*h, Y^{(0)}\big), \tag{2.6a}$$

where either B = O or B = D and where the matrix V and the vector v satisfy the usual consistency conditions (we shall assume that the vector v vanishes in the case of stiffly accurate correctors). The cases B = O and B = D will be referred to as the *extrapolation predictor* (EXP) and the *backward differentiation predictor* (BDP).

If B = D, then both (2.5a) and (2.6a) require the solution of k implicit relations. Similar to the strategy followed in solving the implicit relations in (2.2), we shall perform just one Newton iteration (notice that the righthand side derivatives required in the Newton iteration method are identical to those occurring in (2.3a)). In order to perform this Newton iteration we need an initial guess $Y^{(-1)}$ for $Y^{(0)}$. For the cases (2.5a) and (2.6a) we shall respectively use

$$Y^{(-1)} = ey_n,\ c^* = (E + B)e; \quad Y^{(-1)} = WY_n + wy_n,\ c^* = c,$$

where the W and w are to be determined (we shall assume that w vanishes in the case of stiffly accurate correctors, and that W = V, w = v in the case where B = O). If the corrector is based on collocation, then the matrix W and the vector w can be computed by extrapolating the collocation polynomial defined in $[t_{n-1},t_n]$ to the interval $[t_n,t_{n+1}]$ and can be expressed in terms of the Lagrange interpolation polynomials.

2.5. The iteration parameters

There are various options for choosing the number of iterations m, and the iteration parameters δ_i. In this paper, we consider three cases:

Option 1: fixed-number-of-iterations option
- the number of iterations is fixed and such that the orders of the PDIRK and corrector are equal
- the iteration parameters are chosen such that the stability region in the left halfplane is optimized.

Option 2: minimal-spectral-radius option
- the number of iterations is sufficiently large to closely approximate the corrector solution
- the iteration parameters are such that the spectral radius of the matrix $D^{-1}A - I$ is minimized.

Option 3: minimal-stiff-error-constant option
- the number of iterations is sufficiently large to closely approximate the corrector solution
- the iteration parameters are such that the principal stiff error constant of the PDIRK method is minimized.

Several families of methods constructed according to the fixed-number-of-iterations option were already considered in [8]. An interesting family considered in this paper possesses the stability functions investigated by Wolfbrandt [13] and uses constant iteration parameters δ_i determined by these stability functions. However, because of the fixed number of iterations, these methods are in fact DIRK methods and consequently, they have the disadvantage of possessing stage order q=1. In many stiff problems, such a low stage order may lead to reduced accuracies. In order to get insight into the extent of this accuracy reduction, we shall consider the magnitude of the stiff error constants for the 'fixed-number-of-iterations PDIRK methods' (see Section 4.2, Table 4.1).

For the explicit one-step predictor, [7] presents a number of PDIRK methods constructed according to the minimal-spectral-radius option. The effect of minimizing the spectral radius of the matrix $D^{-1}A - I$ is a strong damping of the stiff iteration error components. On the one hand, the number of iterations m should be sufficiently large to solve more or less the RK corrector, on the other hand, m should be sufficiently small to achieve that the (sequential) costs per step are not excessive when compared with those of the BDFs. In this paper, we shall investigate a few characteristics of the 'minimal-spectral-radius PDIRK methods' as a function of m. In particular, in Section 3 we consider the rate of convergence (Table 3.1) and the effect on the stability of the various predictors (Table 3.2), and in Section 4 we consider the magnitude of the principal stiff error constants (Tables 4.2 and 4.3).

Option 3 offers an alternative to option 2 and directly addresses the truncation error of PDIRK methods when applied to stiff systems. In this paper, we present preliminary results for the simple inhomogeneous test equation $y'(t) = \lambda y(t) + g(t)$. This approach is a special case of a more general treatment of minimizing stiff error constants which will be reported in [1].

3. Convergence and Stability

We shall investigate convergence and stability by means of the scalar test equation $y' = \lambda y$. Note that for this simple test equation the particular strategy used in the inner iterations is not relevant. For a rigorous convergence analysis of parallel RK methods containing the PDIRK methods of this paper we refer to Jackson and Nørsett [9, 10].

3.1. Rate of convergence

From (2.2) it can be deduced that the iteration error satisfies the recursion

$$(3.1) \qquad U - Y_{n+1} = Z(z)\left(U - Y^{(m-1)}\right) = \ldots = Z^m(z)\left(U - Y^{(0)}\right),\ \ Z(z) := zD[I - zD]^{-1}\,[D^{-1}A - I],\ \ z := \lambda h.$$

The region in the complex z-plane where $Z^m(z) \to 0$ for $m \to \infty$ will be called the *region of convergence*. We define the *iteration function* C of the PDIRK method by the spectral radius of Z(z), i.e.,

$$(3.2) \qquad C(z) := \rho(Z(z)) = \rho\left(zD[I - zD]^{-1}\,[D^{-1}A - I]\right).$$

Evidently, the region of convergence is determined by the set of points where $C(z) < 1$. The rate of convergence is larger as the norm of $C(z)$ is smaller in the region of relevant values of z. Thus, adopting the maximum norm, we are led to the minimization of $C(z)$ in this region. In this connection we introduce the following definition:

Definition 3.1. A PDIRK method is said to be *strongly A-convergent* if its iteration function $C(z) \le \eta < 1$ in the whole left halfplane $\mathrm{Re}\, z < 0$. If, in addition, $C(-\infty) = 0$, then the PDIRK method is called *L-convergent*. []

First we consider the constant δ_i case which is of interest in the case of fixed-number-of-iterations methods.

Theorem 3.1. If D has constant, positive diagonal elements, then minimization of $\rho(D^{-1}A - I)$ implies that the norm of $C(z)$ is minimized in (any region of) the left halfplane.

Proof. If $D = \delta I$, we may write $C(z) = |\delta z|\, \rho(A/\delta - I) \,/\, |1 - \delta z|$. In the left halfplane, the maximum of the function $|\delta z/(1 - \delta z)|$ does not depend on δ, provided that $\delta > 0$. Hence, the norm of $C(z)$ is minimized if $\rho(D^{-1}A - I)$ is minimized. []

In the case where D does not have constant diagonal entries, we cannot derive such a simple expression for $C(z)$, and a numerical search is needed to find the matrix D that minimizes the norm of $C(z)$ in the left halfplane. However, our numerical experiments revealed that also in the nonconstant δ_i case the minimization of $\rho(D^{-1}A - I)$ yields fast converging PDIRK methods and that $\|C\| := \max\{C(z): \mathrm{Re}(z) \le 0\}$ is considerably smaller than in the constant δ_i case.

Example 3.1. We consider an example of the fixed-number-of-iterations methods studied in [8] which is based on the third-order Radau IIA corrector. For

$$m = 3,\quad A = \frac{1}{12}\begin{pmatrix} 5 & -1 \\ 9 & 3 \end{pmatrix},\quad D = \delta I,\quad \delta = 0.43586650$$

this leads to a third-order, L-stable PDIRK method. The convergence function associated with this method is given by $C(z) = |\delta z|\, \rho(A/\delta - I) \,/\, |1 - \delta z|$ where $\rho(A/\delta - I) = \delta^{-1}\,(1/6 - 2\delta/3 + \delta^2)^{1/2}$. Setting $\delta = 0.43586650$ we find that $C(z) < 0.59$ in the whole left halfplane. Among the methods with $D = \delta I$ this method is almost optimal (the minimizing value is given by $\delta = 1/2$ leading to $C(z) < \sqrt{1/3} \approx 0.577$).

Next, we consider the case where D minimizes $\rho(D^{-1}A - I)$. In [7] it was shown that the method can be made L-convergent (i.e., has vanishing $\rho(D^{-1}A - I)$) for $\delta_1 = (4 - \sqrt{6})/6$ and $\delta_2 = (4 + \sqrt{6})/10$. The corresponding matrix $Z(z)$ is easily computed, yielding $\|C\| \approx 0.262$. []

Table 3.1 lists the $\|C\|$-values for a number of minimal-spectral-radius PDIRK methods. These methods are based on Radau IIA correctors and on the so-called Lagrange correctors derived in [7]. The Lagrange methods are strongly A-stable, stiffly accurate collocation methods which are completely determined by the collocation vector c (see Table 3.1). Their stage order is one higher than that of the Radau IIA methods which was achieved by using one explicit and k implicit stages. However, they do not possess the superconvergence property of the Radau methods, so that the computation of the *nonstiff* solution components is considerably less accurate.

For the Radau IIA and Lagrange correctors with k implicit stages, the iteration parameters are contained in the matrices D_{kR} and D_{kL}:

(3.3a)
$$D_{2R} = \frac{1}{30}\begin{pmatrix} 20-5\sqrt{6} & 0 \\ 0 & 12+3\sqrt{6} \end{pmatrix},\quad D_{3R} = \begin{pmatrix} \frac{4365}{13624} & 0 & 0 \\ 0 & \frac{1032}{7373} & 0 \\ 0 & 0 & \frac{1887}{5077} \end{pmatrix},\quad D_{4R} = \begin{pmatrix} \frac{3055}{9532} & 0 & 0 & 0 \\ 0 & \frac{531}{5956} & 0 & 0 \\ 0 & 0 & \frac{1471}{8094} & 0 \\ 0 & 0 & 0 & \frac{1848}{7919} \end{pmatrix},$$

(3.3b)
$$D_{2L} = \begin{pmatrix} \frac{3}{4(\sqrt{2}+1)} & 0 \\ 0 & \frac{1}{6(\sqrt{2}-1)} \end{pmatrix},\quad D_{3L} = \begin{pmatrix} \frac{2246}{10669} & 0 & 0 \\ 0 & \frac{2537}{8794} & 0 \\ 0 & 0 & \frac{3026}{8923} \end{pmatrix},\quad D_{4L} = \begin{pmatrix} \frac{5147}{38467} & 0 & 0 & 0 \\ 0 & \frac{1983}{17459} & 0 & 0 \\ 0 & 0 & \frac{3197}{14090} & 0 \\ 0 & 0 & 0 & \frac{3086}{12339} \end{pmatrix}.$$

Table 3.1 shows that these methods can all be made strongly A-convergent, and that only the methods based on a *two*-stage corrector are L-convergent (see also [7]). Furthermore, we observe that the rates of convergence of the Lagrange-based methods are slightly better. Hence, together with their increased stage order, the Lagrange correctors

seem to be attractive alternatives to the Radau correctors in problems where the order of accuracy is determined by the stage order. However, in problems where, apart from the stage order, the nonstiff (or, classical) order is important, the superconvergent Radau correctors are to be preferred. As to the ||C||-values given in Table 3.1, it should be remarked that these are 'worst case' values, that is, in actual computation, where the relevant values of z are located in a restricted region of the left halfplane, the corresponding bound on C(z) may be much smaller.

Table 3.1. ||C||-values for minimal-spectral-radius PDIRK methods based on Radau IIA and Lagrange correctors.

Corrector	k		||C||	strongly A-convergent	L-convergent
Radau IIA	2		0.262	yes	yes
	3		0.401	yes	no
	4		0.527	yes	no
Lagrange	2	$c = (3/4,1)^T$	0.182	yes	yes
	3	$c = (7/12,5/6,1)^T$	0.403	yes	no
	4	$c = (1/6,7/12,11/12,1)^T$	0.404	yes	no

3.2. Region of stability

In order to investigate the stability properties of PDIRK methods we have to specify the predictor formula. The stability of PDIRK methods using the one-step predictor (2.5) was extensively discussed in [8] for the case where y_{n+1} is defined by the (2.4). For the case (2.3b) considered in this paper, we have the following theorems:

Theorem 3.2. For the equation $y' = \lambda y$ the PDIRK solution generated by {(2.3), (2.5)} satisfies the recursion

$$y_{n+1} = R_m(z)y_n, \quad R_m(z) := R(z) - E_m(z),$$

$$E_m(z) := b^TA^{-1}Z^m(z)\big([I - zA]^{-1}[e + za] - [I - zB]^{-1}[I + zE]e\big), \quad R(z) := 1 + zb_0 + zb^T[I - zA]^{-1}[e + za].$$

Here, R(z) is the stability function of the corrector reducing to $R(z) = e_k^T[I - zA]^{-1}[e + za]$ in the stiffly accurate case.

Proof. From the relations

$$(3.4) \qquad Y^{(0)} = [I - zB]^{-1}[I + zE]ey_n, \quad U = [I - zA]^{-1}[e + za]y_n,$$

it follows that

$$(3.5) \qquad U - Y_{n+1} = Z^m(z)\big(U - Y^{(0)}\big) = Z^m(z)\big([I - zA]^{-1}[e + za] - [I - zB]^{-1}[I + zE]e\big)y_n.$$

Hence, from the step point formula (2.3b) we obtain

$$(3.6) \qquad u_{n+1} - y_{n+1} = b^TA^{-1}\big(U - Y_{n+1}\big) = b^TA^{-1}Z^m(z)\big([I - zA]^{-1}[e + za] - [I - zB]^{-1}[I + zE]e\big)y_n.$$

Furthermore, introducing the stability function R(z) of the corrector, we may write

$$(3.7) \qquad u_{n+1} = R(z)y_n,$$

where R(z) is defined in the theorem. From (3.6) and (3.7) the assertion of the theorem is immediate. []

Theorem 3.3. For the equation $y' = \lambda y$ the PDIRK solution generated by {(2.3), (2.6)} satisfies the recursion

$$\begin{pmatrix} Y_{n+1} \\ y_{n+1} \end{pmatrix} = M_m(z) \begin{pmatrix} Y_n \\ y_n \end{pmatrix},$$

where $M_m(z)$ is the amplification matrix

$$M_m(z) := \begin{pmatrix} I & 0 \\ -b^TA^{-1} & 1 \end{pmatrix}^{-1} \begin{pmatrix} Z^m(z)[I-zB]^{-1}V & [I-Z^m(z)][I-zA]^{-1}[e+za] + Z^m(z)[I-zB]^{-1}v \\ 0^T & 1 + b_0z - b^TA^{-1}[e+za] \end{pmatrix}.$$

Proof. By means of the equation for U given in (3.4), relation (3.5) and

(3.8) $\quad Y^{(0)} = [I - zB]^{-1}[VY_n + vy_n]$

we derive that

(3.9) $\quad Y_{n+1} = Z^m(z)\,[I - zB]^{-1}VY_n + \left([I - Z^m(z)][I - zA]^{-1}[e + za] + Z^m(z)\,[I - zB]^{-1}v\right)y_n.$

Together with the step-point formula (2.3b) the one-step recursion of the theorem is easily obtained. []

With the amplification matrix $M_m(z)$ we associate the *stability function*

(3.10) $\quad R_m(z) := \rho(M_m(z)),$

where $\rho(M_m)$ denotes the spectral radius of the matrix M_m. The region in the complex z-plane where $R_m(z) < 1$ for all m greater than or equal to some given integer m_{crit} will be called the *region of stability associated with* m_{crit}.

Table 3.2a. Values of m_{crit} of minimal-spectral-radius PDIRK methods for various PC pairs

Corrector		LSP	EXP	IEP	BDP
Radau IIA	k = 2	1	1	1	1
	k = 3	5	5	2	4
	k = 4	7	7	4	7
Lagrange	k = 2	2	2	2	2
	k = 3	3	3	3	3
	k = 4	6	7	5	6

Table 3.2b. Values of $\alpha=\alpha(m)$ of minimal-spectral-radius PDIRK methods for various PC pairs

Predictor	Corrector	k	m=1	m=2	m=3	m=4	m=5	m=6	m=7
LSP	Radau IIA	2	90						
EXP			90						
IEP			90						
BDP			90						
LSP		3	*	*	81.9	89.94	90		
EXP			*	*	64.7	88.7	90		
IEP			87.5	90					
BDP			65.0	81.8	88.4	90			
LSP			*	*	*	40.3	80.5	88.5	90
EXP			*	*	*	*	70.3	84.2	90
IEP		4	60.2	75.9	86.1	90			
BDP			43.0	14.6	67.1	78.2	84.6	88.6	90
LSP	Lagrange	2	*	90					
EXP			*	90					
IEP			86.5	90					
BDP			89.82	90					
LSP		3	*	*	90				
EXP			*	*	90				
IEP			77.2	*	90				
BDP			83.4	*	90				
LSP		4	*	*	*	60.8	86.7	90	
EXP			*	*	*	*	73.0	88.0	90
IEP			51.6	*	*	86.5	90		
BDP			48.8	*	*	79.9	87.6	90	

For future reference, we have computed the value of m_{crit} for a number of predictor-corrector (PC) pairs. For the correctors we again chose the Radau IIA methods and the Lagrange methods of Section 3.1. The predictors are those defined in Section 2.4 and the matrices D are defined according to the minimal-spectral-radius option (see (3.3)). Table 3.2a shows that m_{crit} increases if the number of stages of the corrector increases. However, in actual computation, the minimal number of iterations may be much smaller because many stiff problems require only A(α)-stability. This means that automatic codes based on PDIRK methods are likely to choose the number of iterations not larger than necessary to ensure a stable performance. Table 3.2b presents the corresponding angles α (in degrees) as a function of m (lack of A(0)-stability is indicated by $*$). The results illustrate the favourable A(α)-stability characteristics of minimal-spectral-radius PDIRK methods after only a few iterations. In general, the implicit predictors IEP and BDP possess (of course) larger stability angles α than the explicit predictors LSP and EXP, even if we take into account that the implicit predictors require extra computational effort roughly comparable with an additional iteration. Furthermore, if we compare IEP and BDP, then IEP has the best stability characteristics (in particular for Radau-based methods). However, the overall efficiency will be reduced because of its low-order of accuracy. Therefore, we drop the low-order predictors LSP and IEP and recommend either the EXP or BDP predictor.

4. The Error Functions for the Linear Inhomogeneous Test Equation

The following theorem presents a result for general RK methods derived in [2]:

Theorem 4.1. For RK methods the global error e_n when applied to the test equation $y'(t) = \lambda y(t) + g(t)$ satisfies

$$e_{n+1} = R(z)e_n + \sum_{j=q+1} Q_j(z)\, h^j\, y_{ex}^{(j)}(t_n), \quad Q_j(z) := \frac{1}{j!}\,[1 - jb^T c^{j-1}] + \frac{1}{j!} z b^T [I - zA]^{-1}\,[c^j - jAc^{j-1}],$$

where $y_{ex}(t)$ denotes the exact solution of the test equation, R(z) is the stability function of the RK method and q is its stage order (i.e., the largest integer such that $1 - jb^T c^{j-1} = c^j - jAc^{j-1} = 0$ for $j = 1, \ldots, q$). []

We shall prove a similar theorem for PDIRK methods employing one-step predictors. As before, the simplicity of the test equation $y'(t) = \lambda y(t) + g(t)$ implies that the particular strategy used in the inner iteration process is not relevant.

In the following, y(t) denotes the locally exact solution at t_n, i.e., $y_n = y(t_n)$. It is straightforwardly verified that for the linear inhomogeneous equation the recursion (3.5) changes to

$$(4.1) \qquad U - Y_{n+1} = Z^m(z)\big(U - Y^{(0)} + h\, z^{-1}\,[g(t_n e + hc) - g(t_n e + hc^*)]\big).$$

Assuming that g is sufficiently differentiable, we may write for any fixed vector v

$$(4.2) \qquad g(t_n e + hv) = \sum_{j=0} \frac{1}{j!}\,(hv)^j\, g^{(j)}(t_n) = \frac{1}{h} \sum_{j=0} \frac{1}{j!}\, h^j\, y^{(j)}(t_n)\,[jv^{j-1} - zv^j].$$

Hence,

$$h\,[g(t_n e + hc) - g(t_n e + hc^*)] = \sum_{j=1} \frac{1}{j!}\,\gamma_j(z)\, h^j\, y^{(j)}(t_n), \quad \gamma_j(z) := jc^{j-1} - zc^j - j(c^*)^{j-1} + z(c^*)^j.$$

Furthermore, it follows from (2.1) that

$$U = [I - zA]^{-1}\,[y(t_n)e + hy'(t_n)a + hAg(t_n e + hc)],$$

so that

$$(4.3a) \qquad U = y_n e + \sum_{j=1} \frac{1}{j!}\, c_j(z)\, h^j\, y^{(j)}(t_n), \quad c_1(z) := c,\ c_j(z) := [I - zA]^{-1}A\,[jc^{j-1} - zc^j],\ j \geq 2.$$

4.1. One-step predictors

Let us assume that $Y^{(0)}$ is provided by a one-step formula, then it can also be expanded in terms of a similar Taylor series with coefficients $c_j^*(z)$:

$$(4.3b) \qquad Y^{(0)} = y_n e + \sum_{j=1} \frac{1}{j!}\, c_j^*(z)\, h^j\, y^{(j)}(t_n).$$

Thus,

$$(4.4)\qquad U - Y_{n+1} = Z^m(z) \sum_{j=1} q_j(z)\, h^j\, y^{(j)}(t_n),\quad q_1(z) := c^* - c_1^*(z),\quad q_j(z) := \frac{1}{j!}\,[c_j(z) - c_j^*(z) + z^{-1}\gamma_j(z)],\ j \ge 2.$$

Assuming that $c_1^*(z)$ does not depend on z, we may choose in (2.3) $c^* = c_1^*$ so that $q_1(z)$ vanishes. Using the relation

$$y^{(j)}(t_n) = y_{ex}^{(j)}(t_n) + \lambda^j[y(t_n) - y_{ex}(t_n)] = y_{ex}^{(j)}(t_n) + \lambda^j[y_n - y_{ex}(t_n)],$$

the iteration error (4.4) can be expanded in terms of derivatives of the exact solution. We obtain

$$(4.4')\qquad U - Y_{n+1} = Z^m(z) \sum_{j=2} q_j(z)\, \left(z^j\,[y_n - y_{ex}(t_n)] + h^j\, y_{ex}^{(j)}(t_n)\right).$$

As

$$(4.5)\qquad u_{n+1} - y_{n+1} = b^T A^{-1}\,[U - Y_{n+1}],$$

we find

$$(4.6a)\qquad u_{n+1} - y_{n+1} = S_m(z)[y_n - y_{ex}(t_n)] + \sum_{j=2} Q_{mj}(z)\, h^j\, y_{ex}^{(j)}(t_n),$$

$$(4.6b)\qquad S_m(z) := b^T A^{-1}\, Z^m(z) \sum_{j=2} q_j(z)\, z^j,\quad Q_{mj}(z) := b^T A^{-1}\, Z^m(z) q_j(z).$$

Applying Theorem 4.1 to the corrector at the point t_n with $e_n = y_n - y_{ex}(t_n)$ yields

$$(4.7)\qquad u_{n+1} - y_{ex}(t_{n+1}) = R(z)[y_n - y_{ex}(t_n)] + \frac{1}{(q+1)!} z b^T [I - zA]^{-1} A\,[c^{q+1} - (q+1)Ac^q]\, h^{q+1} y_{ex}^{(q+1)}(t_n) + O(h^{q+2}),$$

hence,

$$y_{n+1} - y_{ex}(t_{n+1}) = y_{n+1} - u_{n+1} + u_{n+1} - y_{ex}(t_{n+1}) = y_{n+1} - u_{n+1} + R(z)\,[y_n - y_{ex}(t_n)] + O(h^{q+1}).$$

Thus, using (4.6) we obtain

$$(4.8)\qquad y_{n+1} - y_{ex}(t_{n+1}) = (R(z) - S_m(z))\,[y_n - y_{ex}(t_n)] - \sum_{j=2} Q_{mj}(z)\, h^j\, y_{ex}^{(j)}(t_n) + O(h^{q+1}).$$

The functions $Q_{mj}(z)$ will be called the *error functions* of the PDIRK method.

Finally, we show that the function $R(z) - S_m(z)$ is identical with the stability function R_m of the PDIRK method. For that purpose, we consider the particular case where the inhomogeneous term g vanishes. It is easily verified that we then may write

$$(4.9)\qquad y_{n+1} - y_{ex}(t_{n+1}) = R_m(z)\,[y_n - y_{ex}(t_n)] + (R_m(z) - e^z)\, y_{ex}(t_n),\quad g = 0.$$

Now, suppose that the initial value y_0 tends to zero. Then, $y_{ex}(t)$ also tends to zero. Since (4.8) holds for vanishing g too, it follows that $R_m(z) = R(z) - S_m(z)$. Notice that in the case of the predictor (2.5) the functions $S_m(z)$ and $E_m(z)$ as defined in Theorem 3.2 are apparently identical. Thus, we have proved the following PDIRK analogue of Theorem 4.1:

Theorem 4.2. For one-step predictors possessing the expansion (4.3b) with $c^* = c_1^*$ the global error of PDIRK methods when applied to the test equation $y'(t) = \lambda y(t) + g(t)$ satisfies the recursion

$$y_{n+1} - y_{ex}(t_{n+1}) = R_m(z)\,[y_n - y_{ex}(t_n)] - \sum_{j=2} Q_{mj}(z)\, h^j\, y_{ex}^{(j)}(t_n) + O(h^{q+1}),$$

$$R_m(z) = R(z) - S_m(z),\quad S_m(z) := b^T A^{-1}\, Z^m(z) \sum_{j=2} q_j(z)\, z^j,\quad Q_{mj}(z) := b^T A^{-1}\, Z^m(z) q_j(z),$$

where q is the stage order of the corrector, and R(z) and $R_m(z)$ are the stability functions of the corrector and the PDIRK method, respectively. []

This theorem shows that the stage order of PDIRK methods is only one, unless the error function $Q_{m2}(z)$ is identically zero for the m-value used (this is not surprising because formally PDIRK methods are just DIRK methods which are known to have stage order one). However, as all error functions $Q_{mj}(z)$ contain the factor $Z^m(z)$, their maximal values $|Q_{mj}|$ are expected to decrease rapidly with m in any region of the left halfplane, so that effectively the stage order shown in actual computation is much higher.

The following corollary presents an explicit expression of Q_{mj} for the predictor (2.5).

Corollary 4.1. For the predictor (2.5) the error functions are given by

$$Q_{mj}(z) := \frac{1}{j!}\, b^TA^{-1}\, Z^m(z)z^{-1}\Big(jc^{j-1} - [I - zB]^{-1}[j(c^*)^{j-1} - z(c^*)^j]\Big),\quad j = 2, \ldots, q,$$

where $c^* := (B + E)e$.

Proof. In the case (2.5) the expansion (4.3b) becomes

$$Y^{(0)} = [I - zB]^{-1}\left([I + zE]\, y(t_n)e + hEg(t_n)e + hBg(t_ne + hc^*)\right)$$

$$= y(t_n)e + [I - zB]^{-1}\left(E\, hy'(t_n)e + B\sum_{j=1}\frac{1}{j!}\, h^j\, y^{(j)}(t_n)\,[j(c^*)^{j-1} - z(c^*)^j]\right),$$

so that

$$c_1^*(z) = [I - zB]^{-1}\,(Ee + Be - zBc^*) = c^* = (B + E)e,\quad c_j^*(z) = [I - zB]^{-1}\, zB\,[jz^{-1}(c^*)^{j-1} - (c^*)^j],\ j \ge 2.$$

By virtue of Theorem 4.2 we may write

$$Q_{mj}(z) = b^TA^{-1}\, Z^m(z)q_j(z) = \frac{1}{j!}\, b^TA^{-1}\, Z^m(z)\,[c_j(z) - c_j^*(z) + z^{-1}\gamma_j(z)] =$$

$$= \frac{1}{j!}\, b^TA^{-1}\, Z^m(z)z^{-1}\Big((I - zA)^{-1}[jc^{j-1} - zc^j] - [I - zB]^{-1}[j(c^*)^{j-1} - z(c^*)^j]\Big)$$

By means of the simplifying condition C(q) associated with (2.1) (cf. [3]), we obtain the relation $jAc^{j-1} = c^j$ for $j = 2, \ldots, q$ which leads to the result of the corollary. []

4.2. Last step-value predictor with constant iteration parameters

In the case of the predictor LSP (predictor (2.5) with $B = E = O$) with constant iteration parameters ($D = \delta I$), the error functions $Q_{mj}(z)$ can be factorized into factors that depend on z and factors that do not depend on z. This enables us to derive an explicit upper bound for $Q_{mj}(z)$.

Theorem 4.3. Let $D=\delta I$ and let the predictor be given by (2.5). Then the error function bound in a region $\mathbb{R}$ is given by

$$|Q_{mj}|_{\mathbb{R}} = \frac{1}{(j-1)!}\, d(m)\, |b^TA^{-1}D\,(D^{-1}A - I)^m\, c^{j-1}|,\quad j = 2, \ldots, q;\quad d(m) := \left|\frac{(\delta z)^{m-1}}{(1 - \delta z)^m}\right|_{\mathbb{R}}.$$

If $\mathbb{R}$ is the infinite wedge defined by $\mathbb{W} := \{z: \pi/2 \le \phi \le \arg(z) \le \pi,\ -\pi \le \arg(z) \le -\phi\}$, then

$$d(m) = \frac{x_m^{m-1}}{\left(m(1 - x_m\cos(\phi))\right)^{m/2}},$$

where x_m is the positive root of the equation $x^2 - (2 - m)\, x\cos(\phi) - m + 1 = 0$.

Proof. The expression for the error bound $|Q_{mj}|_{\mathbb{R}}$ immediately follows from Corollary 4.1. In order to derive an expression for the function d(m) we first observe that

$$\left|\frac{z}{1 - z}\right| = \frac{|z|}{\sqrt{1 - 2|z|\cos(\arg(z)) + |z|^2}},$$

where $\pi/2 \le \arg(z) \le \pi$ or $-\pi \le \arg(z) \le -\pi/2$. Hence,

$$\left| \frac{z^{m-1}}{(1-z)^m} \right| = \frac{|z|^{m-1}}{[1 - 2|z|\cos(\arg(z)) + |z|^2]^{m/2}},$$

Since the function $z^{m-1}(1-z)^{-m}$ is analytic, its maximum value in $\mathbb{W}$ is assumed at a point on the line $\arg(z)=\phi$. An elementary calculation reveals that the modulus of this point is given by the positive root x_m of the quadratic equation $x^2 - (2 - m)\cos(\phi)x - m + 1 = 0$. This leads us to the bound d(m) given in the theorem. []

This theorem shows that in the case where the relevant z-values are in an infinite wedge $\mathbb{W}$, the optimal choice of the matrix $D = \delta I$ does not depend on $\mathbb{W}$. Furthermore, the function d(m) is slowly varying with m. This can be concluded from the extreme cases where $\mathbb{R}$ is either only the negative axis or the whole left halfplane. We then have, respectively, $x_m=m-1$ and $x_m=\sqrt{m-1}$, which yields

$$d(m) = \frac{1}{m-1}\left(1 - \frac{1}{m}\right)^m \quad \text{and} \quad d(m) = \frac{1}{\sqrt{m-1}}\left(1 - \frac{1}{m}\right)^{m/2}.$$

Thus, within a few iterations the function d(m) slowly converges to zero.

It is of interest to compare the error functions $Q_j(z)$ of conventional DIRK methods (cf. Theorem 4.1) with the error functions $Q_{mj}(z)$ of PDIRK methods. Table 4.1 presents a comparison for two conventional Nørsett-DIRK methods [11] and a few L-stable, fixed-number-of-iterations PDIRK methods constructed according to option 1 [8]. In this table, k denotes the number of processors needed, p* is the order of the method, and m denotes the number of sequential stages per step (both for the Nørsett-DIRK and PDIRK methods). Clearly, the PDIRK methods possess considerably smaller error bounds.

Table 4.1. Values of $|Q_j|_{\mathbb{R}}$ and $|Q_{mj}|_{\mathbb{R}}$ with $\mathbb{R} = \{z: \text{Re } z \le 0\}$ for the Nørsett-DIRK methods and fixed-number-of-iterations PDIRK methods.

Method / PC pair	k	δ_i	m	p*	j=2	j=3	j=4	j=5
Nørsett-DIRK	1		2	3	0.144	0.076	0.024	0.0055
{LSP, Radau IIA}	2	0.43586650	3	3	0.024	0.015	0.005	0.0012
{LSP, Lagrange}	2	0.43586650	3	3	0.038	0.015	0.005	0.0012
Nørsett-DIRK	1		3	4	0.112	0.054	0.015	0.0040
{LSP, Radau IIA}	3	0.278053841	5	5	0.019	0.006	0.0014	0.0003
{LSP, Lagrange}	3	0.572816063	4	4	0.046	0.013	0.0001	0.0012
{LSP, Lagrange}	4	0.278053841	5	5	0.025	0.005	0.0001	0.0001

4.3. Minimal-spectral-radius PDIRK methods

Table 4.2 lists values of $|Q_{mj}|_{\mathbb{R}}$ with $\mathbb{R} = \{z: \text{Re } z \le 0\}$ for minimal-spectral-radius PDIRK methods (option 2), based on {LSP, Radau IIA} pairs and using the iteration parameters given in (3.3). It turns out that for $m > p^*$ the error constants decrease by an almost constant reduction factor r as m increases by 1 and that they are substantially smaller than those of the fixed-number-of-iterations PDIRK methods of Table 4.1 (notice that r is almost independent of j).

Table 4.2. Values of the error constants for minimal-spectral-radius PDIRK methods.

PC pair	k	m	p*	j=2	j=3	j=4	j=5
{LSP, Radau IIA}	2	2	2	0.0249	0.0263	0.0102	0.0027
		3	3	0.0060	0.0062	0.0024	0.0006
		...	3	r ≈ .25	r ≈ .25	r ≈ .25	r ≈ .25
{LSP, Radau IIA}	3	3	3	0.0360	0.0086	0.0027	0.00076
		4	4	0.0138	0.0031	0.0009	0.00025
		5	5	0.0052	0.0012	0.0003	0.00009
		...	5	r ≈ .40	r ≈ .38	r ≈ .39	r ≈ .38
{LSP, Radau IIA}	4	5	5	0.0153	0.00098	0.000031	0.00004
		6	6	0.0079	0.00051	0.000016	0.00002
		7	7	0.0041	0.00027	0.000008	0.00001
		...	7	r ≈ .50	r ≈ .52	r ≈ .50	r ≈ .52

For future reference, we give a survey of the *principal stiff error constants* $|Q_{m2}|_{\mathbb{R}}$ with $\mathbb{R} = \{z: \mathrm{Re}\, z \le 0\}$ for a number of PC pairs. In Table 4.3, p denotes the order of the corrector and the order of the iterated method is in all cases given by $p^* = \min\{p,m\}$. From these results we conclude that the explicit predictor LSP leads to slightly smaller principal error constants than the implicit predictor IEP, provided that we count the application of IEP as an additional iteration. Furthermore, the Lagrange-based methods show considerably smaller error constants. However, we should bear in mind that the *nonstiff* error constants of the Radau-based methods decrease much faster than those of the Lagrange-based methods because of the high (nonstiff) orders of the Radau correctors. Finally, note that the reduction factors are very close to the ||C||-values listed in Table 3.1

Table 4.3. Values of the principal error constant for minimal-spectral-radius PDIRK methods.

Method	k	p	m=k	m=k+1	m=k+2	r
{LSP, Radau IIA}	2	3	0.025	0.0060	0.0015	0.25
{IEP, Radau IIA}	2	3	0.024	0.0059	0.0015	0.25
{LSP, Lagrange}	2	3	0.013	0.0023	0.0004	0.18
{IEP, Lagrange}	2	3	0.006	0.0011	0.0002	0.18
{LSP, Radau IIA}	3	5	0.036	0.0138	0.0052	0.40
{IEP, Radau IIA}	3	5	0.014	0.0053	0.0020	0.41
{LSP, Lagrange}	3	4	0.008	0.0034	0.0014	0.40
{IEP, Lagrange}	3	4	0.004	0.0018	0.0007	0.40
{LSP, Radau IIA}	4	7	0.027	0.0153	0.0079	0.50
{IEP, Radau IIA}	4	7	0.017	0.0088	0.0044	0.50
{LSP, Lagrange}	4	5	0.022	0.0092	0.0037	0.40
{IEP, Lagrange}	4	5	0.013	0.0054	0.0021	0.40

5. Concluding Remarks

In this paper, we have studied special characteristics, such as the rate of convergence, the (linear) stability, and the stiff error constants, of PDIRK methods based on Radau IIA and Lagrange correctors using various types of iteration parameters and predictors. The minimal-spectral-radius methods turn out to be either comparable or superior to fixed-number-of-iterations methods. Confining our considerations to minimal-spectral-radius methods, the following conclusions can be drawn from our analysis:

Rate of convergence:	Lagrange correctors are superior to Radau corrector for $k = 2$ or $k = 4$ For $k = 3$, these correctors are comparable.
Linear stability:	Lagrange correctors are slightly superior to Radau correctors The implicit predictors IEP and BDP are superior to explicit predictors EXP and LSP.
Order reduction:	Lagrange correctors are superior to Radau correctors (both with respect to the stage order and the magnitude of the error constants) The explicit predictor LSP is slightly superior to the implicit predictor IEP.
Nonstiff error constants:	The two-stage Radau corrector is comparable with the two-stage Lagrange corrector Radau correctors are by far superior to Lagrange correctors for $k > 2$ The predictors EXP and BDP are by far superior to the predictors LSP and IEP.

By these conclusions, we are led to recommend PDIRK methods using an {EXP, Radau} PC pair and the minimal-spectral-radius iteration strategy as the most efficient in the class of PDIRK methods.

References

[1] **Burrage, K., van der Houwen, P.J., Hundsdorfer, W. & Sommeijer, B.P.**: Analysis of the error constants of parallel Runge-Kutta methods for stiff ODEs, in preparation.

[2] **Burrage, K., Hundsdorfer, W. & Verwer, J.G. (1986)**: A study of B-convergence of Runge-Kutta methods, Computing 36, 17-34.

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

[4] **Dekker, K. & Verwer, J.G. (1984)**: Stability of Runge-Kutta methods for stiff nonlinear differential equations, North-Holland, Amsterdam.

[5] **Hairer, E., Nørsett, S.P. & Wanner, G. (1987)**: Solving ordinary differential equations, I. Nonstiff problems, Springer-Verlag, Berlin.

[6] **Hairer, E. & Wanner, G. (1991)**: Solving ordinary differential equations, II. Stiff and differential-algebraic problems, Springer-Verlag, Berlin.

[7] **Houwen, P.J. van der & Sommeijer, B.P. (1991)**: Iterated Runge-Kutta methods on parallel computers, to appear in SIAM J. Sci. Stat. Comp..

[8] **Houwen, P.J. van der, Sommeijer, B.P. & Couzy, W. (1992)**: Embedded diagonally implicit Runge-Kutta algorithms on parallel computers, to appear in Math. Comp..

[9] **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.

[10] **Jackson, K.R. & Nørsett, S.P.**: The potential for parallelism in Runge-Kutta methods, Part II: RK predictor-corrector formulas, in preparation.

[11] **Nørsett, S.P. (1974)**: Semi-explicit Runge-Kutta methods, Report Mathematics and Computation No.6/74, Depart. of Mathematics, University of Trondheim.

[12] **Shampine, L.F. (1980)**: Implementation of implicit formulas for the solution of ODEs, SIAM J. Sci. Stat. Comp. 1, 103-118.

[13] **Wolfbrandt, A. (1977)**: A study of Rosenbrock processes with respect to order conditions and stiff stability, Ph. D. Thesis, Chalmers University of Technology, Göteborg.