PETER J. VAN DER HOUWEN
and
WOLTER A. VAN DER VEEN

CWI, Amsterdam

On Solving Implicit Differential Equations on Parallel Computers

Conferenza tenuta da P.J. van der Houwen il 5 giugno 1995

ABSTRACT. We construct and analyse integration methods for solving initial value problems for implicit differential equations (IDEs) that can be efficiently used on parallel computer systems. We construct an IDE method for general IDEs of arbitrarily high index, and two methods that can be applied to partitioned IDEs. The partitioned IDE methods both exploit the special form of the problem and converge faster than the general IDE method. The first partitioned IDE method is suitable for higher-index problems, the second partitioned IDE method only applies to index 1 problems, but is considerably less expensive on parallel computers. This paper presents the results presented in June 1995 at the Seminario Matematico e Fisico organized by the Mathematics Department of the Polytechnics University of Milano.

CR Subject Classification (1991): G.1.7.

Keywords and Phrases: numerical analysis, implicit ODEs, DAEs, Runge-Kutta methods, parallelism.

1. Introduction

Consider the initial value problem (IVP) for a system of implicit differential equations (IDEs) of the form

$$\phi(\dot{\mathbf{y}}(t), \mathbf{y}(t)) = \mathbf{0}, \quad \mathbf{y}(t_0) = \mathbf{y}_0; \quad \mathbf{y},\ \phi \in \mathbb{R}^d. \tag{1.1}$$

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

In particular, we shall consider the partitioned case where

$$K := \phi_{\mathbf{u}}(\mathbf{u},\mathbf{v}) = \begin{pmatrix} K_{11} & 0 \\ 0 & 0 \end{pmatrix}, \quad J := -\phi_{\mathbf{v}}(\mathbf{u},\mathbf{v}) = \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix}. \tag{1.2}$$

Here, $d = d_1 + d_2$ with K_{11} and J_{11} representing d_1-by-d_1 matrices (in the definition of J, the minus sign is inserted so that *explicit* differential equations $\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y})$ yield the familiar formula $J = \partial\mathbf{f}/\partial\mathbf{y}$).
If often happens that the problem (1.1) is given in the form

$$Q(\mathbf{y})\dot{\mathbf{y}}(t) = \mathbf{f}(\mathbf{y}(t)),\ \mathbf{y}(t_0) = \mathbf{y}_0, \quad \mathbf{y},\mathbf{f} \in \mathbb{R}^d \tag{1.3}$$

with Q a constant, singular matrix of rank d_1. By the transformation $\mathbf{z} = S_2^{-1}\mathbf{y}$, this problem can be transformed to

$$S_1 Q S_2 \dot{\mathbf{z}} - S_1\mathbf{f}(S_2\mathbf{z}) = \mathbf{0}, \quad \mathbf{z}(t_0) = S_2^{-1}\mathbf{y}_0, \quad \mathbf{z},\ \mathbf{f} \in \mathbb{R}^d \tag{1.3'}$$

with nonsingular matrices S_1 and S_2 such that

$$S_1 Q S_2 = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix},$$

where the dimension of I is the rank of Q (cf. [1, p. 406]). Problem (1.3') is of the partitioned form $\{(1.1),(1.2)\}$ with $K_{11} = I$. Introducing the partitioning $\mathbf{z} = (\mathbf{u}^T,\mathbf{v}^T)^T$, where $\mathbf{u}$ and $\mathbf{v}$ are respectively of dimension d_1 and d_2, it can be written in the familiar form of an IVP for the semiexplicit differential-algebraic equation (DAE)

$$\begin{aligned} &\frac{d\mathbf{u}}{dt} = \mathbf{f}(\mathbf{u},\mathbf{v}),\ \mathbf{u}(t_0) = \mathbf{u}_0, \\ &\mathbf{g}(\mathbf{u},\mathbf{v}) = \mathbf{0},\ \mathbf{v}(t_0) = \mathbf{v}_0, \end{aligned} \qquad \mathbf{u},\ \mathbf{f} \in \mathbb{R}^{d_1};\quad \mathbf{v},\ \mathbf{g} \in \mathbb{R}^{d_2}. \tag{1.4}$$

In this paper, we shall analyse integration methods for solving the IVP (1.1) that can be efficiently used on parallel computer systems. We construct an IDE method for general IDEs of arbitrarily high index, and two methods that can be applied to partitioned problems of the type (1.2). These partitioned IDE methods both exploit the special form of the problem and converge faster than the general IDE method. The first partitioned IDE method is suitable for higher-index problems, the second partitioned IDE method only applies to index 1 problems, but is considerably less expensive.

2. The numerical scheme

Let us start with the case where (1.1) can be presented in the explicit differential equation form

$$\dot{\mathbf{y}}(t) = \mathbf{f}(\mathbf{y}(\mathbf{t})),\ \mathbf{y}(t_0) = \mathbf{y}_0; \quad \mathbf{y},\ \mathbf{f} \in \mathbb{R}^d.$$

A large class of implicit step-by-step methods can be represented in the form

$$\mathbf{Y} - h(A \otimes I)\mathbf{F}(\mathbf{Y}) = \mathbf{W}, \quad \mathbf{y}_{n+1} = (\mathbf{e}_s^T \otimes I)\mathbf{Y}. \tag{2.1}$$

Here, A denotes a nonsingular s-by-s matrix, $\mathbf{W}$ is an sd-dimensional vector containing information computed in preceding integration steps, I is the d-by-d identity matrix, h is the stepsize $t_{n+1} - t_n$, and $\otimes$ denotes the Kronecker product. The s vector components $\mathbf{Y}_i$ of the sd-dimensional solution vector $\mathbf{Y}$ represent numerical approximations to the exact solution vectors $\mathbf{y}(\mathbf{e}t_n + \mathbf{c}h)$, $\mathbf{c}$ being the abscissa vector with $c_s = 1$ and $\mathbf{e}$ representing the s-dimensional vector with unit entries. Furthermore, $\mathbf{e}_s$ is the sth unit vector and $\mathbf{y}_n$ is the numerical approximation to $\mathbf{y}(t_n)$. In (2.1), $\mathbf{F}(\mathbf{V})$ contains the derivative values $(\mathbf{f}(\mathbf{V}_i))$ for any vector $\mathbf{V} = (\mathbf{V}_i)$. In the following, we shall use the notation I for an identity matrix. However, its dimension will always be clear from the context.

An important class of methods leading to implicit relations of the form (2.1) are the (stiffly accurate) Runge-Kutta (RK) methods, where $\mathbf{W} := \mathbf{e} \otimes \mathbf{y}_n$. In fact, in our numerical experiments we shall use such RK methods.

In order to derive the analogue of (2.1) for IDEs, we observe that if ϕ is invertible with respect to $\dot{\mathbf{y}}$, then this analogue should be equivalent to (2.1). This leads us to use (2.1) for expressing the derivative stage vector $\mathbf{F}(\mathbf{V})$ in terms of $\mathbf{Y}$, and to substitute this expression into the equation $\mathbf{\Phi}(\mathbf{F}(\mathbf{Y}), \mathbf{Y}) = \mathbf{0}$, with $\mathbf{\Phi}$ defined in the same way as $\mathbf{F}$. This yields the method

$$\mathbf{R}(\mathbf{Y}) = \mathbf{0},\ \mathbf{R}(\mathbf{Y}) := \mathbf{\Phi}((h^{-1}A^{-1} \otimes I)(\mathbf{Y} - \mathbf{W}), \mathbf{Y}),\ \mathbf{y}_{n+1} = (\mathbf{e}_s^T \otimes I)\mathbf{Y}. \tag{2.2}$$

Thus, the method (2.2) is completely specified by the pair $\{A, \mathbf{W}\}$. As an example, we consider the IVP (1.3). If we apply (2.2) with $\mathbf{W} := \mathbf{e} \otimes \mathbf{y}_n$, then it assumes the form

$$(h^{-1}A^{-1} \otimes Q)(\mathbf{Y} - \mathbf{W}) - \mathbf{F}(\mathbf{Y}) = 0, \quad \mathbf{y}_{n+1} = (\mathbf{e}_s^T \otimes I)\mathbf{Y}, \qquad (2.3)$$

which is equivalent to the RK method discussed in [1, p. 406].

REMARK 2.1 – As explained in [1, p. 407], the RK solution $\mathbf{Y}$ defined by (2.3) is algebraically identical to $(I \otimes S_2)\mathbf{Z}$, where $\mathbf{Z}$ is the RK solution obtained by applying (2.2) with $\mathbf{W} := \mathbf{e} \otimes \mathbf{z}_n$ to (1.3′), or equivalently, to (1.4). This equivalence holds for any method $\{A, \mathbf{W}\}$.

The implicit equation in (2.2) will be solved iteratively by generating sequences of iterates $\{\mathbf{Y}^{(j)}\}$. Our starting point is the iteration method

$$N(\mathbf{Y}^{(j)} - \mathbf{Y}^{(j-1)}) = -(hA \otimes I)\mathbf{R}(\mathbf{Y}^{(j-1)}), \quad j = 1, \ldots, m, \qquad (2.4)$$

where N is a nonsingular matrix. The iteration error associated with (2.4) satisfies the recursion

$$N(\mathbf{Y}^{(j)} - \mathbf{Y}) = N(\mathbf{Y}^{(j-1)} - \mathbf{Y}) - (hA \otimes I)(\mathbf{R}(\mathbf{Y}^{(j-1)}) - \mathbf{R}(\mathbf{Y})).$$

so that ignoring second-order terms leads to

$$\mathbf{Y}^{(j)} - \mathbf{Y} = M(\mathbf{Y}^{(j-1)} - \mathbf{Y}), \; M := N^{-1}(N - N_0), \; N_0 := I \otimes K - A \otimes hJ. \qquad (2.5)$$

where the Jacobian matrices K and J are both evaluated at the step point t_n. The conventional choice for N is the modified Newton iteration matrix N_0 resulting in a zero amplification matrix M. The advantage of the choice $N = N_0$ is that, even in strongly nonlinear problems, a few iterations usually suffice to solve the implicit system in (2.2). However, a disadvantage is that solving the linear Newton systems can be quite expensive. For example, when direct methods are used, the LU-decomposition of the sd-by-sd matrix N_0 requires as many as $O(s^3d^3)$ arithmetic operations.

In this paper, we shall consider several choices of more "convenient" iteration matrices N. Since a necessary and sufficient condition for

linear convergence of the iteration method (2.4) requires the spectral radius $\rho(M)$ to be less than 1, we shall try to combine a small spectral radius with a reduction of the complexity of the linear Newton systems. In particular, we shall look for matrices N that reduce the computational complexity on parallel computer systems (for example, matrices N with a block-triangular structure).

REMARK 2.2 – In an actual implementation of (2.4), it may be recommendable to remove the h^{-1} factor in the residual defined in (2.2) by defining the "derivative" iterate $\dot{\mathbf{Y}}^{(j)} := (h^{-1}A^{-1} \otimes I)(\mathbf{Y}^{(j)} - \mathbf{W})$. Then, the iteration scheme becomes

$$N(A \otimes I)(\dot{\mathbf{Y}}^{(j)} - \dot{\mathbf{Y}}^{(j-1)}) = -(A \otimes I)\Phi(\dot{\mathbf{Y}}^{(j-1)}, \mathbf{W} + \\ + (hA \otimes I)\dot{\mathbf{Y}}^{(j-1)}), \qquad j = 1, \ldots, m,$$
$$\mathbf{Y}^{(j)} = \mathbf{W} + h(A \otimes I)\dot{\mathbf{Y}}^{(j)}. \tag{2.4'}$$

The sequences $\{\mathbf{Y}^{(j)}\}$ generated by the schemes (2.4) and (2.4') are algebraicly identical, but (2.4') can be used as $h \to 0$.

2.1. General IDE method

Consider the iteration matrix

$$N = I \otimes K - B \otimes hJ = \begin{pmatrix} K - hT_{11}J & 0 & 0 & \ldots \\ -hT_{21}J & K - hT_{22}J & 0 & \ldots \\ -hT_{31}J & -hT_{32}J & K - hT_{33}J & \ldots \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}, \tag{2.6}$$

where $B = (B_{ij})$ is a diagonal matrix D or a lower triangular matrix T. On substitution into (2.4) we obtain

$$(I \otimes K - B \otimes hJ)(\mathbf{Y}^{(j)} - \mathbf{Y}^{(j-1)}) = -(hA \otimes I)\mathbf{R}(\mathbf{Y}^{(j-1)}). \tag{2.7}$$

Formally, this method can be applied to problems of any index as long as the matrix $I \otimes K - B \otimes hJ$ is nonsingular (that is, the blocks $K - hB_{ii}J$ should be nonsingular). However, in order to have convergence to the exact solution as $h \to 0$, the matrix K should be nonsingular. The

method (2.7) will be referred to as the *general IDE method*. Furthermore, we shall say that the method is in *diagonal mode* if $B = D$ and in *triangular mode* if $B = T$.

Each iteration with (2.7) requires the solution of a linear system with the block-triangular Newton matrix N. Hence, the system splits into s subsystems of dimension d, reducing the computational costs considerably. For example, if direct solution methods are used, then the s LU-decompositions associated with the s systems require $O(sd^3)$ flops which is factor s^2 less than the number of flops needed when the modified Newton matrix N_0 is used. Moreover, these LU-decompositions can be done in parallel, so that the effective costs on a parallel system are a factor s^3 smaller. Similar types of linear systems occur in the parallel diagonal-implicitly iterated RK methods and the parallel triangular-implicitly iterated RK methods analysed in [2] and [3] for solving IVPs for ODEs.

It follows from (2.5) that the error amplification matrix corresponding to (2.7) is given by

$$M = (I \otimes K - B \otimes hJ)^{-1}((A - B) \otimes hJ). \tag{2.8}$$

By means of this matrix, we can derive convergence results for nonsingular K and for nonsingular J, respectively.

2.1.1. Convergence results

For nonsingular K, the amplification matrix (2.8) can be written as

$$M = (I - B \otimes hK^{-1}J)^{-1}((A - B) \otimes hK^{-1}J), \tag{2.8'}$$

so that the eigenvalues of M are given by those of the matrix

$$Z(z) := z(I - zB)^{-1}(A - B), \tag{2.9}$$

where $z \in \sigma(hK^{-1}J)$. Similarly, if J is nonsingular, then we write the amplification matrix (2.8) in the form

$$M = (I \otimes h^{-1}J^{-1}K - B \otimes I)^{-1}((A - B) \otimes I). \tag{2.8''}$$

The eigenvalues of M are again those of $Z(z)$, but now with $z^{-1} \in \sigma(h^{-1}J^{-1}K)$.

Matrices of the type (2.9) have extensively been studied in [2] and [3]. For a large number of RK matrices A, diagonal and lower triangular matrices B have been found such that $\rho(Z(z)) < 1$ if $\mathrm{Re}(z) \leq 0$. In the following, matrices B that possess this property will be said to lie in the set $\mathbb{B}(A)$ associated with the RK matrix A.

We summarize the result established in this subsection in the following theorem (from now on, we shall tacitly assume the nonsingularity of matrices as soon as their inverses are used):

THEOREM 2.1 – *If $B \in \mathbb{B}(A)$, then the general IDE method (2.7) converges for all h, provided that either* $\mathrm{Re}(\sigma(K^{-1}J)) \leq 0$ *or* $\mathrm{Re}(\sigma(J^{-1}K)) \leq 0$.

EXAMPLE 2.1 – Consider the implicit IVP (1.3) where $K = Q(\mathbf{y})$ and $J = (\mathbf{f}(\mathbf{y}) - Q(\mathbf{y})\dot{\mathbf{y}})_{\mathbf{y}}$. If K is nonsingular in the neighbourhood of the solution, then Theorem 2.1 requires that the matrix $K^{-1}J = Q^{-1}(\mathbf{y})(\mathbf{f}(\mathbf{y}) - Q(\mathbf{y})\dot{\mathbf{y}})_{\mathbf{y}}$ has its eigenvalues in the nonpositive halfplane. If, instead, J is nonsingular, then Theorem 2.1 requires the matrix $J^{-1}K = ((\mathbf{f}(\mathbf{y}) - Q(\mathbf{y})\dot{\mathbf{y}})_{\mathbf{y}})^{-1}Q(\mathbf{y})$ to have its eigenvalues in the nonpositive halfplane. As a special case of (1.3), we have the DAE of index 2, defined by

$$\begin{pmatrix} I & O \\ O & O \end{pmatrix} \begin{pmatrix} \dot{\mathbf{u}} \\ \dot{\mathbf{v}} \end{pmatrix} - \begin{pmatrix} \mathbf{g}(\mathbf{u},\mathbf{v}) \\ \mathbf{h}(\mathbf{u}) \end{pmatrix} = \mathbf{0},$$

with $\mathbf{g}_{\mathbf{v}}$ and $\mathbf{h}_{\mathbf{u}}$ nonsingular. The matrices J and $J^{-1}K$ are given by

$$J = \begin{pmatrix} \mathbf{gu} & \mathbf{gv} \\ \mathbf{hu} & O \end{pmatrix}, \quad J^{-1}K = \begin{pmatrix} O & O \\ \mathbf{g}_{\mathbf{v}}^{-1} & O \end{pmatrix},$$

so that $J^{-1}K$ has zero eigenvalues implying convergence of the iteration method.

2.2. IDE methods for partitioned problems

In this section, we consider problems of the partitioned form {(1.1), (1.2)}. For such problems, it is convenient to write $\mathbf{y} = (\mathbf{u}^T, \mathbf{v}^T)^T$, where $\mathbf{u}$ and $\mathbf{v}$ are respectively of dimension d_1 and d_2, and to replace the stage vector $\mathbf{Y}$ by the permuted stage vector $\mathbf{X} = P\mathbf{Y} := (\mathbf{U}^T, \mathbf{V}^T)^T$, where $\mathbf{U}$ and $\mathbf{V}$ are stage vectors associated with $\mathbf{u}$ and $\mathbf{v}$ in the same way as $\mathbf{Y}$ is associated with $\mathbf{y}$. Let us introduce the permuted iterates $\mathbf{X}^{(j)} := P\mathbf{Y}^{(j)}$, then the permuted versions of (2.4) and (2.5) become

$$\tilde{N}(\mathbf{X}^{(j)} - \mathbf{X}^{(j-1)}) = -P(hA \otimes I)\mathbf{R}(P^{-1}\mathbf{X}^{(j-1)}), \quad \tilde{N} := PNP^{-1}, \tag{2.10}$$

$$\mathbf{X}^{(j)} - \mathbf{X} = \tilde{M}(\mathbf{X}^{(j-1)} - \mathbf{X}), \quad \tilde{M} := PMP^{-1} = \tilde{N}^{-1}(\tilde{N} - \tilde{N}_0),$$

$$\tilde{N}_0 = P(I \otimes K - A \otimes hJ)P^{-1}. \tag{2.11}$$

Before selecting suitable matrices $\tilde{N}$, we consider the general IDE method, when applied to problems characterized by (1.2). By replacing in (2.7) $\mathbf{Y}^{(j)}$ by $P^{-1}\mathbf{X}^{(j)}$ and by observing that the permutation matrix P has the property that for any matrix C and any matrix J with a partitioning as in (1.2), we have

$$P(C \otimes J)P^{-1} = \begin{pmatrix} C \otimes J_{11} & C \otimes J_{12} \\ C \otimes J_{21} & C \otimes J_{22} \end{pmatrix}.$$

it can be verified that the general IDE method (2.7) takes the form

$$\begin{pmatrix} I \otimes K_{11} - B \otimes hJ_{11} & -B \otimes hJ_{12} \\ -B \otimes hJ_{21} & -B \otimes hJ_{22} \end{pmatrix} (\mathbf{X}^{(j)} - \mathbf{X}^{(j-1)}) =$$

$$= \begin{pmatrix} -hA \otimes I & O \\ O & -hA \otimes I \end{pmatrix} P\mathbf{R}(P^{-1}\mathbf{X}^{(j-1)}). \tag{2.7'}$$

Since

$$\tilde{N}_0 = \begin{pmatrix} I \otimes K_{11} - A \otimes hJ_{11} & -A \otimes hJ_{12} \\ -A \otimes hJ_{21} & -A \otimes hJ_{22} \end{pmatrix}.$$

the matrix $\tilde{N}$ associated with (2.7′) can be written as

$$\tilde{N} = \begin{pmatrix} I \otimes K_{11} - B \otimes hJ_{11} & -B \otimes hJ_{12} \\ -B \otimes hJ_{21} & -B \otimes hJ_{22} \end{pmatrix} = \\ = \tilde{N}_0 + h \begin{pmatrix} (A-B) \otimes J_{11} & (A-B) \otimes J_{12} \\ (A-B) \otimes J_{21} & (A-B) \otimes J_{22} \end{pmatrix}. \tag{2.6′}$$

From (2.11) it follows that the convergence is expected to be faster as the magnitude of the matrix $\tilde{N} - \tilde{N}_0$ is smaller. Hence, taking into account that we should not loose computational efficiency, we are led to the alternative iteration matrices:

$$\tilde{N} = \begin{pmatrix} I \otimes K_{11} - B \otimes hJ_{11} & -B \otimes hJ_{12} \\ -A \otimes hJ_{21} & -A \otimes hJ_{22} \end{pmatrix} = \\ \tilde{N}_0 + h \begin{pmatrix} (A-B) \otimes J_{11} & (A-B) \otimes J_{12} \\ O & O \end{pmatrix} \tag{2.12}$$

$$\tilde{N} = \begin{pmatrix} I \otimes K_{11} - B \otimes hJ_{11} & -A \otimes hJ_{12} \\ O & -A \otimes hJ_{22} \end{pmatrix} = \tilde{N}_0 + h \begin{pmatrix} (A-B) \otimes J_{11} & O \\ A \otimes J_{21} & O \end{pmatrix}. \tag{2.13}$$

where in (2.13) J_{22} is assumed nonsingular. These iteration matrices generate the methods

$$\begin{pmatrix} I \otimes K_{11} - B \otimes hJ_{11} & -B \otimes hJ_{12} \\ I \otimes J_{21} & I \otimes J_{22} \end{pmatrix} (\mathbf{X}^{(j)} - \mathbf{X}^{(j-1)}) = \\ = \begin{pmatrix} -hA \otimes I & O \\ O & I \end{pmatrix} P\mathbf{R}(P^{-1}\mathbf{X}^{(j-1)}), \tag{2.14}$$

$$\begin{pmatrix} I \otimes K_{11} - B \otimes hJ_{11} & -A \otimes hJ_{12} \\ O & I \otimes J_{22} \end{pmatrix} (\mathbf{X}^{(j)} - \mathbf{X}^{(j-1)}) = \\ = \begin{pmatrix} -hA \otimes I & O \\ O & I \end{pmatrix} P\mathbf{R}(P^{-1}\mathbf{X}^{(j-1)}), \tag{2.15}$$

and will be referred to as the *partitioned IDE method* and the *partitioned index 1 IDE method* (note that the nonsingularity of J_{22} in (2.15) requires the problem to be of index 1).

From (2.12) and (2.13), we conclude that for many problems the two partitioned IDE methods converge faster than the general IDE method (2.7′) because they are "closer" to the true Newton process. However, if B is such that the matrix $A - B$ has entries of small magnitude, then we may expect that the general IDE method and the partitioned IDE method (2.12) are both superior to the partitioned index 1 IDE method (2.15).

Both partitioned methods lead to linear systems, of which the matrices of coefficients can be transformed to block-triangular form. However, they differ by their degree of implicitness, and therefore by their amount of intrinsic parallelism. Note that the partitioned IDE methods (2.14) and (2.15) both reduce to the general IDE method (2.7′) if $d_2 = 0$ (that is, if there is no partitioning).

2.2.1. Partitioned IDE method

Each iteration with (2.14) requires the solution of s systems of dimension $d := d_1 + d_2$ whose matrices of coefficients are of the form

$$C_i := \begin{pmatrix} K_{11} - hB_{ii}J_{11} & hB_{ii}J_{12} \\ J_{21} & J_{22} \end{pmatrix} \qquad i = 1, \ldots, s.$$

Thus, in order to apply the partitioned IDE method, the matrices C_i should be nonsingular.

Furthermore, it follows from (2.7′) and (2.14) that for partitioned problems, the computational complexities of the general IDE and partitioned IDE methods are comparable.

In order to derive convergence conditions, we rewrite its amplification matrix $\tilde{M}$ in one of the following two forms:

$$M = \begin{pmatrix} \tilde{M}_{11} & O \\ \tilde{M}_{21} & O \end{pmatrix}, \quad \tilde{M} = \begin{pmatrix} O & \tilde{M}_{12} \\ O & \tilde{M}_{22} \end{pmatrix},$$

where the submatrices $\tilde{M}_{11}$ and $\tilde{M}_{22}$ are defined by

$$\tilde{M}_{11} := (I - B \otimes hK_{11}^{-1}S_{22})^{-1}((A - B) \otimes hK_{11}^{-1}S_{22}),$$

$$S_{22} := J_{11} - J_{12}J_{22}^{-1}J_{21},$$

$$\tilde{M}_{22} := (B \otimes I - I \otimes h^{-1}S_{21}^{-1}K_{11}J_{21}^{-1}J_{22})^{-1}((A - B) \otimes I),$$

$$S_{21} := J_{12} - J_{11}J_{21}^{-1}J_{22}.$$

Obviously, the eigenvalues of $\backslash O(M, \tilde{\ })$, and therefore the eigenvalues of M, are given by those of $\tilde{M}_{11}$ or $\tilde{M}_{22}$. A comparison of $\tilde{M}_{11}$ with the amplification matrix obtained in (2.8) reveals that they have the same structure. Hence, the same matrices T as before can be used to make $\rho(M) = \rho(\tilde{M}_{11}) < 1$ for all h, provided that the matrix $K_{11}^{-1}S_{22}$ has its eigenvalues in the nonpositive halfplane. The eigenvalues of $\tilde{M}_{22}$ are given by the eigenvalues of $Z(z)$, defined in (2.9), with $z^{-1} \in \sigma(h^{-1}S_{21}^{-1}K_{11}J_{21}^{-1}J_{22})$.

THEOREM 2.2 – *Let $B \in \mathbb{B}(A)$ and let (1.2) hold. Then, the partioned IDE method (2.14) converges for all h, provided that either* $\mathrm{Re}(\sigma(K_{11}^{-1}S_{22})) \leq 0$ *or* $\mathrm{Re}(\sigma(S_{21}^{-1}K_{11}J_{21}^{-1}J_{22})) \leq 0$.

We remark that the Theorems 2.1 and 2.2 lead to equivalent convergence conditions when applied to partitioned problems satisfying (1.2). For example, Theorem 2.1 claims convergence if all eigenvalues λ of the matrix $J^{-1}K$ are in the nonpositive halfplane. It is easily verified that these eigenvalues λ also satisfy the relation $K_{11}\mathbf{u} = \lambda S_{22}\mathbf{u}$. Hence, λ^{-1} is an eigenvalue of $K_{11}^{-1}S_{22}$ which Theorem 2.2 requires to be in the nonnegative halfplane. Thus, the two convergence conditions are equivalent when both are applicable.

EXAMPLE 2.2 – The IDE in (1.3) can be written as the DAE (cf. [1, p. 486])

$$\begin{pmatrix} I & O \\ O & O \end{pmatrix} \begin{pmatrix} \dot{\mathbf{u}} \\ \dot{\mathbf{v}} \end{pmatrix} - \begin{pmatrix} \mathbf{v} \\ Q(\mathbf{u})\mathbf{v} - \mathbf{f}(\mathbf{u}) \end{pmatrix} = \mathbf{0},$$

which satisfies (1.2) with $K_{11} = I$ and $J_{11} = O$, $J_{12} = I$, $J_{22} = Q$. Assuming that Q is nonsingular, we have $S_{22} = Q^{-1}(\mathbf{u})(\mathbf{f}(\mathbf{u}) - Q(\mathbf{u})\mathbf{v})_{\mathbf{u}}$. Theorem 2.2 requires that S_{22} has its eigenvalues in the nonpositive halfplane, the same condition as derived in Example 2.1.

2.2.2. Partitioned index 1 IDE method

From (2.15) it follows that in each iteration we can first solve in parallel the s (uncoupled) d_2-dimensional systems for the last sd_2 components of $\mathbf{X}^{(j)} - \mathbf{X}^{(j-1)}$ and next the s system of dimension d_1 for the first sd_1 components. The latter s systems can also be solved concurrently if B is diagonal. In the case of direct solvers, the LU-decompositions of J_{22} and the matrices $I - B_{ii}J_{11}$, $i = 1, \ldots, s$, can again be done in parallel. It is here where the sequential (or effective) costs of the partitioned index 1 IDE method (2.15) may be substantially less than those for the general IDE and partitioned IDE methods (2.7) and (2.14), whose LU-costs are $\approx 2(d_1+d_2)^3/3$ flops, whereas the partitioned index 1 IDE method requires only $2(\max\{d_1, d_2\})^3/3$ flops, i.e. a speed-up factor $(d_1 + d_2)^3 \ (\max\{d_1, d_2\})^{-3} = (1 + \min\{d_1 d_2^{-1}, d_2 d_1^{-1}\})^3$. For $d_1 = d_2$, a maximum speed-up factor of 8 is obtained.

Proceeding as in the preceding subsection, we obtain for the amplification matrix $\tilde{M}$

$$\tilde{M} = \begin{pmatrix} \tilde{M}_{11} & O \\ -I \otimes hJ_{22}^{-1}J_{21} & O \end{pmatrix},$$

where

$$\tilde{M}_{11} := (I - B \otimes hK_{11}^{-1}J_{11}^{-1})\,(A \otimes hK_{11}^{-1}S_{22} - B \otimes hK_{11}^{-1}J_{11}),$$

so that M and $\tilde{M}_{11}$ have the same eigenvalues. If either J_{11} vanishes or if we choose $B = O$, then

$$\rho(M) = \rho(\tilde{M}_{11}) = \rho(A \otimes hK_{11}^{-1}S_{22}) = h\,\rho(A)\,\rho(K_{11}^{-1}S_{22}),$$

and if either J_{12} or J_{21} vanishes, then

$$\rho(M) = \rho(\tilde{M}_{11}) = \rho(I - B \otimes hK_{11}^{-1}J_{11})^{-1}\,((A - B) \otimes hK_{11}^{-1}J_{11})).$$

This leads to the theorem:

THEOREM 2.3 – *Let (1.2) be satisfied. Then, the partitioned index 1 IDE method (2.15) converges if*

$$\textbf{(a)} \quad B = O \text{ or } J_{11} = O, \qquad h < \frac{1}{\rho(A)\rho(K_{11}^{-1}S_{22})}, \qquad S_{22} := J_{11} - J_{12}J_{22}^{-1}J_{21},$$

$$\textbf{(b)} \quad J_{12} = O \text{ or } J_{21} = O, \quad B \in \mathbb{B}(A), \ \mathrm{Re}\,(\sigma(K_{11}^{-1}J_{11})) \leq 0.$$

EXAMPLE 2.3 – Consider the DAE of Example 2.2. Since $J_{11} = O$, Theorem 2.3 (a) applies with $K_{11} = I$ and $S_{22} = Q^{-1}(\mathbf{u})(\mathbf{f}(\mathbf{u}) - Q(\mathbf{u})\mathbf{v})_{\mathbf{u}}$, so that we have convergence if $h < (\rho(A)\,\rho(S_{22}))^{-1}$.

EXAMPLE 2.4 – Consider the DAE of index 1 defined by

$$\begin{pmatrix} I & O \\ O & O \end{pmatrix} \begin{pmatrix} \mathbf{u}' \\ \mathbf{v}' \end{pmatrix} - \begin{pmatrix} \mathbf{g}(\mathbf{u},\mathbf{v}) \\ \mathbf{h}(\mathbf{v}) \end{pmatrix} = \mathbf{0},$$

with $J_{22} = \mathbf{h}_{\mathbf{v}}$ nonsingular. Since $J_{21} = \mathbf{h}_{\mathbf{u}} = O$, Theorem 2.3 (b) applies with $K_{11} = I$, so that we have convergence if $B \in \mathbb{B}(A)$ and $\mathrm{Re}\,(\sigma(J_{11})) \leq 0$.

We did not succeed in deriving sharp estimates for more general situations than the special cases covered by Theorem 2.3. Therefore, we followed an alternative approach, by deriving the spectral radius of the amplification matrix for a related iteration scheme. The related scheme generates a sequence $\{\mathbf{X}^{(k,m)}\}$, $k = 1,2,\ldots$, and is defined by

$$\mathbf{X}^{(k,0)} = \mathbf{X}^{(k-1,m)}, \tilde{N}(\mathbf{X}^{(kj)} - \mathbf{X}^{(k,j-1)}) =$$

$$= -P\mathbf{R}^*(P^{-1}\mathbf{X}^{(k,j-1)}, P^{-1}\mathbf{X}^{(k-1,m)}, \quad j = 1,\ldots,m,$$

$$\mathbf{R}^*(\mathbf{U},\mathbf{V}) := (I \otimes K - A \otimes hJ^*)\,(\mathbf{U} - \mathbf{V}) + (hA \otimes I)\mathbf{R}(\mathbf{V}), \quad J^* := \begin{pmatrix} J_{11} & J_{12} \\ O & J_{22} \end{pmatrix}. \tag{2.16}$$

where $\tilde{N}$ is defined as in (2.13). Let us compare this related scheme with the iteration scheme (2.15) we actually apply. Writing (2.15) in

the form

$$\tilde{N}(\mathbf{X}^{(k)} - \mathbf{X}^{(k-1)}) = -P(hA \otimes I)\mathbf{R}(P^{-1}\mathbf{X}^{(k-1)}) = \\ = -P\mathbf{R}^*(P^{-1}\mathbf{X}^{(k-1)}, P^{-1}\mathbf{X}^{(k-1)}) \tag{2.15'}$$

with $\tilde{N}$ also defined by (2.13), we see that (2.15') and (2.16) only differ in the frequency by which the quantity $\mathbf{R}^*$ is updated. In (2.15'), $\mathbf{R}^*$ is updated in each iteration, whereas (2.16) updates $\mathbf{R}^*$ each m iterations. This leads us to conclude that the partitioned index 1 IDE method (2.15) will converge as $j \to \infty$, whenever (2.16) converges as $j = 0, 1, \ldots, m$ and $k \to \infty$.

In order to derive a convergence condition for (2.16), let k be fixed and let $\mathbf{X}^{(k)}$ be the solution of the equation

$$\mathbf{R}^*(P^{-1}\mathbf{X}^{(k)}, P^{-1}\mathbf{X}^{(k-1,m)}) = \mathbf{0}\,. \tag{2.17}$$

It is easily verified that

$$\mathbf{X}^{(k,j)} - \mathbf{X}^{(k)} = \tilde{M}^*(\mathbf{X}^{(k,j-1)} - \mathbf{X}^{(k)}), \qquad \tilde{M}^* = \begin{pmatrix} \tilde{M}_{11}^* & O \\ O & O \end{pmatrix},$$

where

$$\tilde{M}_{11}^* := (I - B \otimes hK_{11}^{-1}J_{11})^{-1}\,((A - B) \otimes hK_{11}^{-1}J_{11})\,. \tag{2.18}$$

As before, this leads us to require that the matrix $K_{11}^{-1}J_{11}$ has its eigenvalues in the nonpositive halfplane. Since this condition implies convergence of (2.16) for any fixed k, it follows from (2.17) that for sufficiently large m the iterates $\mathbf{X}^{(k)}$ satisfy the relation

$$\mathbf{R}^*(P^{-1}\mathbf{X}^{(k)}, P^{-1}\mathbf{X}^{(k-1)}) = \mathbf{0}\,. \tag{2.19}$$

Next, we consider the convergence of the iterates $P^{-1}\mathbf{X}^{(k)}$ to the solution $\mathbf{Y}$ in (2.2), On subtraction of (2.2) and (2.19), and ignoring second-order terms, it follows that

$$P(I \otimes K - A \otimes hJ^*)P^{-1}(\mathbf{X}^{(k)} - P\mathbf{Y}) = \\ P(A \otimes (hJ - hJ^*))P^{-1}(\mathbf{X}^{(k-1)} - P\mathbf{Y})\,.$$

An elementary calculation yields

$$\mathbf{X}^{(k)} - P\mathbf{Y} = \tilde{M}(\mathbf{X}^{(k-1)} - P\mathbf{Y}), \quad \tilde{M} = \begin{pmatrix} \tilde{M}_{11} & O \\ -I \otimes hJ_{22}^{-1}J_{21} & O \end{pmatrix}.$$

where

$$\tilde{M}_{11} := (I - A \otimes hK_{11}^{-1}J_{11})^{-1} \ (A \otimes hK_{11}^{-1}J_{12}J_{22}^{-1}J_{21}). \tag{2.20}$$

Thus, the eigenvalues of the amplification matrix M are given by the eigenvalues of

$$Z^*(\alpha h) := \alpha h(I - \alpha hK_{11}^{-1}J_{11})^{-1}K_{11}^{-1}(J_{11} - S_{22}),$$

$$S_{22} := J_{11} - J_{12}J_{22}^{-1}J_{21}, \ \alpha \in \sigma(A), \tag{2.21}$$

so that the convergence condition for the sequence $\{\mathbf{X}^{(k)}\}$ becomes $\rho(Z^*(\alpha h)) < 1$.
Summarizing, we have the result

THEOREM 2.4 – *Let $B \in \mathbb{B}(A)$, let (1.2) be satisfied, let $Z^*(\alpha h)$ be defined by (2.21), and let $\mathrm{Re}(\sigma(K_{11}^{-1}J_{11})) \leq 0$. Then, the related method (2.16) converges, provided that $\rho(Z^*(\alpha h)) < 1$ for all $\alpha \in \sigma(A)$.*

We recall that it is expected that this convergence theorem also applies to the partitioned index 1 IDE method (2.15).

EXAMPLE 2.6 – Consider the DAE of index 1 of Example 2.4. Since $K_{11} = I$ and $J_{21} = \mathbf{h_u} = O$, it follows that $Z^*(\alpha h) = O$, so that Theorem 2.4 implies that (2.16) converges if $\mathrm{Re}(\sigma(J_{11})) \leq 0$. This condition is identical with the one obtained in Example 2.4 for the method partitioned index 1 IDE method (2.15).

The usefulness of Theorem 2.4 depends on the possibility of estimating the quantity $\rho(Z^*(\alpha h))$. By using properties of the logarithmic norm,

it is sometimes possible to derive sufficiently sharp estimates. Let $H := I - \alpha h K_{11}^{-1} J_{11}$. Then,

$$\rho(Z^*(\alpha h)) \leq h|\alpha| \|H^{-1}\| \, \|K_{11}^{-1}(J_{11} - S_{22})\| \leq \frac{h|\alpha| \|K_{11}^{-1}(J_{11} - S_{22})\|}{\max\{-\mu[-H], -\mu[H]\}},$$

where $\mu[H]$ denotes the logarithmic norm of H.

3. Numerical examples

We solved these systems using the general IDE method (2.7), and the partitioned IDE methods (2.14) and (2.15) when applicable. In all cases, the corrector is defined by the four-stage Radau IIA corrector and the predictor formula is given by $\mathbf{Y}^{(0)} := E\mathbf{Y}_n^{(m)}$, where $\mathbf{Y}_n^{(m)}$ denotes the final approximation to the stage vector obtained in the preceding step, and E is the extrapolation matrix of maximal order. The matrix B is chosen either according to the diagonal matrix D derived in [2] or the lower triangular matrix T derived in [3]. These matrices are given by

$$B = D := \begin{pmatrix} 0.3205 & 0 & 0 & 0 \\ 0 & 0.0892 & 0 & 0 \\ 0 & 0 & 0.1817 & 0 \\ 0 & 0 & 0 & 0.2334 \end{pmatrix},$$

$$B = T := \begin{pmatrix} 0.1130 & 0 & 0 & 0 \\ 0.2344 & 0.2905 & 0 & 0 \\ 0.2167 & 0.4834 & 0.3083 & 0 \\ 0.2205 & 0.4668 & 0.4414 & 0.1176 \end{pmatrix}.$$

Both matrices have the property that $\rho(Z(z)) < 1$ whenever $\mathrm{Re}\,(z) \leq 0$. However, for $B = D$ the maximal value of $\|Z(z)^j\|$ in the lefthand halfplane is greater than 1 for $j \leq 4$, whereas for $B = L$, it is less than 1 for all j. As a consequence, $B = T$ should lead to a much more robust iteration scheme. Furthermore, the matrices $A - B$ that play a role in how close the iteration methods are to the true Newton iteration

process are given by

$$A - D = \begin{pmatrix} -0.2075 & -0.0403 & 0.0258 & -0.0099 \\ 0.2344 & 0.1177 & -0.0479 & 0.0160 \\ 0.2167 & 0.4061 & 0.0073 & -0.0242 \\ 0.2205 & 0.3882 & 0.3288 & -0.1709 \end{pmatrix},$$

$$A - T = - \begin{pmatrix} 0 & 0.0403 & -0.0258 & 0.0099 \\ 0 & 0.0836 & 0.0479 & -0.0160 \\ 0 & 0.0773 & 0.1192 & 0.0242 \\ 0 & 0.0786 & 0.1126 & 0.0551 \end{pmatrix}.$$

Evidently, the magnitude of $A - T$ is considerably smaller than that of $A - D$.

Since this paper aims at a comparison of algorithmic properties of the three IDE methods and the effect of the diagonal and triangular modes, we avoided effects of stepsize and iteration strategies by performing the experiments with fixed stepsizes h and fixed numbers of iterations m. The tables of results below list the minimal number of correct significant digits defined by.

$$csd := -\log_{10} \left\| \frac{\mathbf{y}_N - \mathbf{y}(t_N)}{\mathbf{y}(t_N)} \right\|_\infty,$$

where $\mathbf{y}_N$ denotes the numerical solution at the end point t_N, and where the division of vectors should be understood as componentwise division.

3.1. The Colpitts oscillator

Our first test problem is the IVP for the Colpitts oscillator specified in [4]. This IVP of index 0 is described by an implicit ODE system of the form (1.3) with four linear differential equations and with constant, nonsingular capacity matrix Q. For such problems, the general

IDE method and the two partitioned IDE methods are identical. Therefore, Table 3.1 only compares the effect of the diagonal and triangular modes. The results show that the diagonal mode performs much better. Apparently, this problem is relatively easy, so that the increased robustness of the triangular mode is not demonstrated.

Table 3.1. General IDE method applied to the Colpitts oscillator in the form (1.3).

3.1a. $B = D$.

h	$m=3$	$m=4$	$m=5$	$m=6$
0.04	4.0	5.7	6.1	7.4
0.02	4.8	7.8	7.9	9.4

3.1b. $B = T$.

h	$m=3$	$m=4$	$m=5$	$m=6$
0.04	3.0	4.2	5.6	6.4
0.02	4.2	5.8	7.3	8.7

3.2. The transistor amplifier

The second test problem is an IVP for the transistor amplifier given in [1] (see also [5]). This nonlinear, eight-dimensional problem of index 1 can be represented in the implicit form (1.3) with a constant, nonpartitioned (but singular) capacity matrix Q, as well as in semi-explicit form (1.4) with $d_1 = 5$ and $d_2 = 3$. In the implicit form (1.3), only the general IDE method can be applied, whereas the semi-explicit form (1.4) allows application of all three IDE methods. In order to facilitate a mutual comparison, the *csd*-values in the tables of results refer to the accuracies of the numerical solution of the untransformed problem (1.3).

Table 3.2. General IDE method applied to the Transistor amplifier in the form (1.3).

3.2a. $B = D$.

h	$m=4$	$m=5$	$m=6$	$m=7$
$4 \cdot 10^{-4}$	*	*	*	5.7
$2 \cdot 10^{-4}$	6.0	7.9	8.0	8.6

3.2b. $B = T$.

h	$m=4$	$m=5$	$m=6$	$m=7$
$4 \cdot 10^{-4}$	*	6.8	6.5	6.5
$2 \cdot 10^{-4}$	8.0	8.6	8.8	9.3

Table 3.3. General IDE qmethod applied to the Transistor amplifier in the semi explicit form (1.4).

3.3a. $B = D$.

h	$m=4$	$m=5$	$m=6$	$m=7$
$4 \cdot 10^{-4}$	*	*	*	5.1
$2 \cdot 10^{-4}$	4.6	6.9	7.1	7.3

3.3b. $B = T$.

h	$m=4$	$m=5$	$m=6$	$m=7$
$4 \cdot 10^{-4}$	*	5.5	6.0	6.5
$2 \cdot 10^{-4}$	6.6	7.8	7.8	8.4

Table 3.4. Partitioned IDE method applied to the Transistor amplifier in the form (1.4).

3.4a. $B = D$.

h	$m=4$	$m=5$	$m=6$	$m=7$
$4 \cdot 10^{-4}$	3.3	3.6	3.9	4.2
$2 \cdot 10^{-4}$	5.1	5.7	6.3	7.0

3.4b. $B = T$.

h	$m=4$	$m=5$	$m=6$	$m=7$
$4 \cdot 10^{-4}$	3.8	4.0	4.3	4.7
$2 \cdot 10^{-4}$	5.4	6.2	6.9	7.6

Table 3.5. Partitioned index 1 IDE method applied to the Transistor amplifier in the form (1.4).

3.5a. $B = D$.

h	$m=4$	$m=5$	$m=6$	$m=7$
$4 \cdot 10^{-4}$	4.1	5.1	4.8	5.6
$2 \cdot 10^{-4}$	5.5	6.3	8.1	7.3

3.5b. $B = T$.

h	$m=4$	$m=5$	$m=6$	$m=7$
$4 \cdot 10^{-4}$	4.1	4.5	5.0	5.6
$2 \cdot 10^{-4}$	5.9	6.4	7.0	7.7

Table 3.2 lists results for he general IDE method when applied to the implicit form (1.3). It clearly shows the greater robustness of the triangular mode. Apparently, the Transistor amplifier presents a much more difficult problem than the Colpitts oscillator, so that the diagonal mode fails for a small numbers of iterations.

The tables 3.3, 3.4 and 3.5 present results for the three IDE methods when applied to the semiexplicit form (1.4). As expected, the triangular mode is again superior to the diagonal mode. Furthermore, we see that for larger stepsizes, the partitioned IDE methods are more robust than the general IDE method.

Notice that algebraically, the results in the Tables 3.2 and 3.3 should be identical. They only differ by the way they are computed. The differences in these tables are due to rounding errors caused by the transformation from the variable $\mathbf{y}$ to $\mathbf{z} = S_2^{-1}\mathbf{y}$ and back (cf. Remark 2.1).

References

[1] HAIRER, E. and WANNER, G., Solving ordinary differential equations, II. Stiff and Differential-algebraic problems, Springer-Verlag, Berlin, 1991.

[2] VAN DER HOUWEN, P.J. and SOMMEIJER, B.P., *Iterated Runge-Kutta methods on parallel computers*, SIAM J. Sci. Stat. Comput., **12** (1991), 1000–1028.

[3] VAN DER HOUWEN, P.J. and DE SWART, J.J.B., *Triangularly implicit iteration methods for ODE-IVP solvers*, submitted for pubblication, 1995.

[4] KAMPOWSKY, W., RENTROP, P. and SCHMIDT, W., *Classification and numerical simulation of electric circuits*, Surveys on Mathematical for Industry, **2** (1992), 23–65.

[5] DE SWART, J.J.B., VAN DER VEEN, W.A. and SOMMEIJER, B.P., Test set for IVPODE and IVPDAE solvers, available via anonymous ftp at the site *ftp.cwi.nl* in the directory *publ/IVPtestset* in the (compressed) file *testset.ps.Z.*,

Peter J. van der Houwen, Wolter A. van der Veen
CWI
P.O. Box 94079
1090 GB Amsterdam
The Nederlands

Pervenuta in Redazione il 28 luglio 1995