Baltzer Journals June 30, 1998

Non-stiff integrators for differential-algebraic systems of index 2

MARTIN ARNOLD[1] AND ANDER MURUA[2,*]

[1] *DLR German Aerospace Centre, Institute of Robotics and System Dynamics, P.O. Box 1116, D – 82230 Wessling, Germany*
E-mail: martin.arnold@dlr.de

[2] *Konputazio Zientziak eta A. A., Informatika Fakultatea EHU/UPV, Donostia/San Sebastián, Spain*
E-mail: ander@si.ehu.es

Non-stiff differential-algebraic equations (DAE's) can be solved efficiently by partitioned methods that combine well known non-stiff integrators from ODE–theory with an implicit method to handle the algebraic part of the system. In the present paper we consider partitioned one-step and partitioned multi-step methods for index-2 DAE's in Hessenberg form and the application of these methods to constrained mechanical systems. The methods are presented from a unified point of view. The comparison of various classes of methods is completed by numerical tests for benchmark problems from the literature.

Subject classification: AMS(MOS) 65L05.

Keywords: DAE, higher index, partitioned methods, non-stiff DAE's.

1 Introduction

Most numerical methods for differential-algebraic equations (DAE's) are based on standard methods from the theory of ordinary differential equations (ODE's). It is well known that the robust and numerically stable application of these ODE–methods to higher index DAE's has to be based on the structure of the DAE (e. g. index, semi-explicit form, Hessenberg form, ... [6]). In "general purpose" codes like DASSL ([6]) or RADAU5 ([11]) the structure of the DAE is considered in the implementation of the method and also (choosing appropriate input parameters) at run-time ([17], [9, Chapters 7 and 8]).

An alternative are *partitioned* methods that consider the DAE's structure already at the level of the construction of the discretization method and combine different methods for different parts of the DAE. In the present paper we consider partitioned methods for

*The work of this author has been partially supported by the University of the Basque Country, research grant UPV140.226–EA105/96.

non-stiff index-2 DAE's in Hessenberg form:

$$\begin{aligned} y'(t) &= \varphi(y(t),z(t)) \, , \\ 0 &= \gamma(y(t)) \, . \end{aligned} \tag{1}$$

These methods combine efficient non-stiff integrators from ODE–theory with a method to handle the algebraic part of (1).

The prototype of these methods is the *half-explicit Euler method* ([11, p. 519])

$$\begin{aligned} y_{n+1} &= y_n + h\varphi(y_n,z_n) \, , \\ 0 &= \gamma(y_{n+1}) \end{aligned} \tag{2}$$

that defines the algebraic components $z_n := \zeta$ as solution of the system of n_z non-linear equations

$$0 = \gamma(y_n + h\varphi(y_n,\zeta)) \tag{3}$$

such that y_{n+1} may be obtained in an explicit way from the differential part of (2).

Throughout the present paper we suppose that the initial value problem

$$y(0) = y_0 \; , \quad z(0) = z_0$$

for (1) has a sufficiently differentiable solution $y : [0,T] \to \mathbb{R}^{n_y}$, $z : [0,T] \to \mathbb{R}^{n_z}$. For all (η,ζ) in a small neighbourhood of the solution trajectory the functions $\varphi(\eta,\zeta)$ and $\gamma(\eta)$ are supposed to be sufficiently differentiable and to satisfy the *index-2 condition*

$$\text{“ } \gamma_y(\eta)\varphi_z(\eta,\zeta) \;\text{ is non-singular ”.} \tag{4}$$

Obviously, the initial values must satisfy the algebraic constraint of (1), and in addition, they have to satisfy the *hidden constraint*

$$0 = \frac{d}{dt}\gamma(y(t)) = \gamma_y(y)y'(t) = \gamma_y(y)\varphi(y,z) \, , \tag{5}$$

being obtained by differentiating the original constraint w. r. t. t. That is, the initial values (y_0, z_0) must be *consistent*.

The first partitioned methods for index-2 DAE's (1) were introduced by Hairer et al. ([9, pp. 20f]) who considered *half-explicit Runge–Kutta* (HERK) methods that combine an s-stage explicit Runge–Kutta method from ODE–theory for the differential part of (1) with the solution of s systems of n_z non-linear equations to obtain the stage vectors Z_{ni} of the algebraic components. Lubich ([12]) combined the approach (2) with Gragg's extrapolation method to get a very efficient *half-explicit extrapolation method* for (1). In the field of multi-step methods Arévalo et al. ([1]) proposed so-called *β-blocked methods* to allow a stable numerical integration of index-2 DAE's (1) by non-stiff linear multi-step methods from ODE–theory. The state of the art in 1996 (including the partitioned HERK methods described below) is summarized in Sec. VII.6 of the monograph [11].

To get partitioned methods for (1) that are competitive to fully implicit methods (BDF, implicit Runge–Kutta) the partitioned methods have to be based on the most efficient methods from ODE–theory. But up to 1995 this basic requirement was met only by Lubich's half-explicit extrapolation method that is successfully applied to the dynamical simulation of constrained mechanical systems. The open problems included:

(P1) If standard high order explicit Runge–Kutta methods from ODE–theory are extended to HERK methods as described in [9] then the order drops down to 2 ([9, p. 70]).

(P2) Second order ODE's may be solved efficiently by Runge–Kutta–Nyström methods ([10, Sec. II.14]). Is there an equivalent for index-2 DAE's (1) originating from 2nd order DAE's?

(P3) Higher order Adams methods get unstable if they are applied to the index-2 DAE (1) in the standard way ([11, Sec. VII.3]) or in its "β-blocked" form ([1]).

In the last two years there has been substantial progress in the solution of problems (P1)–(P3):

Modifying the approach of Hairer et al. the new class of *partitioned half-explicit Runge–Kutta methods with explicit stages* (PHERK methods) was introduced independently by both authors of the present paper ([2, 14]). The 5th order PHERK methods PHEM56 ([14]) and HEDOP5 ([2]) were constructed on the basis of the very efficient 5th order explicit Runge–Kutta method of Dormand and Prince ([10, pp. 178f]).

The main application of partitioned methods for (1) are the model equations of non-stiff constrained mechanical systems in index-2 formulation, which are usually of the form

$$\begin{aligned} q' &= v\,, \\ M(q)v' &= f(q,v) - G(q)^T\lambda\,, \\ 0 &= G(q)v \end{aligned} \tag{6}$$

with $y = (\, q^T \,,\, v^T \,)^T$, $z = \lambda$. The index-1 formulation of (6) (obtained by replacing the algebraic constraint $0 = G(q)v$ by its time derivative) defines implicitly a second order ODE. Recently ([16]), it has been shown how the special structure of these systems can be exploited by applying Runge–Kutta–Nyström methods to the index-1 formulation of (6). In Sec. 3 below we will illustrate how these ideas may be extended to the index-2 case (*half-explicit Runge–Kutta–Nyström methods*, HERKN methods).

In [4] we proposed the class of partitioned linear multi-step methods for index-2 DAE's (1). Following this approach Adams–Moulton methods of arbitrary high order may be extended to numerically stable methods for index-2 DAE's (1) without any order reduction (*partitioned Adams methods*).

The main purpose of the present paper is to summarize the new classes of partitioned methods from a unified point of view, and to compare these methods in numerical tests. Furthermore, we include some new results, that have not been published before.

The remaining part of the paper is split in 3 sections describing PHERK methods (Sec. 2), HERKN methods (Sec. 3), and partitioned Adams methods (Sec. 4), respectively, and the final Sec. 5 that summarizes the results of numerical comparisons for two benchmark problems from the literature.

2 Partitioned half-explicit Runge–Kutta methods

The application to (1) of a step $(y_0, z_0) \to (y_1, z_1)$ of an s-stage *partitioned half-explicit Runge–Kutta (PHERK)* method is defined by

$$
\begin{aligned}
Y_1 &= y_0, \quad Z_1 \;=\; z_0, \\
i &= 2, \ldots, s+1, \\
&\qquad Y_i \;=\; y_0 + h \sum_{j=1}^{i-1} a_{ij} \varphi(Y_j, Z_j), \\
&\qquad \bar{Y}_i \;=\; y_0 + h \sum_{j=1}^{i} \bar{a}_{ij} \varphi(Y_j, Z_j), \quad \gamma(\bar{Y}_i) = 0, \\
y_1 &= Y_{s+1}, \quad z_1 \;=\; Z_{s+1},
\end{aligned}
\tag{7}
$$

where it is assumed that $a_{s+1,i} = \bar{a}_{s,i}$ for all i, so that the numerical approximation $y_1 = Y_{s+1}$ satisfies the algebraic constraint of (1). However the hidden constraint (5) will not hold in general.

In contrast to the HERK methods of [9] we do *not* require $\gamma(\bar{Y}_1) = 0$ in (7). Furthermore, the PHERK methods have additional parameters $\bar{a}_{ij}$ that may be chosen to get a high order of convergence. Because of these two reasons PHERK methods do not suffer from the order reduction phenomenon mentioned in the Introduction (see (P1)).

Similar to the half-explicit Euler method (2) the stage vector Y_i in (7) is computed explicitly at the ith stage, and Z_i is obtained by solving the equation that results replacing in $\gamma(\bar{Y}_i) = 0$ the expression for $\bar{Y}_i$. Note that the Jacobian with respect to Z_i of such algebraic equation is $h\bar{a}_{ii}\gamma_y(\bar{Y}_i)\varphi_z(Y_i, Z_i)$. Thus, we will always assume that $\bar{a}_{ii} \neq 0$ for $i \geq 2$. This condition guarantees the existence of the numerical solution as stated in Lemma 1 below.

PHERK methods are particularly efficient for constrained mechanical systems of the form (6). This is due to the fact that the algebraic system of equations needed to solve at each stage is reduced to a linear system of the form

$$
\begin{pmatrix} M(Q_i) & G(Q_i)^T \\ G(\bar{Q}_i) & 0 \end{pmatrix} \begin{pmatrix} V_i' \\ \Lambda_i \end{pmatrix} = \begin{pmatrix} f(Q_i, V_i) \\ r_i \end{pmatrix}
\tag{8}
$$

such that the iterative solution of non-linear equations is completely avoided in this application.

An interesting particular case of PHERK methods (7) is that of methods satisfying

$$
\bar{a}_{ij} = a_{i+1,j}, \quad 2 \leq i \leq s, \quad 1 \leq j \leq i,
\tag{9}
$$

so that $\bar{Y}_i = Y_{i+1}$ for $2 \leq i \leq s$ ([2]). This simplification reduces the numerical work per step. In general, some linear combinations are avoided, but some more savings are possible when applied to constrained mechanical systems of the form (6) (the stage vector Q_i in (8) coincides with $\bar{Q}_{i-1}$, which allows saving some evaluations of $G(q)$).

Lemma 1
Let us consider initial values (y_0, z_0) *satisfying the algebraic constraint of the system (1), and assume that (4) holds in a neighbourhood of* (y_0, z_0)*. Then, if*

$$\varepsilon = \max\left(|h|, ||\gamma_y(y_0)\varphi(y_0, z_0)||\right)$$

is sufficiently small, the method (7) has for $h \le h_0$ *a locally unique solution, which smoothly depends on* h *and* (y_0, z_0)*. Furthermore,*

$$Y_i = y_0 + O(h), \quad Z_i = z_0 + O(\varepsilon).$$

In order to understand the behaviour of the application of successive steps of the PHERK methods, we investigate how the errors in the hidden constraint are propagated:

Lemma 2
Under the conditions of Lemma 1,

$$\gamma_y(y_1)\varphi(y_1, z_1) = \sigma\gamma_y(y_0)\varphi(y_0, z_0) + O(h), \quad \text{where } \sigma = -\sum_{i=2}^{s+1} w_{s+1,i}\,\bar{a}_{i1}, \tag{10}$$

and $(w_{ij})_{i,j=2}^{s+1}$ *is the inverse of the matrix* $(\bar{a}_{ij})_{i,j=2}^{s+1}$*.*

Proof
Differentiating the algebraic equation $0 = \gamma(\bar{Y}_i)$ in (7) with respect to h and taking the limit as $h \to 0$, one obtains $0 = \sum \bar{a}_{ij}\gamma_y(y_0)\varphi(y_0, Z_j^0)$, where $Z_i^0 = \lim_{h\to 0} Z_i$, so that $Z_i = Z_i^0 + O(h)$. Taking into account that $Z_1 = Z_1^0 = z_0$, we arrive at

$$\gamma_y(y_0)\varphi(y_0, Z_i^0) = -\sum_{j=2}^{i} w_{ij}\bar{a}_{j1}\gamma_y(y_0)\varphi(y_0, z_0),$$

which together with $y_1 = y_0 + O(h)$, $z_1 = Z_{s+1} = Z_{s+1}^0 + O(h)$ leads to (10). □

It is clear from Lemma 2 that the method cannot be convergent if $|\sigma| > 1$, and the best propagation properties correspond to methods with $\sigma = 0$.

The convergence of PHERK methods has been studied in [2] and [14] by following similar approaches. In [15], a new methodology is proposed, which in general gives sharper global error estimates, in particular for the algebraic components. In all of those approaches, series expansions indexed by certain classes of rooted trees with vertices of two different colours (*DAE2-trees* [9, 11]) are used to characterize the convergence behaviour in a systematic way. Here, we will restrict ourselves to give sufficient conditions for a PHERK method to be of order 5 for the differential component (Theorem 3 below).

It is interesting to note that, when the function γ in (1) is linear, the application of a PHERK method (7) is equivalent to the application of the *underlying explicit Runge–Kutta* method (that is, the Runge–Kutta method with coefficients a_{ij} and $b_i = a_{s+1,i}$) to the equivalent index-1 DAE that is obtained from (1) replacing the original constraint $0 = \gamma(y)$ by the hidden constraint (5). This implies that for a given pth order PHERK method the underlying explicit Runge–Kutta method has to be of order p as well. This suggest dividing the construction of pth order PHERK methods in two steps: First, choose a standard pth order explicit Runge–Kutta method. This determines the coefficients a_{ij} of the PHERK method. And second, determine the coefficients $\bar{a}_{ij}$ in order that the corresponding PHERK method is of the required order when applied to systems of the form (1). Theorem 3 below gives the theoretical basis to do that for order 5. Let us first fix some notation:

$$\begin{aligned} c_i &= \sum_{j=1}^{i-1} a_{ij}, \quad \bar{c}_i = \sum_{j=1}^{i} \bar{a}_{ij}, \quad d_i = \sum_{j=2}^{i-1} a_{ij} c_j - \frac{c_i^2}{2}, \quad 2 \le i \le s+1, \\ b_i &= a_{s+1,i} = \bar{a}_{si}, \quad 1 \le i \le s, \quad \bar{b}_i = \sum_{j=i}^{s} b_j (1-c_j) w_{ji}, \quad 2 \le i \le s. \end{aligned}$$

Note that the last equality above is equivalent to $\sum_j \bar{b}_j \bar{a}_{ji} = b_i(1-c_i)$, $(2 \le i \le s)$. We will consider PHERK methods whose underlying Runge–Kutta method satisfies the standard simplifying assumption

$$\sum_{j=i+1}^{s} b_j a_{ji} = b_i(1-c_i), \quad 1 \le i \le s. \tag{11}$$

Theorem 3
Let us consider the PHERK scheme (7) such that its underlying Runge–Kutta method is of order 5 for ODE's and satisfies (11).

1. *If $|\sigma| < 1$ and the following conditions hold,*

$$\sum_{i=2}^{s} \bar{b}_i \bar{c}_i^{l-1} = \frac{1}{l}, \quad 2 \le l \le 5, \quad \sum_{i,j=2}^{s} \bar{b}_i \bar{c}_i \bar{a}_{ij} d_j = 0, \tag{12}$$

$$\sum_{j=2}^{i} \bar{a}_{ij} c_j = \frac{\bar{c}_i^2}{2}, \quad \sum_{j=2}^{i} \bar{a}_{ij} (c_j^2 + \mu d_j) = \frac{\bar{c}_i^3}{3}, \quad 2 \le i \le s+1, \tag{13}$$

where μ is a certain real number, then the PHERK method is of order 5 for the differential component y, and of order 3 for the algebraic component z.

2. *If in addition*

$$\sum_{i=2}^{s} w_{s+1,i} \bar{c}_i^4 = 4, \quad \sum_{i,j=2}^{s} w_{s+1,i} \bar{c}_i \bar{a}_{ij} d_j = 0, \tag{14}$$

hold, then the approximation for the algebraic component is of order 4.

Proof
It can be shown by following the techniques presented in ([15]) that the hypothesis of the first part of the theorem implies that

$$\varepsilon_{n+1} = |\sigma| \varepsilon_n + O(h^3 + h^2 \varepsilon_n + h \varepsilon_n^2), \tag{15}$$

$$(y_{n+1} - y(t_{n+1})) = (I + O(h))(y_n - y(t_n)) + O(h^6 + h^3 \varepsilon_n + h \varepsilon_n^2), \tag{16}$$

where $(y(t), z(t))$ is the solution of (1) with consistent initial values (y_0, z_0), (y_n, z_n) is the approximation given by the PHERK method at $t_n = t_0 + nh$, and $\varepsilon_n = ||\gamma_y(y_n)\varphi(y_n, z_n)||$. It is not difficult to see that (15) and (16) imply the first statement of the theorem. Moreover, it can be shown that if the additional condition (14) hold, then the h^3-term in (15) can be replaced by h^4, which leads to the second statement of the theorem. □

It can be seen from Theorem 3 that for each 6-stage 5th order explicit Runge–Kutta method satisfying (11), there exists a 7-parametric family (with $\bar{c}_2, \bar{c}_3, \bar{c}_4, \bar{c}_5, \bar{c}_7, \bar{a}_{77}, \sigma$ as free parameters) of 6-stage 5th order PHERK methods that extend that Runge–Kutta scheme. The two PHERK methods PHEM56 ([14]) and HEDOP5 ([2]) that extend the successful 6-stage 5th order explicit Runge–Kutta method DOPRI5 are particular instances of that family of methods. For both methods, $\sigma = 0$ and the additional condition (14) is fulfilled. According to our numerical experiments, both methods perform very similarly (when γ in (1) is linear, they give exactly the same results). However, HEDOP5 satisfies (9) such that, as mentioned above, the numerical effort per step of integration is slightly reduced (especially for systems of the form (6)).

The extension of DOPRI5 to a PHERK method satisfying (9) is possible because (13) (with $\bar{a}_{ij} = a_{i+1,j}$) is already satisfied by DOPRI5 ((12) is a consequence of the underlying Runge–Kutta method being of order 5, since (9) implies $\bar{c}_i = c_{i+1}$ and $\bar{b}_i = b_{i+1}$). Then, in order to determine such an extension of DOPRI5, only the coefficients $\bar{a}_{7i}$ $(1 \le i \le 7)$ remain to be determined. They must be chosen in such a way that $|\sigma| < 1$ and (13) also holds for $i = s+1 = 7$. In order to achieve order 4 for the algebraic component, (14) must be required as well. For HEDOP5, $\sigma = 0$, $\bar{c}_7 = 19/20$, $\bar{a}_{77} = -3280/75413$ are chosen, which uniquely determine the remaining coefficients:

$$\begin{aligned}
\bar{a}_{71} &= \frac{-18611506045861}{19738176307200}, \quad \bar{a}_{72} = \frac{59332529}{14479296}, \quad \bar{a}_{73} = \frac{-2509441598627}{893904224850}, \\
\bar{a}_{74} &= \frac{2763523204159}{3289696051200}, \quad \bar{a}_{75} = \frac{-41262869588913}{116235927142400}, \quad \bar{a}_{76} = \frac{46310205821}{287848404480}.
\end{aligned}$$

For the efficient application of these methods, it is convenient to have available a cheap 4th order embedded approximation (for local error estimation) and continuous formula. In [14] it is shown how such 4th order cheap formulae can be obtained for the PHERK methods of our 7-parametric family of 5th order methods.

3 Half-explicit Runge–Kutta–Nyström methods

Let us consider the application of a PHERK method (7) to the model equations (6) of constrained mechanical systems. The numerical work needed to perform each stage of the method essentially corresponds to evaluating and solving the linear system (8). The important observation here is that the coefficient matrix of that linear system only depends on Q_i and $\bar{Q}_i$.

We will show that more efficient methods may be obtained by modifying the PHERK scheme as follows: If the two components q and v of the differential variable are treated in a different way, then it is possible to have $Q_i = Q_{i-1}$ and $\bar{Q}_i = \bar{Q}_{i-1}$ for some values of i, so that considerable savings of the numerical work per stage can be achieved. This idea is in fact an extension of those proposed in [16], where it is shown that explicit Runge–Kutta–Nystöm methods can be constructed that are more efficient than explicit Runge–Kutta methods when applied to the index-1 formulation of mechanical systems (6). Here, we will present these new methods, the *half-explicit Runge–Kutta–Nyström* methods, in a more general setting.

Let us consider the following particular case of the index-2 DAE (1)

$$\begin{aligned} q' &= v\ , \\ v' &= \psi(q,v,z)\ , \\ 0 &= \gamma(q,v)\ . \end{aligned} \tag{17}$$

Clearly, the system (6) is of this form. Note that the index-1 DAE obtained from replacing the original constraint of (17) by the hidden constraint implicitly defines a second order ODE. We will extend explicit Runge–Kutta–Nyström methods for general second order ODE's in a similar way that PHERK methods extend explicit Runge–Kutta methods (see also (P2) in the Introduction).

The application to (17) of a step $(q_0, v_0, z_0) \to (q_1, v_1, z_1)$ of an s-stage half-explicit Runge–Kutta–Nyström method is defined by

$$\begin{aligned} Q_1 &= q_0, \quad V_1 = v_0, \quad V_1' = \psi(q_0, v_0, z_0), \\ i &= 2, \ldots, s+1, \\ &\qquad Q_i = q_0 + hc_iv_0 + h^2 \sum_{j=1}^{i-1} \alpha_{ij} V_j', \quad V_i = v_0 + h \sum_{j=1}^{i-1} a_{ij} V_j', \\ &\qquad \bar{Q}_i = q_0 + h\bar{c}_iv_0 + h^2 \sum_{j=1}^{i-1} \bar{\alpha}_{ij} V_j', \quad \bar{V}_i = v_0 + h \sum_{j=1}^{i} \bar{a}_{ij} V_j', \\ &\qquad V_i' = \psi(Q_i, V_i, Z_i), \quad 0 = \gamma(\bar{Q}_i, \bar{V}_i), \\ q_1 &= Q_{s+1}, \quad v_1 = V_{s+1}, \quad z_1 = Z_{s+1}, \end{aligned} \tag{18}$$

where $a_{s+1,i} = \bar{a}_{s,i}$ and $\alpha_{s+1,i} = \bar{\alpha}_{s,i}$ for all i, so that the numerical approximation satisfies the algebraic constraint of (17). We also assume that $\bar{a}_{ii} \neq 0$ which allows us to obtain an existence and uniqueness result analogous to Lemma 1. An analogue of Lemma 2 also holds for the method (18) with σ defined as in (10).

It is interesting to note that the application of the PHERK method corresponding to the coefficients $a_{ij}, \bar{a}_{ij}$ to the system (17) is equivalent to the application of a HERKN scheme with coefficients $c_i = \sum a_{ij}$, $\bar{c}_i = \sum \bar{a}_{ij}$, $\alpha_{ij} = \sum a_{ik}a_{kj}$, $\bar{\alpha}_{ij} = \sum \bar{a}_{ik}a_{kj}$. We consider HERKN methods with general coefficients (that is, not generated from a PHERK method) mainly because this allows us to choose $Q_i = Q_{i-1}$ and $\bar{Q}_i = \bar{Q}_{i-1}$ for some values of i, so that the numerical work of evaluating and solving the linear system (8) corresponding to the ith stage is considerably reduced.

The procedure we propose to construct pth order HERKN methods that are efficient for systems of the form (6) is the following: First, determine the coefficients c_i, α_{ij}, a_{ij} such that the corresponding Runge–Kutta–Nyström method is of order p for general second order ODE's and $Q_i = Q_{i-1}$ for some stages i (for instance, choosing one of those proposed in [16]). And second, determine the coefficients $\bar{c}_i, \bar{\alpha}_{ij}, \bar{a}_{ij}$ in order to extend the RKN method to a pth order HERKN method with $\bar{Q}_i = \bar{Q}_{i-1}$ for the same stages i.

The theorem below gives sufficient conditions for such an extension to be of 5th order. There we keep the notation adopted in Section 2. Its proof, which we omit here, was obtained using the techniques presented in [15] for PHERK methods. Rooted trees with three types of vertices (instead of two) are used. As in the proof of Theorem 3, simplifying assumptions are used in a way that is usual in the construction of Runge–Kutta type methods.

Theorem 4
Let us consider the scheme (18) such that its underlying RKN method (determined by the coefficients c_i, α_{ij}, a_{ij}) is of order 5 for general second order ODE's, and (11) and the following condition are fulfilled,

$$\sum_{j=1}^{i-1} \alpha_{ij} = \frac{c_i^2}{2} \quad (2 \le i \le s+1),$$

1. *If $|\sigma| < 1$, (12), (13) and the following conditions hold,*

$$\bar{\alpha}_{si} = b_i(1-c_i), \quad 1 \le i \le s, \quad \sum_{i,j=2}^{s} \bar{b}_i \bar{\alpha}_{ij} \bar{c}_i^2 = \frac{1}{15}, \quad \sum_{i,j=2}^{s} \bar{b}_i \bar{\alpha}_{ij} d_j = 0,$$

$$\sum_{j=2}^{i-1} \bar{a}_{ij} c_j = \frac{\bar{c}_i^3}{6} + \nu \sum_{j=2}^{i-1} \bar{a}_{ij} d_j, \quad 2 \le i \le s+1, \quad (\textit{for some real } \nu),$$

   then the HERKN method is of order 5 for the differential component y, and of order 3 for the algebraic component z.

2. *If in addition (14) and*

$$\sum_{i,j=2}^{s} w_{s+1,i} \bar{\alpha}_{ij} \bar{c}_j^2 = 1/12, \qquad \sum_{i,j=2}^{s} w_{s+1,i} \bar{\alpha}_{ij} d_j = 0,$$

   hold, then the approximation for the algebraic component is of order 4.

To compare our new approach with the PHERK methods of Sec. 2 we constructed on the basis of Theorem 4 the 5th order HERKN method HERKN5 that extends the 5th order RKN method `rkn5456` constructed in [16]. This method, which satisfies $Q_6 = Q_7$, $\bar{Q}_6 = \bar{Q}_7$, is provided with cheap 4th order embedded approximation and dense output formula.

4 Partitioned Adams methods

Linear multi-step methods for index-2 DAE's (1) may be stabilized using a BDF approximation for the computation of the algebraic components z ([4]). In this section we use this approach to construct partitioned k-step Adams methods that are numerically stable for all $k \ge 2$ and converge in y with the classical order $q = k+1$ from ODE–theory.

If a (classical) implicit linear multi-step method with generating polynomials $\varrho(\xi) = \sum_j \alpha_j \xi^j$ and $\sigma(\xi) = \sum_j \beta_j \xi^j$ is applied to (1) then y_{n+k} and z_{n+k} are obtained from ([11, Sec. VII.3])

$$\begin{array}{rcl} \frac{1}{h}\sum_{j=0}^{k} \alpha_j y_{n+j} & = & \sum_{j=0}^{k} \beta_j \varphi(y_{n+j}, z_{n+j}) \ , \\ 0 & = & \gamma(y_{n+k}) \ . \end{array} \tag{19}$$

Solving the differential part of (19) w. r. t. $y_{n+k} =: \eta$ and inserting this expression into the algebraic part we obtain similarly to (3) the system of n_z non-linear equations

$$0 = \gamma(\Phi(\eta, Y_n, Z_n; \varrho, \sigma, h)) \tag{20}$$

with

$$\begin{aligned} &\Phi(\eta, Y_n, Z_n; \varrho, \sigma, h) := \\ &\quad = \frac{1}{\alpha_k}\Big(-\sum_{j=0}^{k-1} \alpha_j y_{n+j} + h \sum_{j=0}^{k-1} \beta_j \varphi(y_{n+j}, z_{n+j}) + h\beta_k \varphi(\eta, \Psi(\eta, Y_n, Z_n; \varrho, \sigma, h))\Big) \end{aligned} \tag{21}$$

and the initial values

$$Y_n := (y_n, \ldots, y_{n+k-1})^T \ , \quad Z_n := (z_n, \ldots, z_{n+k-1})^T \ .$$

Eqs. (20) and (21) define implicitly the function $\Psi(\eta, Y_n, Z_n; \varrho, \sigma, h)$ such that (19) may be formally rewritten as

$$z_{n+k} = \Psi(y_{n+k}, Y_n, Z_n; \varrho, \sigma, h)$$

where y_{n+k} is a fixed point of the mapping

$$\eta \mapsto \Phi(\eta, Y_n, Z_n; \varrho, \sigma, h) \ . \tag{22}$$

For sufficiently small stepsizes $h > 0$ this fixed point and the function Ψ are (locally) uniquely defined if the initial values Y_n, Z_n are sufficiently close to the analytical solution ([11, Theorem VII.3.1]).

Gear et al. ([8]) proposed a predictor-corrector scheme for the application of (classical) linear multi-step methods to non-stiff higher index DAE's. In our setting such a predictor-corrector scheme is defined by the fixed point iteration

$$y_{n+k}^{(l+1)} := \Phi(y_{n+k}^{(l)}, Y_n, Z_n; \varrho, \sigma, h) \ , \quad (l \geq 0)$$

with the predictor $y_{n+k}^{(0)}$ as initial value. As part of the evaluation steps in a P(EC)^L– or P(EC)^{L}E–scheme systems of the form (20) have to be solved w. r. t. Ψ to get

$$z_{n+k}^{(l)} := \Psi(y_{n+k}^{(l)}, Y_n, Z_n; \varrho, \sigma, h)$$

and $\varphi(y_{n+k}^{(l)}, z_{n+k}^{(l)})$. Similar to the half-explicit one-step methods of Sections 2 and 3 the predictor-corrector scheme combines an explicit update formula for y_{n+k} with the solution of systems of n_z non-linear equations to get z_{n+k}. In the application to the model equations (6) of constrained mechanical systems $z_{n+k}^{(l)}$ is obtained as solution of a system of n_z linear equations that has the basic form (8) with Q_i, $\bar{Q}_i$ being substituted by $q_{n+k}^{(l+1)}$. Since both G and G^T are evaluated with one and the same argument $q_{n+k}^{(l+1)}$ the matrix of coefficients in this system of linear equations is symmetric such that the numerical effort may be further reduced.

From (22) we see that y_{n+k} in (19) is completely defined by Y_n, Z_n, ϱ, σ, and h and does not depend on z_{n+k}. Therefore, y_{n+k} remains unchanged if we consider instead of (19) the *partitioned linear multi-step method* ([4])

$$\begin{aligned} y_{n+k} &= \Phi(y_{n+k}, Y_n, Z_n; \varrho, \sigma, h) \ , \\ z_{n+k} &= \Psi(y_{n+k}, Y_n, Z_n; \hat{\varrho}, \hat{\sigma}, h) \end{aligned} \tag{23}$$

with generating polynomials (ϱ, σ) and $(\hat{\varrho}, \hat{\sigma})$. Eqs. (23) cover the classical approach (19) as special case with $\varrho = \hat{\varrho}$ and $\sigma = \hat{\sigma}$.

The numerical efforts of (19) and (23) are nearly identical if both methods are implemented as predictor-corrector schemes in P(EC)^{L}E–mode ([4]). The important difference between (19) and (23) are the additional coefficients $\hat{\alpha}_j$, $\hat{\beta}_j$ of the polynomials $\hat{\varrho}$ and $\hat{\sigma}$ in (23). These new coefficients may be chosen to guarantee both stability and a high order of convergence for (23) such that the well known instability phenomenon of the classical approach (19) is overcome (see (P3) in the Introduction, and [11, Theorem VII.3.6], [1]).

Theorem 5 [4, Theorem 3b]
Consider a partitioned linear multi-step method (23) that is based

- *on a zero-stable implicit linear multi-step method with generating polynomials* (ϱ, σ) *and classical order* p *([10, Chapter III]) and*
- *on an implicit linear multi-step method with generating polynomials* $(\hat{\varrho}, \hat{\sigma})$ *and classical order* $\hat{p}$ *that is strictly stable at infinity (i. e.* $\hat{\sigma}(\xi) = 0 \Rightarrow |\xi| < 1$*).*

This partitioned method is convergent with order $q = \min(p, \hat{p} + 1)$ *in* y *and with order* $\hat{q} = \min(p, \hat{p})$ *in* z *if* $\hat{q} \geq 2$ *and the initial values satisfy*

$$\|y_i - y(t_i)\| = \mathcal{O}(h^{p+1}) \ , \ \ \|z_i - z(t_i)\| = \mathcal{O}(h^{\hat{p}}) \ , \ \ (i = 0, 1, \ldots, k-1) \ .$$

Proof
The proof is given in [4], it is based on the convergence analysis for the classical approach (19) that was carried out by Hairer and Wanner in Sec. VII.3 of [11]. In addition to standard arguments from ODE–theory we have to study the error propagation in z. To this end (20) is linearized along the analytical solution to obtain

$$\begin{aligned} 0 = \frac{h}{\hat{\alpha}_k} \sum_{j=0}^{k} \hat{\beta}_j \gamma_y(y_{n+j}) \varphi(y_{n+j}, z_{n+j}) \, + \\ + \mathcal{O}(h) \sum_{j=0}^{k-1} \|y_{n+j} - y(t_{n+j})\| + \mathcal{O}(h^2) \sum_{j=0}^{k} \|z_{n+j} - z(t_{n+j})\| + \mathcal{O}(h^{\hat{p}+1}) \ . \end{aligned} \tag{24}$$

As for the PHERK methods (see Lemma 2) the errors in the sequence $(\gamma_y(y_n)\varphi(y_n, z_n))_{n \geq 0}$ are propagated in a stable way if a contractivity condition is satisfied. In the case of partitioned linear multi-step methods (23) we see from (24) that a stable error propagation is guaranteed if all roots of $\hat{\sigma}$ are within the unit circle. □

Remark 1
a) The assumptions of Theorem 5 are satisfied for arbitrary convergent implicit linear k-step methods (ϱ, σ) from ODE–theory with order $p \geq 2$ that are combined with $(\hat{\varrho}, \hat{\sigma})$ being defined by k-step BDF ($\hat{\sigma}(\xi) = \beta_k \xi^k$). In the non-stiff case the combination of an implicit k-step Adams method (ϱ, σ) with k-step BDF $(\hat{\varrho}, \hat{\sigma})$ was found to be particularly useful because of the high order and the small error constants of the Adams methods. These partitioned Adams methods *converge for any $k \geq 2$ with order $q = k+1$ in y and order $\hat{q} = k$ in z.*

b) Throughout the present paper we restrict ourselves to index-2 DAE's (1) in Hessenberg form. However, Theorem 5 may be extended to semi-explicit index-2 DAE's of the more general form

$$\begin{aligned} y'(t) &= \varphi(y(t), z(t)) \,, \\ 0 &= \gamma(y(t), z(t)) \end{aligned} \tag{25}$$

with a Jacobian γ_z that has constant rank in a neighbourhood of the solution. In the special case $\varrho = \hat{\varrho}$, $\sigma = \hat{\sigma}$ this fact was observed by Hairer and Wanner who pointed out that (25) may be transformed to Hessenberg form with a transformation leaving (19) invariant ([11, p. 482]). In general, the partitioned methods (23) are not *invariant w. r. t. this transformation but nevertheless convergence with order q in y and order $\hat{q}$ in z can be proved for the application of partitioned Adams methods to (25) ([3, Remark 34]).*

5 Numerical tests

The implementation of partitioned methods for index-2 DAE's is based on the ideas that Lubich et al. used in their extrapolation code MEXX ([13]). In this section we present the results of a numerical comparison of the partitioned methods described in Sections 2 – 4.

Up to now partitioned methods are mainly applied to the model equations of constrained mechanical systems and the currently available variable stepsize codes use explicitly the special structure of these equations. In classical mechanics the model equations are formulated as index-3 DAE's with holonomic constraints

$$g(q) - g(q_0) = 0 \tag{26}$$

on the level of position coordinates ([11, p. 464]). To avoid numerical instability we substitute (26) by its time derivative

$$0 = \frac{d}{dt}(g(q(t)) - g(q_0)) = \frac{\partial}{\partial q} g(q(t)) q'(t) = G(q)v$$

with $G(q) := \frac{\partial}{\partial q} g(q)$ such that the index is reduced to 2 before discretization (see (6)). The original holonomic constraints (26) are invariants of the index-2 formulation (6). A numerical solution satisfying these invariants is obtained by projection on the manifold $\{\, q \,:\, g(q) - g(q_0) = 0 \,\}$ ([13]).

The half-explicit Runge–Kutta codes have been written on the basis of standard implementations of explicit Runge–Kutta methods ([10]), among other features they offer automatic stepsize control and dense output ([14]). Since all methods being considered in Sections 2 – 4 extend well known very efficient ODE–methods to index-2 DAE's (1) we refer to [10, Secs. II.10 and III.7] for a comparison of the basis methods in the ODE case. We want to illustrate with test results for benchmark problems that this efficiency is not lost by the extension to partitioned DAE methods. Furthermore, the test results show the differences between the various classes of partitioned methods.

In this survey paper we do not discuss the details of the implementation. The codes PHEM56 ([14], `ftp://ftp.unige.ch/pub/doc/math/mechanic/phem56.f`) and HEDOP5 ([2], part of MBSPACK, `ftp://ftp.mathematik.tu-darmstadt.de/pub/department/software/mbspack/`) are available in the Internet, the code HERKN5 may be obtained contacting the 2nd author. Up to now there is no variable stepsize implementation of partitioned Adams methods available.

In a first test we compare variable stepsize codes applied to a benchmark problem that models a wheel suspension being excited by a rough floor ([18]). The wheel suspension is described by $n_y = 28$ differential and $n_z = 12$ algebraic components. For various tolerances we compare the simulation results at $T = 0.6$ with a numerical solution that has been computed with high accuracy. In the work-precision diagrams in Figs. 1 and 2 the cpu-time is plotted vs. the (relative) global error that is measured in $\|.\|_\infty$.

Fig. 1 shows the results for the PHERK methods PHEM56 ("$\circ$") and HEDOP5 ("$+$"), for the HERKN method HERKN5 ("$*$"), and for HEM5 ([5], "$\times$"), the most efficient HERK method that was constructed following the approach of [9]. In Fig. 2 the PHERK method HEDOP5 ("$+$") is compared with the extrapolation code MEXX ([13], "$\diamond$") and with the general purpose code DASSL ([6], "$\circ$") that is applied to the Gear–Gupta–Leimkuhler formulation ([8]) of the model equations. Furthermore, we included the results for the application of the 5th order explicit Runge–Kutta method of Dormand and Prince ([10, p. 178f]) to the index-1 formulation of the model equations (MDOP5, "$*$").

Figs. 1 and 2 are characteristic of typical simulation results for non-stiff problems. The non-stiff integrators are competitive or even superior to DASSL. The extrapolation code MEXX gives excellent results for small tolerances but is less efficient than the half-explicit Runge–Kutta codes if less accuracy is requested. The new partitioned methods are substantially faster than the traditional HERK method HEM5, they reach the (excellent) performance of MDOP5. Depending on the problem PHEM56, HEDOP5, and HERKN5 may even be slightly faster than MDOP5.

Because of the savings in evaluations of G the code HEDOP5 is up to 10% faster than PHEM56, and the HERKN method is clearly superior to the PHERK methods. Note, however, that HERKN methods can not be applied to index-2 DAE's of the general form (1) since they are tailored to DAE's of the form (6).

The comparison is completed by Fig. 3 that shows results for three 5th order methods: for HEDOP5 ("$+$"), for the 4-step partitioned Adams method ("$\circ$") and for the 5-step BDF ("$*$"). The methods are implemented with fixed stepsize, for the multi-step methods we use a P(EC)^{2}E–scheme. The methods are applied to a truck model with $n_y = 18$ differential and $n_z = 2$ algebraic solution components ([7]). In the work-precision diagram we plotted for various stepsizes the numerical effort (that is measured in the number of systems (8) that have to be solved) vs. the error at $T = 1.0$. Because of the smaller error constant the partitioned Adams method is clearly superior to BDF. The results of HEDOP5 underline again that the PHERK method reaches the high efficiency of the method of Dormand and Prince.

Summarizing the test results of Figs. 1 – 3 we see that the solution of problems (P1)–

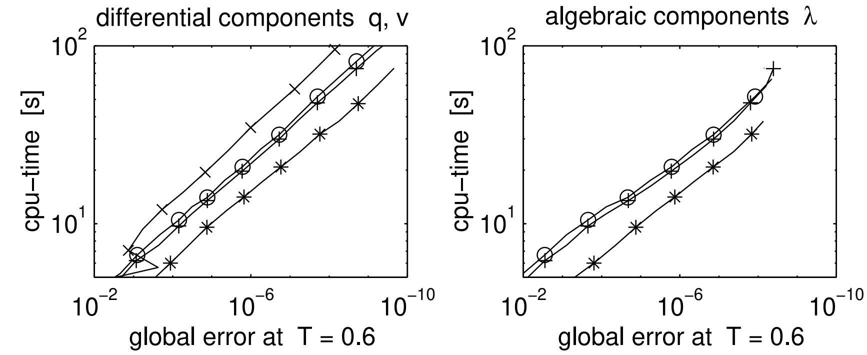

Figure 1: Comparison of integrators PHEM56 ("∘"), HEDOP5 ("+"), HERKN5 ("$*$"), and HEM5 ("×"). Benchmark Wheel suspension ([18]).

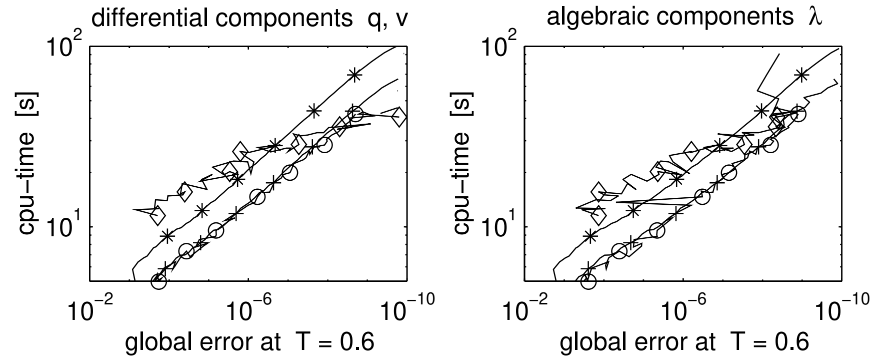

Figure 2: Comparison of integrators HEDOP5 ("+"), MEXX ("◇"), DASSL ("∘"), and MDOP5 ("$*$"). Benchmark Wheel suspension ([18]).

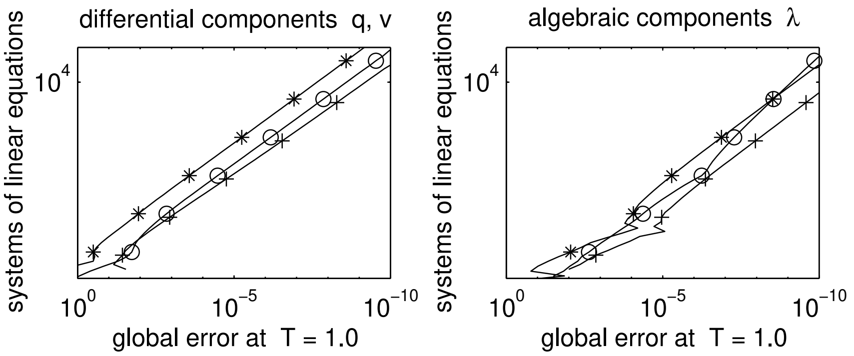

Figure 3: Numerical effort and discretization error of 5th order methods: BDF with $k = 5$ ("$*$"), partitioned Adams method with $k = 4$ ("∘"), and HEDOP5 ("+"). Benchmark Truck ([7]).

(P3) from the Introduction gives us very efficient partitioned integrators for non-stiff index-2 DAE's (1). There are minor — and problem dependent — differences between the new developed classes of methods but in the non-stiff case they are always competitive and often even superior to general purpose codes being based on BDF or implicit Runge–Kutta methods.

Summary

Standard non-stiff integrators from ODE–theory may be extended to partitioned methods for non-stiff index-2 DAE's in Hessenberg form. Both theoretical investigations and numerical tests prove the efficiency of partitioned half-explicit Runge–Kutta methods, of half-explicit Runge–Kutta–Nyström methods, and of partitioned Adams methods. The methods are especially advantageous in the application to constrained mechanical systems since non-linear systems of equations may be completely avoided in this special case.

For this application partitioned one-step methods have been implemented as variable stepsize codes PHEM56, HEDOP5, and HERKN5. Currently, these three 5th order methods and the extrapolation method MEXX ([13]) are the most efficient integrators for non-stiff index-2 DAE's in Hessenberg form.

Acknowledgements

We are grateful to Prof. E. Hairer (Genève) who gave us the initial motivation for the renewed interest in partitioned DAE integrators. Furthermore, we would like to thank him, Prof. Ch. Lubich (Tübingen), and Dr. B. Simeon (Darmstadt) for many helpful comments on the topic of the present paper and for the support in the development of the codes PHEM56 and HEDOP5.

References

[1] C. Arévalo, C. Führer, and G. Söderlind. Stabilized multistep methods for index 2 Euler–Lagrange DAEs. *BIT*, 36:1–13, 1996.

[2] M. Arnold. Half-explicit Runge–Kutta methods with explicit stages for differential-algebraic systems of index 2. *BIT Numerical Mathematics*, 38:415–438, 1998.

[3] M. Arnold. *Zur Theorie und zur numerischen Lösung von Anfangswertproblemen für differentiell-algebraische Systeme von höherem Index.* Fortschritt-Berichte VDI Reihe 20, Nr. 264. VDI–Verlag, Düsseldorf, 1998.

[4] M. Arnold. The stabilization of linear multistep methods for constrained mechanical systems. *Applied Numerical Mathematics*, to appear in 1998.

[5] V. Brasey. A half-explicit method of order 5 for solving constrained mechanical systems. *Computing*, 48:191–201, 1992.

[6] K.E. Brenan, S.L. Campbell, and L.R. Petzold. *Numerical solution of initial–value problems in differential–algebraic equations.* SIAM, Philadelphia, 2nd edition, 1996.

[7] E. Eich-Soellner and C. Führer. *Numerical Methods in Multibody Dynamics.* Teubner–

Verlag, Stuttgart, 1998.

[8] C.W. Gear, G.K. Gupta, and B. Leimkuhler. Automatic integration of Euler–Lagrange equations with constraints. *J. Comp. Appl. Math.*, 12&13:77–90, 1985.

[9] E. Hairer, Ch. Lubich, and M. Roche. *The numerical solution of differential-algebraic systems by Runge–Kutta methods.* Lecture Notes in Mathematics, 1409. Springer-Verlag, Berlin Heidelberg New York, 1989.

[10] E. Hairer, S.P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations. I. Nonstiff Problems.* Springer-Verlag, Berlin Heidelberg New York, 2nd edition, 1993.

[11] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations. II. Stiff and Differential-Algebraic Problems.* Springer-Verlag, Berlin Heidelberg New York, 2nd edition, 1996.

[12] Ch. Lubich. h^2-extrapolation methods for differential-algebraic systems of index 2. *Impact Comput. Sc. Eng.*, 1:260–268, 1989.

[13] Ch. Lubich, Ch. Engstler, U. Nowak, and U. Pöhle. Numerical integration of constrained mechanical systems using MEXX. *Mech. Struct. Mach.*, 23:473–495, 1995.

[14] A. Murua. Partitioned half-explicit Runge–Kutta methods for differential-algebraic systems of index 2. *Computing*, 59:43–61, 1997.

[15] A. Murua. Formal series and numerical integrators. Part II: Application to index 2 differential-algebraic systems. *Applied Numerical Mathematics*, to appear in 1998.

[16] A. Murua. Runge–Kutta–Nyström methods for general second order ODEs with application to multi-body systems. *Applied Numerical Mathematics*, to appear in 1998.

[17] L.R. Petzold and P. Lötstedt. Numerical solution of nonlinear differential equations with algebraic constraints II: practical implications. *SIAM J. Sci. Stat. Comput.*, 7:720–733, 1986.

[18] B. Simeon. On the numerical solution of a wheel suspension benchmark problem. *Comp. Appl. Math.*, 66:443–456, 1996.