Electronic Transactions on Numerical Analysis.
Volume 1, pp. 1-10, September 1993.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

ANALYSIS OF THE LINEARLY IMPLICIT MID–POINT RULE FOR DIFFERENTIAL–ALGEBRAIC EQUATIONS *

CLAUS SCHNEIDER†

Abstract. The error of the linearly implicit mid–point rule after $2m+1$ steps is expanded in powers of m^2. We prove that the well-known expansion for ordinary differential equations (an expansion in negative powers of m^2) is perturbed by additional terms with non-negative powers of m^2 for semi–explicit differential–algebraic equations of index one. Hence, extrapolation in m^{-2} will be of limited value only. The complete expansion shows these limits and, furthermore, can be used to derive an order 8 method of Rosenbrock type.

Key words. Differential–algebraic equations, linearly implicit mid–point rule, Rosenbrock–type methods, extrapolation.

AMS subject classifications. 65L05, 65B05, 58F99.

1. Introduction.

We study the behaviour of the linearly implicit mid–point (LIMP) rule when applied to semi–explicit differential–algebraic equations (DAE's)

$$y' = f(y,z), 0 = g(y,z). \tag{1.1}$$

We assume that f and g are sufficiently smooth functions and that $\partial g/\partial z$ has a bounded inverse in a neighbourhood of the exact solution, i.e., the DAE is of index one. Finally, the initial values $y(x_0) = y_0, z(x_0) = z_0$ should be consistent, i.e., $g(y_0, z_0) = 0$.

We derive an expansion of the error which shows how far extrapolation is justified and the limits of this procedure as well. Furthermore, we introduce some other means of eliminating error terms which finally yields a method of order 8.

In Section 2, we briefly present the linearly implicit mid–point rule, reformulate it as a Rosenbrock–type (ROW) method, and collect some known results. In Section 3, the expansion of the error is given and discussed, thus leading to extrapolation, elimination, and the highest achievable order. The proofs are provided in the last three sections where we have to analyze the 'growth' and the structure of order conditions.

2. The Linearly Implicit Mid–Point Rule for DAE's.

The LIMP rule (and its extrapolation) were introduced in [1]. An extension to differential–algebraic equations is straight–forward (cf. [8], p. 473, e.g.).

Let

$$J_h := \begin{pmatrix} I - hf_y^0 & -hf_z^0 \\ -hg_y^0 & -hg_z^0 \end{pmatrix},$$

with partial derivatives of f and g at (y_0, z_0). Then the first approximations y_1, z_1 are computed by one linearized Euler step:

$$J_h \begin{pmatrix} y_1 - y_0 \\ z_1 - z_0 \end{pmatrix} = h \begin{pmatrix} f(y_0,z_0) \\ g(y_0,z_0) \end{pmatrix}, \tag{2.1}$$

* Received March 18, 1993. Accepted for publication June 21, 1993. Communicated by V. Mehrmann.

† Johannes Gutenberg–Universität, Fachbereich Mathematik, Saarstraße 21, D–55099 Mainz, Germany, (Schneider@mat.mathematik.Uni-Mainz.DE)

ETNA
Kent State University
etna@mcs.kent.edu

whereas the approximations $y_{i+1}, z_{i+1}, i = 1(1)2m$ are defined by

$$J_h \begin{pmatrix} y_{i+1} - y_i \\ z_{i+1} - z_i \end{pmatrix} = \left(J_h - \begin{pmatrix} 2I & 0 \\ 0 & 0 \end{pmatrix} \right) \begin{pmatrix} y_i - y_{i-1} \\ z_i - z_{i-1} \end{pmatrix} + 2h \begin{pmatrix} f(y_i, z_i) \\ g(y_i, z_i) \end{pmatrix} \tag{2.2}$$

Finally, a smoothing step provides the numerical solution at $x_0 + H (H := 2mh)$:

$$\begin{aligned} y_h(x_0 + H) &:= (y_{2m+1} + y_{2m-1})/2, \\ z_h(x_0 + H) &:= (z_{2m+1} + z_{2m-1})/2. \end{aligned} \tag{2.3}$$

In order to study the errors $y_h(x_0 + H) - y(x_0 + H)$ and $z_h(x_0 + H) - z(x_0 + H)$ of the LIMP rule, we reformulate the method as an ROW method following ideas of [4]. Set $k_0 := 2m(y_1 - y_0), k_j := m(y_{j+1} - y_{j-1}), j = 1(1)2m$. Then the steps (2.1-3) may be collected as one step (from x_0 to $x_0 + H$) of a $(2m + 1)$-stage ROW method which is given by the data $\alpha, A \in \mathbb{R}^{(2m+1)\times(2m+1)}, b \in \mathbb{R}^{2m+1}$. ROW methods are explained for ODE's and DAE's, e.g., in [8]. In our case

$$(\alpha)_{ij} = \frac{1}{m} \begin{cases} 1/2 & , \quad j = 0 \quad \text{and } i \text{ odd}, \\ 1 & , \quad 0 < j < i \quad \text{and } i - j \text{ odd}, \\ 0 & , \quad \text{otherwise}, \end{cases} \tag{2.4}$$

$$A = I/(2m) + S\alpha \text{ with the shift } S := \begin{pmatrix} 0 & & & & & \\ 1 & 0 & & & & \\ 0 & 1 & 0 & & & \\ \vdots & \vdots & \vdots & \vdots & . & \\ 0 & . & . & 0 & 1 & 0 \end{pmatrix}, \tag{2.5}$$

$$b^T = \text{ last row of } A = (1/2, 0, 1, 0, 1, \ldots, 0, 1, 0, 1/2)/m. \tag{2.6}$$

Hence, the method is stiffly accurate (cf. [8] p. 448) and

$$b^T A^{-1} = (0, 0, \ldots 0, 1) \in \mathbb{R}^{2m+1}. \tag{2.7}$$

For the analysis to come, we finally need that

$$(\alpha A^{-1})_{ij} = 2m(\alpha)_{ij}(-1)^{[j/2]}(-1)^{[(i-1)/2]}, \tag{2.8}$$

where [r] denotes the integer part of $r \in \mathbb{R}$ and i, j vary between 0 and $2m$. Now we can apply Roche's analysis of ROW methods (see [10] and[11]) while studying the order of the method and the possibilities to increase the order by some sort of 'extrapolation'.

It is easy to see that $b^T e = 1, e^T := (1, 1, \ldots, 1) \in \mathbb{R}^{2m+1}, b^T A^{-1} e = 1$, and that $b^T Ae = 1/2 + 1/(4m^2) \neq 1/2$. Hence the LIMP rule yields an ROW method which is convergent of order 1 (cf. [3] and [11]). This order can be increased by eliminating the $1/(4m^2)$–error term with extrapolation. In the ODE case, Bader & Deuflhard proved in [1] that the LIMP rule allows an expansion of the error in positive powers of m^{-2} which could be successfully exploited for extrapolation. For stiff equations, however, perturbations are introduced (cf. [6]) making extrapolation less useful. At DAE's (of index 1), finally, m^{-2}-extrapolation achieves order 5 at most (cf. [5] and [8] p. 473), in general. On the other hand, a perturbed asymptotic expansion could be derived for the linearly implicit Euler rule for DAE's (cf. [3]), but the techniques used in that paper failed for the LIMP rule, and the complete expansion of its error remained unknown.

3. Expansion of the Error. The following theorem will show that the error of the LIMP rule allows an expansion in negative powers of m^2 which is perturbed by oscillating terms of the same kind and additional error terms involving non-negative powers of m^2. This explains the limited advantage of the standard extrapolation for DAE's. Furthermore, it shows that the perturbation is quite different from that of the linearly implicit Euler rule.

THEOREM 3.1. *Let $q \in \mathbb{N}$. For sufficiently smooth f and g and regular g_z^0 hold the following expansions of the error introduced by $2m+1$ steps of the LIMP rule with stepsize $h(H = 2mh)$ — or the error introduced by one step of the associated ROW method with stepsize H, respectively:*

$$\begin{aligned} y_h(x_0+H) - y(x_0+H) = H\sum_{i=1}^{q-1} H^i p_{1,i}(m^2)/(m^2)^{[(i+1)/2]} + \\ (-1)^m H^3 \sum_{i=1}^{q-3} H^i p_{2,i+2}(m^2)/(m^2)^{[(i+3)/2]} + \\ (-1)^m H^7 \sum_{i=1}^{q-7} H^i r_{1,i}(m^2) + H^8 \sum_{i=1}^{q-8} H^i r_{2,i}(m^2) + O(H^{q+1}), \end{aligned} \tag{3.1}$$

where $p_{1,i}, p_{2,i}, r_{1,i}, r_{2,i}$ are polynomials of degree [(i-1)/2] with coefficients depending on x_0, f, g but not *on H or m. Likewise,*

$$\begin{aligned} z_h(x_0+H) - z(x_0+H) = H^2\sum_{i=1}^{q-2} H^{i-1} \hat{p}_{1,i}(m^2)/(m^2)^{[(i+1)/2]} + \\ (-1)^m H^3 \sum_{i=1}^{q-3} H^{i-1} \hat{p}_{2,i+1}(m^2)/(m^2)^{[(i+2)/2]} + \\ (-1)^m H^5 \sum_{i=1}^{q-5} H^{i-1} \hat{r}_{1,i}(m^2) + H^8 \sum_{i=1}^{q-8} H^{i-1} \hat{r}_{2,i}(m^2) + \\ O(H^{max\{q,2\}}), \end{aligned} \tag{3.2}$$

where the functions with a ˆ have the same properties as those without one.

We will prove the Theorem 3.1 in the sections to come, but first, some remarks, interpretations, and comments are in order.

REMARKS .

(i) $q = 3$*: One step of Richardson's extrapolation in m^{-2} starting with (y_{h_1},z_{h_1}; $H = 2m_1h_1$) and (y_{h_2},z_{h_2}; $H = 2m_2h_2$) yields an ROW method of order three, i.e., the local error in the approximation of y is $O(H^4)$ and the error for z is $O(H^3)$.*

(ii) $q = 5$*: Two extrapolations lead to order 5 when m_1, m_2, and m_3 are all odd or all even; otherwise order three cannot be improved because of the $(-1)^m$-terms.*

(iii) $q = 7$*: A third extrapolation step as in (ii) will give order 8 in the y–component only. In the z-component there remains $H^5(-1)^m\hat{r}_{1,1}$, which is not affected by extrapolation.*

(iv) $q = 9$*: In general, it is impossible to achieve order 9 for the ROW method because of $r_{2,1}$ and $\hat{r}_{2,1}$ which do not depend on m at all.*

ETNA
Kent State University
etna@mcs.kent.edu

However, the expansion shows how to construct a method of order 8 : Perform three extrapolation steps with even m_j and the same with odd $m_j, j = 1(1)4$. This will eliminate the terms containing m^{-2}, m^{-4}, m^{-6}. Then average the two results in order to eliminate $\hat{r}_{1,1}(m^2)$ and $r_{1,1}(m^2)$, as well. Hence we arrive at order 6 for z and 8 for y, i.e., order 6 for the method. Obviously, this process would be far too expensive for practical applications, but in theory, even m^{-8} could be eliminated by a fourth step in the previous extrapolation. Then the approximation of y is of order 9, but for the z-component, there are still remaining $\hat{r}_{1,3}(m^2)$ and $\hat{r}_{1,4}(m^2)$ which are both terms in m^2 only. They finally could be eliminated in order to obtain the H^8 for z, i.e., an ROW method of order 8. Unfortunately, the elimination of the m^2–terms cannot be achieved by the kind of extrapolation used up to now. Fortunately, formulas for the recursive elimination of unconventional terms are given in [13]. They can be used to remove the m^2–terms. We do not go into the details of this process, but we want to emphasize that even standard extrapolation with $m_{j-k}, \cdots, m_j$ amplifies the m^2-terms in the extrapolated result by $\sum_{i=j-k}^{j} m_i^2$ (interpolate $1/x^2$ at m_i^{-2} with a polynomial in x^2 and evaluate at 0). This clearly shows that the m_i used for extrapolation should be as small as possible if we want to avoid an unpleasant growth of the error constants. Furthermore, extrapolating more than four steps would eliminate some error terms but increase others at the same time. Therefore, higher accuracy requires a smaller H, i.e., more ROW steps and more updates of the Jacobian which makes sense because we are trying to solve nonlinear equations $g(y, z) = 0$, too. Before proving the Theorem, some final remarks.

Remarks .

(i) In the trivial case of linear constraints g in (1.1), the unperturbed expansion of the ODE case is still valid.

(ii) If g happens to be linear in z only, then the unpleasant non-negative powers of m^2 do not appear at all, but the oscillating terms in m^{-2} are still present. Hence, extrapolation can be successfully performed with even (odd) m_j only. Otherwise the order is bounded by three again.

(iii) The ROW method of order 8 needs quite a lot of work and is probably of theoretical interest only.

The assertions of the first two remarks may be deduced from the results in Section 6. There it is shown which trees lead to the unpleasant terms in the expansions (3.1) and (3.2), and it is easy to see that the elementary differentials of these critical trees vanish in the given cases. Now the proofs will be derived in three steps progress from nice trees to a tedious counting procedure.

4. How the Trees are Growing. The order conditions of Runge–Kutta and ROW methods may be derived from trees (cf. [2], [7], [8]). There are trees with meagre roots leading to conditions for order k which have the form $b^T x^{(k)} = \sigma, \sigma = \sigma(x^{(k)})$. We will call them *y-conditions* and $x^{(k)}$ an *order-vector for y and order k*. Additional conditions for order k of the z-component are derived from trees with fat roots. They are given by equations like $b^T A^{-1} u^{(k)} = \lambda, \lambda = \lambda(u^{(k)})$ (cf. [10] and [11]). In our case $b^T A^{-1} u^{(k)} = (u^{(k)})_{2m}$. Consequently, we will call them *z-conditions* and $u^{(k)}$ an *order–vector for z and order k*. If $k = 1$ then there exists only one order-vector for $y : x^{(1)} = (1, 1, ..., 1)^T =: e; \sigma(x^{(1)}) = 1$. The z–conditions start with $k = 2$ and one order–vector for $z : u^{(2)} = (\alpha e)^2$ (componentwise squaring); $\lambda(u^{(2)}) = 1$.

It is easy to construct y–conditions of order $q > 1$ (cf. [13] and [9] for a detailed proof):

$$
\text{(4.1)}\qquad
\begin{array}{l}
\text{1) If } b^T x^{(q-1)} = \sigma \text{ is a } y\text{–condition for order } q-1, \text{ then} \\
\quad b^T A x^{(q-1)} = \sigma/q \text{ is a } y\text{–condition for order } q. \\
\text{2) If } b^T A^{-1} u^{(q-1)} = \lambda \text{ is a } z\text{–condition for order } q-1, \text{ then} \\
\quad b^T u^{(q-1)} = \lambda/q \text{ is a } y\text{–condition for order } q.
\end{array}
$$

Hence, $Ax^{(q-1)}$ and $u^{(q-1)}$ are order–vectors for y and order $q; \sigma(Ax^{(q-1)}) = \sigma(x^{(q-1)})/q, \sigma(u^{(q-1)}) = \lambda(u^{(q-1)})/q$.

The real problem are the z–conditions. For order $q > 1$ they are constructed by combining previous order–vectors in the following way (cf. [8], [10], [11]):

$$
\text{(4.2)}\qquad
\begin{array}{l}
\text{If } k_1, k_2 \geq 0, k_1 + k_2 > 1, k_1 + k_2 \leq q \text{ and} \\
\displaystyle\sum_{i=1}^{k_1} \nu_i + \sum_{j=1}^{k_2} \mu_j = q \text{ with } 1 \leq \nu_1 \leq \nu_2 \leq \ldots \leq \nu_{k_1}, \\
\qquad\qquad\qquad\qquad 2 \leq \mu_1 \leq \mu_2 \leq \ldots \leq \mu_{k_2},
\end{array}
$$

and if $b^T x^{(\nu_i)} = \sigma_{\nu_i}$ are y–conditions for order $\nu_i, b^T A^{-1} u^{(\mu_j)} = \lambda_{\mu_j}$ are z–conditions for order μ_j, then

$$
\begin{array}{rl}
u^{(q)} := [(\alpha x^{(\nu_1)}) * (\alpha x^{(\nu_2)}) * & \ldots \quad *(\alpha x^{(\nu_{k_1})}) * \\
(\alpha A^{-1} u^{(\mu_1)}) * & \ldots \quad *(\alpha A^{-1} u^{(\mu_{k_2})})]
\end{array}
$$

is an order–vector for z and order q ('$*$' means componentwise multiplication), and

$$
b^T A^{-1} u^{(q)} = \prod_{i=1}^{k_1} \sigma_{\nu_i} \prod_{j=1}^{k_2} \lambda_{\mu_j} =: \lambda(u^{(q)})
$$

yields a z–condition for order q.

In the sequel, we have to show that the 'extrapolated' LIMP rule actually achieves order 8, i.e., it satisfies 1731 order conditions constructed by the rules just described. Lucky enough, we are only interested in the structure of these conditions for the LIMP rule with respect to m and not in the details. Obviously, this structure is influenced by the following operations: Multiplication of an order–vector by $\alpha, \alpha A^{-1}, A, b^T$, or $b^T A^{-1}$, and by componentwise multiplication of order-vectors. We will show that order–vectors which are produced according to (4.1-2) by α or A of the LIMP rule (given in (2.4-5)), do not fulfill the associated order conditions exactly, in general, but the error allows an expansion like (3.1) or (3.2) with $H := 1$ and q depending on the actual order. Hence, this error can be eliminated — within the asserted limits — by extrapolation and other linear combinations of approximations computed by the LIMP rule with various stepsizes.

5. Structure of the Order–Vectors. In the first step we will see that an oscillating term is introduced by αA^{-1}. Remember that $u_j^{(2)} = (\alpha e)_j^2 = j^2/(4m^2)$ and hence $\alpha A^{-1} u^{(2)}$ is one part of some order–vectors for z and order $\geq 3, y$ and order ≥ 4, respectively.

LEMMA 5.1. *Let $r \in \mathbb{N}$ and $v_j := j^r, j = 0(1)2m$. Then*

$$
(\alpha A^{-1} v)_i = \begin{cases} 0 & , \quad i = 0, \\ p_{r/2}(i^2) - C(-1)^{i/2} & , \quad i \text{ even and } i > 0, \\ q_{r/2}(i^2) & , \quad i \text{ odd }, \end{cases}
$$

ETNA
Kent State University
etna@mcs.kent.edu

for r even, and

$$(\alpha A^{-1}v)_i = \begin{cases} i \cdot p_{(r-1)/2}(i^2) & , \quad i \text{ even}, \\ i \cdot q_{(r-1)/2}(i^2) + D(-1)^{(i+1)/2} & , \quad i \text{ odd}, \end{cases}$$

for r odd, $i = 0(1)2m$. Here p and q are polynomials of degree $r/2$ or $(r-1)/2$, respectively, with leading coefficient 1.

Proof. Using (2.4) and (2.8) we get

$$(\alpha A^{-1}v)_i = 2 \begin{cases} \sum_{\mu=1}^{i/4} [(4\mu-1)^r - (4\mu-3)^r] & , \quad i \equiv 0 \bmod 4, \\ \sum_{\mu=1}^{(i-1)/4} [(4\mu)^r - (4\mu-2)^r] & , \quad i \equiv 1 \bmod 4, \\ 1 + \sum_{\mu=1}^{(i-2)/4} [(4\mu+1)^r - (4\mu-1)^r] & , \quad i \equiv 2 \bmod 4, \\ \sum_{\mu=0}^{(i-3)/4} [(4\mu+2)^r - (4\mu)^r] & , \quad i \equiv 3 \bmod 4. \end{cases}$$

In the brackets there are expressions of the form $(a+1)^r - (a-1)^r$ which have to be expanded in powers of a. Summing up these terms with the Euler–MacLaurin summation formula for the mid–point rule (as quadrature rule !) (cf. [14] p.135) yields the desired structure of $(\alpha A^{-1}v)_i$. Furthermore, it can be seen that C and D are non–vanishing constants depending on r and the first $[(r-1)/2]$ Bernoulli–numbers. □

The next lemma shows how the oscillation is amplified by αA^{-1} . For $u^{(4)} := (\alpha e) * \alpha A^{-1}[(\alpha e) * \alpha A^{-1}(\alpha e)^2]$ this amplification appears at first.

LEMMA 5.2. *Let $r \in \mathbb{N}$ and $v_j := j^r(-1)^{j/2}(-C)$, j even, $v_j := j^r(-1)^{(j+1)/2}D$, j odd, $j = 0(1)2m$. Then*

$$(\alpha A^{-1}v)_i = i \cdot p_{r/2}(i^2) \cdot \begin{cases} D(-1)^{i/2} & , \quad i \quad \text{even}, \\ C(-1)^{(i+1)/2} & , \quad i \quad \text{odd}, \end{cases}$$

for r even, and

$$(\alpha A^{-1}v)_i = p_{(r+1)/2}(i^2) \cdot \begin{cases} D(-1)^{i/2} & , \quad i \text{ even}, \\ C(-1)^{(i+1)/2} & , \quad i \text{ odd}, \end{cases}$$

for r odd, $i = 0(1)2m$. The leading coefficient of p is $1/(r+1)$.

Proof. Now

$$(\alpha A^{-1}v)_i = 2 \begin{cases} C(-1)^{(i+1)/2} & \sum_{\mu=1}^{(i-1)/2} (2\mu)^r & , \quad i \quad \text{odd}, \\ D(-1)^{i/2} & \sum_{\mu=1}^{i/2} (2\mu-1)^r & , \quad i \quad \text{even}. \end{cases}$$

and — no surprise — the mid–point rule again proves the lemma. □

Combining the two lemmas we obtain the impact of αA^{-1} on an expression which is already perturbed by an oscillation.

LEMMA 5.3. *Let $r \in \mathbb{N}, k, l, n \in \mathbb{N}_0$ and*

$$\hat{v}_j := \begin{cases} p_l(j^2) + (-1)^{j/2}\rho_n(j^2) & , j \text{ even}, \\ q_l(j^2) + (-1)^{(j+1)/2}\tau_k(j^2)\cdot j & , j \text{ odd}, \end{cases} \tag{5.1}$$

$$\hat{w}_j := \begin{cases} j\cdot p_l(j^2) + (-1)^{j/2}\rho_n(j^2)\cdot j & , j \text{ even}, \\ j\cdot q_l(j^2) + (-1)^{(j+1)/2}\tau_k(j^2) & , j \text{ odd}, \end{cases} \tag{5.2}$$

with polynomials p, q, ρ, τ of degrees given by their respective subscripts. Furthermore, define $v_j := j^r\hat{v}_j, w_j := j^r\hat{w}_j, j = 0(1)2m$. Then

$$(\alpha A^{-1}v)_i = \begin{cases} q_{l+r/2}(i^2) + (-1)^{i/2}\tau_{k+1+r/2}(i^2) & , i \text{ even}, \\ p_{l+r/2}(i^2) + (-1)^{(i+1)/2}\rho_{n+r/2}(i^2)\cdot i & , i \text{ odd}, \end{cases}$$

$$(\alpha A^{-1}w)_i = \begin{cases} i\cdot q_{l+r/2}(i^2) + (-1)^{i/2}\tau_{k+r/2}(i^2)\cdot i & , i \text{ even}, \\ i\cdot p_{l+r/2}(i^2) + (-1)^{(i+1)/2}\rho_{n+1+r/2}(i^2) & , i \text{ odd}, \end{cases}$$

for even r, and

$$(\alpha A^{-1}v)_i = \begin{cases} i\cdot q_{l+(r-1)/2}(i^2) + (-1)^{i/2}\tau_{k+1+(r-1)/2}(i^2)\cdot i & , i \text{ even}, \\ i\cdot p_{l+(r-1)/2}(i^2) + (-1)^{(i+1)/2}\rho_{n+(r+1)/2}(i^2) & , i \text{ odd}, \end{cases}$$

$$(\alpha A^{-1}w)_i = \begin{cases} q_{l+(r+1)/2}(i^2) + (-1)^{i/2}\tau_{k+(r+1)/2}(i^2) & , i \text{ even}, \\ p_{l+(r+1)/2}(i^2) + (-1)^{(i+1)/2}\rho_{n+(r+1)/2}(i^2)\cdot i & , i \text{ odd}, \end{cases}$$

for r odd, $i = 0(1)2m$. The leading coefficients of p_l and q_l become the leading coefficients of the new $p_\bullet$ and $q_\bullet$.

It is interesting to notice that the pairs $(q, \tau)[(p, \rho)]$ move from the odd [even] positions in v and w to the even [odd] ones after the operations described in the Lemma 5.3, but there is a different increase of the degrees: The highest exponent of the oscillations is *increased by* $r+1$ whereas there is an *increase of only* r in the regular terms, and this different kind of increase finally leads to the non-negative powers of m^2! Anyway, the structure of $\hat{v}[\hat{w}]$ (i.e., even [odd] exponents at even indices, even [odd] exponents of the regular term and odd [even] ones in the oscillations at odd indices) is maintained for even r and present again for odd r after the application of αA^{-1} to $w[v]$.

In the next step we shall prove that multiplication by α or A similarly preserves the structure of $\hat{v}$ and $\hat{w}$. Actually, there is nothing to prove for αv and αw because of the connection between α and αA^{-1} given in (2.8) and the previous Lemmas.

LEMMA 5.4. *Consider $\hat{v}, \hat{w}$ from (5.1),(5.2) but admit $\hat{w}_0$ to have an arbitrary value. Then*

$$(A\hat{v})_i = \frac{1}{2m}\begin{cases} \hat{v}_0 & , i = 0, \\ i\cdot\hat{p}_l(i^2) + (-1)^{i/2}\hat{\rho}_{n-1}(i^2)\cdot i & , i \text{ even and } i > 0, \\ i\cdot\hat{q}_l(i^2) + (-1)^{(i+1)/2}\hat{\tau}_k(i^2) & , i \text{ odd}, \end{cases}$$

ETNA
Kent State University
etna@mcs.kent.edu

$$(A\hat{w})_i = \frac{1}{2m}\begin{cases} \hat{w}_0 & , \quad i=0, \\ \hat{p}_{l+1}(i^2) + (-1)^{i/2}\hat{\rho}_n(i^2) & , \quad i \text{ even and } i>0, \\ \hat{q}_{l+1}(i^2) + (-1)^{(i+1)/2}\hat{\tau}_{k-1}(i^2)\cdot i & , \quad i \text{ odd,} \end{cases}$$

Hence the structure of $\hat{v}, \hat{w}$ is simply swapped. The highest exponent of the oscillations is decreased by one *whereas it is increased by one at the regular term. The leading coefficients of $\hat{p}_\bullet, \hat{q}_\bullet$ are those of $p_\bullet, q_\bullet$ but divided by the highest exponent of $\hat{p}_\bullet, \hat{q}_\bullet$ (as in integration of polynomials).*

Proof. Multiplying a row of A with a vector may be considered as applying the trapezoidal rule! After sorting the oscillating terms according to their sign it is now useful to interpret the multiplication by A as the difference between a trapezoidal approximation and one obtained by the mid–point rule. Hence, integration of polynomials, Euler–MacLaurin, and Bernoulli prove the assertion.□

Remark . *The dot product of b with $\hat{v}$ or $\hat{w}$ also maintains their structure, because b^T = last row of A — see (2.6).*

In the ODE case we could finish this section now, because there are no fat nodes, i.e., no αA^{-1} and therefore no oscillations. In the DAE case, however, it remains to study the componentwise multiplication of vectors containing oscillations like $\hat{v}$ and $\hat{w}$. A short calculation shows that this operation too preserves the structure described in Lemma 5.3, but unfortunately, the oscillating terms may lose their changing signs as a result of this operation and, hence, become part of the regular terms. If the degree of these former oscillations is the same as or higher than the degree of the former regular terms, then the sign–independent non–negative exponents of m appear in the Theorem 3.1. This concludes Section 5. The final step of the proof will be given in Section 6 where we will compute the exponents of H at which the different sums start.

6. Counting.

The rules (4.1), (4.2) and the results of the previous Section prove inductively the structure of order–vectors for the LIMP rule. By separating regular terms from terms which are still oscillating or which had their origin in oscillations, we obtain the following result:

$$v_i := \frac{\varphi}{(2m)^k}\begin{cases} p_{k/2}(i^2) + \hat{p}_{l_1}(i^2) + (-1)^{i/2}\rho_{n_1}(i^2) & , \quad i \text{ even,} \\ q_{k/2}(i^2) + \hat{q}_{l_2}(i^2) + (-1)^{(i+1)/2}\tau_{n_2}(i^2)\cdot i & , \quad i \text{ odd,} \end{cases}$$

for k even, and

$$v_i := \frac{\varphi}{(2m)^k}\begin{cases} const. & , \quad i=0 \\ i\cdot p_{(k-1)/2}(i^2) + i\cdot\hat{p}_{l_1}(i^2) + (-1)^{i/2}\rho_{n_1}(i^2)\cdot i & , \quad i \text{ even and } i>0, \\ i\cdot q_{(k-1)/2}(i^2) + i\cdot\hat{q}_{l_2}(i^2) + (-1)^{(i+1)/2}\tau_{n_2}(i^2) & , \quad i \text{ odd,} \end{cases}$$

for k odd, $i = 0(1)2m$. The leading coefficient of the polynomials $p_\bullet, q_\bullet$ is 1.

An order–vector u for z and order k has the structure given by v with φ replaced by $\lambda(u) = \sigma(u)\cdot(k+1)$. An order-vector x for y and order $k+1$ has the structure given by v with φ replaced by $\sigma(x)\cdot(k+1)$.

Hence, $b^T A^{-1} u = u_{2m} = \lambda(u) + \sum_{\mu=1}^{k/2} \beta_\mu m^{-2\mu} + \sum_{\mu=0}^{l_1} \gamma_\mu m^{2(\mu-k/2)} + (-1)^m \sum_{\mu=0}^{n_1} \delta_\mu m^{2(\mu-k/2)}$ holds for an order-vector u for z and order k, k even, and Lemma 5.4 (with the last row of A, i.e., with b^T) yields for the same case

$$\begin{aligned} b^T u &= \lambda(u)/(k+1) + \sum_{\mu=1}^{k/2} \hat{\beta}_\mu m^{-2\mu} + \sum_{\mu=0}^{l_1} \hat{\gamma}_\mu m^{2(\mu-k/2)} \\ &\quad + (-1)^m \sum_{\mu=0}^{n_1-1} \hat{\delta}_\mu m^{2(\mu-k/2)}, \\ \sigma(u) &= \lambda(u)/(k+1). \end{aligned}$$

Therefore, the respective order conditions do not hold, in general, but they allow an expansion in m^2, and this property does not depend on the order k — as the lemmas showed. However, we still must discuss the degrees n_1, n_2, and l_1, l_2. Obviously, they depend on the 'history' of the order–vector — that means they depend on the number of operations (multiplying by $\alpha A^{-1}, A$, or componentwise) which were performed while constructing them recursively. It suffices to count this influence in the worst case in order to recognize the first appearance of an exponent greater or equal k. From the lemmas we know that multiplication by αA^{-1} introduces the oscillation and increases its influence. Therefore, the sequence $u^{(j)}$ of order–vectors for z and order j defined by

$$u^{(2)} := (\alpha e)^2, \quad u^{(j)} := (\alpha e) * \alpha A^{-1} u^{(j-1)}, j > 2,$$

deliver the fastest amplification of the oscillating terms. At $j = 3$ we have 3 as highest exponent in the regular terms, 1 in the evenly indexed oscillations, and no oscillation at all at odd indices. Therefore, at $j = 5$, the oscillation also has exponent 5. Hence, we have proved the H^3 and H^5 factors up to the oscillations in (3.2). In order to obtain the H^4 and H^8 at these places in (3.1) we have to study $b^T u^{(j)}$, an order condition for y and order $j + 1$. However, the dot product with b decreases the degree of the oscillation. Nevertheless, the first oscillation occurs in $b^T u^{(3)}$ and it lasts through $b^T u^{(7)}$ where it reaches the exponent of the regular term. Clearly, $\alpha A^{-1} u^{(4)}$ (which is *not* an order–vector) has 4 as highest exponent in the regular term *and* in the oscillations because $(\alpha e)_i = i/(2m)$. Therefore $(\alpha A^{-1} u^{(4)})^2$, which *is* an order–vector for z and order 8, is the first case of the difficulties introduced by componentwise multiplication of perturbed terms. This fact explains the H^8 in front of the last sum of (3.2). Then $b^T (\alpha A^{-1} u^{(4)})^2$ cannot do better and still maintain the no–longer–oscillating perturbations, leading to the H^9 which is present in all terms of the last sum in (3.1). Thus, we have finally proved all statements of the Theorem 3.1.

A first application of the results could be replacing every asterisk by a 7 in Table 4.5 of [8, p.473].

Acknowledgement. It is a pleasure to thank M. Landvogt (Mainz University) who helped me with first important steps of this work by computing nearly 30 order–vectors as functions of m. Some of the calculations were easy but most required a number of iterations until our results matched. In addition, the referee's report helped me to improve the readability of the paper.

REFERENCES

[1] G. Bader and P. Deuflhard, *A semi–implicit mid–point rule for stiff systems of ordinary differential equations*, Numer. Math., 41 (1983), pp. 373–398.

[2] J.C. Butcher, *The numerical analysis of ordinary differential equations. Runge–Kutta and general linear methods*, John Wiley & Sons, Brisbane, 1987.

[3] P. Deuflhard, E. Hairer and J. Zugck, *One–step and extrapolation methods for differential–algebraic systems*, Numer. Math., 51 (1987), pp. 501–516.

[4] E. Hairer, G. Bader and Ch. Lubich, *On the stability of semi–implicit methods for ordinary differential equations*, BIT, 22 (1982), pp. 211–232.

[5] E. Hairer and Ch. Lubich, *On extrapolation methods for stiff and differential–algebraic equations*, Teubner Texte zur Mathematik, No. 104, Teubner, Leipzig, 1988, pp. 64–73.

[6] E. Hairer and Ch. Lubich, *Extrapolation at stiff differential equations*, Numer. Math., 52 (1988), pp. 377–400.

[7] E. Hairer, S.P. Nørsett and G. Wanner, *Solving ordinary differential equations I*, Springer, Heidelberg, 1987.

[8] E. Hairer and G. Wanner, *Solving ordinary differential equations II*, Springer, Berlin, 1991.

[9] W. Link, *Verallgemeinerte Rosenbrock–Wanner–Verfahren für Anfangswertaufgaben bei partitionierten differential–algebraischen Systemen*, Doctoral Thesis, Mainz, 1991.

[10] M. Roche, *Runge–Kutta and Rosenbrock methods for differential–algebraic equations and stiff ODE's*, Doctoral Thesis, Genève, 1988.

[11] M. Roche, *Rosenbrock methods for differential–algebraic equations*, Numer. Math., 52 (1988), pp. 45–63.

[12] C. Schneider, *Vereinfachte Rekursionen zur Richardson–Extrapolation in Spezialfällen*, Numer. Math., 24 (1975), pp. 177–184.

[13] C. Schneider, *Rosenbrock–type methods adapted to differential–algebraic systems*, Math. Comp., 56 (1991), pp. 201–213.

[14] J.F. Steffensen, *Interpolation*, 2nd ed., Chelsea Publishing Company, New York, 1950.