ELSEVIER

Journal of Computational and Applied Mathematics 132 (2001) 107–125

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.nl/locate/cam

High-order convergent deferred correction schemes based on parameterized Runge–Kutta–Nyström methods for second-order boundary value problems

T. Van Hecke [1], M. Van Daele *

Vakgroep Toegepaste Wiskunde en Informatica, Universiteit Gent, Krijgslaan 281-S9, B9000 Gent, Belgium

Received 19 January 2000; received in revised form 12 May 2000

Abstract

Iterated deferred correction is a widely used approach to the numerical solution of first-order systems of nonlinear two-point boundary value problems. Normally, the orders of accuracy of the various methods used in a deferred correction scheme differ by 2 and, as a direct result, each time deferred correction is used the order of the overall scheme is increased by a maximum of 2. In [16], however, it has been shown that there exist schemes based on parameterized Runge–Kutta methods, which allow a higher increase of the overall order. A first example of such a high-order convergent scheme which allows an increase of 4 orders per deferred correction was based on two mono-implicit Runge–Kutta methods. In the present paper, we will investigate the possibility for high-order convergence of schemes for the numerical solution of second-order nonlinear two-point boundary value problems not containing the first derivative. Two examples of such high-order convergent schemes, based on parameterized Runge–Kutta-Nyström methods of orders 4 and 8, are analysed and discussed.

Keywords: Boundary value problem; Deferred correction; Runge–Kutta–Nyström

1. Introduction

In the present paper we will be concerned with the numerical solution of second-order two-point boundary value problems (BVPs) of the form

$$y'' = f(x, y), \quad a \leqslant x \leqslant b, \quad g(y(a), y'(a), y(b), y'(b)) = 0 \tag{1.1}$$

with $y \in \mathbb{R}^d$, $f : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{2d}$. The specific form of the boundary conditions will not be central in this paper. Although it is straightforward to convert a second-order

* Corresponding author. Tel.: +32-9-264-4809; fax: +32-9-264-4995.
E-mail address: marnix.vandaele@rug.ac.be (M.V. Daele).
[1] Present address: Hogeschool Gent, Schoonmeersstraat 52, B9000 Gent, Belgium.

0377-0427/01/$ - see front matter
PII: S0377-0427(00)00603-8

system to a first-order one, it is well-known that several advantages (including substantial gains in efficiency, lower storage requirements, etc.) are realized when the equations can be treated in their original second-order form. A popular class of numerical methods which allows this is the class of Runge–Kutta–Nyström (RKN) methods. In this paper we will use these methods in a deferred correction (DC) framework.

Iterated DC is a widely used technique for the solution of first-order BVPs. For second-order BVPs, Daniel and Martin [10] describe Numerov's method in combination with the deferred correction technique. As so far the technique has not been applied on RKN methods for second-order BVPs. A single step of a typical DC scheme based on implicit Runge–Kutta(–Nyström) formulae can be defined as follows:

Let ϕ, ϕ^* be two Runge–Kutta(–Nyström) formulae of order p and p^* respectively where $p < p^*$. Consider the algorithm defined by

$$\phi(\eta) = 0,$$

$$\phi(\bar{\eta}) = -\phi^*(\eta). \tag{1.2}$$

Then, providing that ϕ and ϕ^* have certain special properties, the DC scheme defined by (1.2) is of order $\min(p^*, 2p)$. Algorithms of this type based on parameterized implicit Runge–Kutta methods have been derived in [5–7] (based on mono-implicit methods) and in [2,9] (based on Lobatto methods). Two codes TWPBVP and ACDC which implement DC schemes are available from NETLIB.

In what follows we will consider the rather more general DC scheme

$$\phi(\eta) = 0,$$

$$\phi(\bar{\eta}) = \psi(\eta). \tag{1.3}$$

A general framework for proving accuracy results for DC schemes of the form (1.3) was given in an influential paper by Skeel [13]. We present his main theorem.

Consider the approximate numerical solution of (1.1) on a mesh

$$\pi : a = x_0 < x_1 < \cdots < x_{N+1} = b.$$

Denote by Δz the restriction of the continuous solution $z(x)$ of (1.1) to the finite grid π. Then we have the following theorem:

Theorem 1. *Let ϕ be a stable numerical method and assume that the following conditions hold for the DC scheme* (1.3):

(i) $\|\eta - \Delta z\| = \mathrm{O}(h^p)$,
(ii) $\|\psi(\Delta z) - \phi(\Delta z)\| = \mathrm{O}(h^{r+p})$,
(iii) $\psi(\Delta u) = \mathrm{O}(h^r)$

for arbitrary functions u having at least r continuous derivatives. Here $\|.\|$ is a suitable finite norm defined in [13] *and h is the maximum grid spacing.*

If $\phi(\bar{\eta}) = \psi(\eta)$ then

$$\|\bar{\eta} - \Delta z\| = \mathrm{O}(h^{\min(2p, r+p)}).$$

The feature that is common to most of the DC schemes that have been derived so far is that $r = 2$ and for these schemes the order of accuracy is increased by 2 for each application of the DC. In [2] a sufficient condition to achieve this increase in accuracy was given and this was basically that the Runge–Kutta formulae ϕ and ϕ^* should be symmetric and that they should be written in a special way that is appropriate for BVPs. This condition is of course straightforward to satisfy. The main reason why it is hard to get more than two orders of accuracy improvement per iteration is the difficulty in satisfying condition (iii) for $r > 2$. In [16], the question was addressed of whether it is possible for first-order systems to choose ϕ and ϕ^* such that high-order convergent schemes emerge, i.e., schemes for which $r > 2$ can be achieved. The necessary and sufficient conditions to obtain a certain value for r were established and a pair of MIRK formulae of orders 4 and 8 was constructed to show that it is possible to achieve this high-order convergence and that the resulting DC schemes have potential advantages over the other DC schemes that have been derived so far.

In this paper on second-order systems, we will again focus on this question of high-order convergence. We will analyse the necessary and sufficient conditions to obtain a certain value for r and we will construct an example to illustrate our results. Further, we will also discuss linear stability properties. Before performing this analysis we first briefly recall the concepts of parameterized (mono-)implicit RKN methods.

2. Parameterized (mono-)implicit Runge–Kutta–Nyström methods

For the numerical solution of second-order IVPs we presented the following representation of s-stage implicit RKN methods, known as *parameterized implicit* RKN (PIRKN) methods [14,17,18]:

$$
\begin{aligned}
y_{n+1} &= y_n + h_n y'_n + h_n^2 \sum_{i=1}^{s} \bar{b}_i f(x_n + c_i h_n, Y_{n,i}),\\
y'_{n+1} &= y'_n + h_n \sum_{i=1}^{s} b_i f(x_n + c_i h_n, Y_{n,i})\\
Y_{n,i} &= (1 - v_i) y_n + v_i y_{n+1} + (c_i - v_i - w_i) h_n y'_n + w_i h_n y'_{n+1}\\
&\quad + h_n^2 \sum_{j=1}^{s} x_{ij} f(x_n + c_j h_n, Y_{n,j}),
\end{aligned}
$$

for $i = 1, \ldots, s$. Hence, a s-stage PIRKN method is completely determined by the tableau

$$
\begin{array}{c|c|c|ccccc}
c_1 & v_1 & w_1 & x_{11} & x_{12} & \ldots & x_{1s}\\
c_2 & v_2 & w_2 & x_{21} & x_{22} & \ldots & x_{2s}\\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\
c_s & v_s & w_s & x_{s1} & x_{s2} & \ldots & x_{ss}\\
\hline
 & & & \bar{b}_1 & \bar{b}_2 & \ldots & \bar{b}_s\\
 & & & b_1 & b_2 & \ldots & b_s
\end{array}
\tag{2.1}
$$

Comparing the representation (2.1) with the description of a general IRKN method by means of its Butcher tableau $(c, A, \bar{b}, b)$ [4,11], it is easy to verify that the relationship $A = X + v \cdot \bar{b}^{\mathrm{T}} + w \cdot b^{\mathrm{T}}$ holds. Thus, there exist many PIRKN methods which correspond to a given RKN method $(c, A, \bar{b}, b)$. However, given v and w, there is only one corresponding PIRKN method, i.e. the one for which $X = A - v \cdot \bar{b}^{\mathrm{T}} - w \cdot b^{\mathrm{T}}$. Therefore, we call v and w the parameters of the PIRKN method.

In general, the use of a s-stage PIRKN method for the solution of (1.1) requires the solution of systems of equations of dimension $\tilde{N} \times \tilde{N}$, where $\tilde{N} \approx (2+s)(N+1)d$ (its exact value depends upon the specific form of the boundary conditions). However, a major reduction of the computational cost is possible when X is lower triangular. In that case the s internal stages $Y_{n,i}$ can be expressed in terms of y_n and y'_n and we only need to solve numerically systems of algebraic equations of dimension $\tilde{N} \times \tilde{N}$ where $\tilde{N} \approx 2d(N+1)$. In analogy with the case of RK methods, we call RKN methods for which X is lower triangular mono-implicit (MIRKN) methods.

We further restrict ourselves to symmetric methods, a natural choice in the context of BVPs. The symmetry of ϕ and ϕ^* then also guarantees the symmetry of the DC scheme (1.2). The conditions of symmetry for a RKN method (2.1) are well known and can be expressed as follows: if x_n and x_{n+1}, y_n and y_{n+1} as well as y'_n and y'_{n+1} are swapped, and h is replaced by $-h$, then the original method is obtained. The IRKN tableau (2.1) is a particularly convenient representation for the investigation of symmetry. The condition for symmetry comes down to the requirement that there must exist a permutation σ of the stages such that for each stage $Y_{n,i}$ there exists a stage $Y_{n,\sigma(i)}$ for which

$$\bar{b}_{\sigma(i)} = b_i - \bar{b}_i, b_{\sigma(i)} = b_i, \quad c_{\sigma(i)} = 1 - c_i, \quad v_{\sigma(i)} = 1 - v_i,$$

$$w_{\sigma(i)} = v_i + w_i - c_i, \quad x_{\sigma(i)\sigma(j)} = x_{ij}. \tag{2.2}$$

We further recall [11] that a RKN method is of order q if for the local initial value problem with sufficiently smooth solution

$$y''(x) = f(x, y), \quad y(x_n) = y_n, \quad y'(x_n) = y'_n,$$

the numerical solution obeys

$$y(x_{n+1}) - y_{n+1} = \mathrm{O}(h_n^{q+1}),$$

$$y'(x_{n+1}) - y'_{n+1} = \mathrm{O}(h_n^{q+1}).$$

To allow us to have order q, the so-called order conditions [11] have to be fulfilled by the coefficients in the Butcher tableau $(c, A, \bar{b}, b)$. For PIRKN methods, these order conditions can be reexpressed in terms of the parameterized tableau $(c, v, w, X, \bar{b}, b)$. For given values of the parameters v and w the original set of order conditions (expressed in terms of A) is equivalent to another set of equations (expressed in terms of X).

The number of equations to be solved can however be reduced by imposing the stage order conditions:

$$C(\eta): X.c^q = \frac{c^{q+2} - v}{(q+2)(q+1)} - \frac{w}{q+1}, \quad q = 0, 1, 2, \ldots, \eta \tag{2.3}$$

with $c^q = (c_1^q, \ldots, c_s^q)^{\mathrm{T}}$. This condition ensures that the order of the internal stages is at least $\eta + 2$, i.e., $Y_{n,i} = y(x_n + c_i h_n) + \mathrm{O}(h_n^{\eta+3})$, $i = 1, 2, \ldots, s$.

3. DC with PIRKN methods

In this section we will examine the difficulties in satisfying condition (iii) of Theorem 1. The results are obtained by (brute force) Taylor series expansion as in [15,16]. Although we feel a little disappointed that we have not been able to obtain our results by using a more elegant approach (e.g. using rooted trees) we are happy that the results we obtain can be presented in a way which is easy to understand.

Suppose we want to approximate the solution of the BVP (1.1) on the mesh π and let $h = \max_i h_i$ where $h_i := x_{i+1} - x_i$. Let Δy and $\Delta y'$ be the restriction to the grid π of the continuous functions $y(x)$ and $y'(x)$ respectively where $y(x)$ is the solution of (1.1) and let η and η^* be approximations to $\Delta z = (\Delta y, \Delta y')^{\mathrm{T}}$. Further, we denote $u = (\theta, \kappa)^{\mathrm{T}}$ with θ and κ arbitrary functions defined over the integration interval.

As already mentioned, in our case ϕ will correspond to a PIRKN method of order p and $\psi := \phi - \phi^*$ where ϕ^* corresponds to a PIRKN method of order $p^* > p$ (we will systematically denote the quantities that relate to ϕ^* with a *-superscript : s^*, a_{ij}^*, $\bar{b}_i^*$, b_i^*, c_i^*, or s^*, x_{ij}^*, $\bar{b}_i^*$, b_i^*, c_i^*, v_i^*, w_i^* in parameterized form).

For ϕ we have

$$\phi(\Delta u)_n := \begin{pmatrix} \frac{\theta_{n+1}-\theta_n}{h_n} - \kappa_n - h_n \sum_{i=1}^{s} \bar{b}_i f(x_n + c_i h_n,\, Y_{n,i}) \\ \frac{\kappa_{n+1}-\kappa_n}{h_n} - \sum_{i=1}^{s} b_i f(x_n + c_i h_n,\, Y_{n,i}) \end{pmatrix} \tag{3.1}$$

with

$$\begin{aligned} Y_{n,i} &= (1 - v_i)\,\theta_n + v_i \theta_{n+1} + h_n (c_i - v_i - w_i)\kappa_n + h_n w_i \kappa_{n+1} \\ &\quad + h_n^2 \sum_{j=1}^{s} x_{ij} f(x_n + c_j h_n, Y_j) \\ &= \theta_n + h_n v_i \theta_n' + \frac{h_n^2}{2} v_i \theta_n'' + \frac{h_n^3}{3!} v_i \theta_n^{(3)} + \cdots + \frac{h_n^q}{q!} v_i \theta_n^{(q)} + \cdots \\ &\quad + h_n (c_i - v_i)\kappa_n + h_n^2 w_i \kappa_n' + \frac{h_n^3}{2} w_i \kappa_n'' + \cdots + \frac{h_n^q}{(q-1)!} w_i \kappa_n^{(q-1)} + \cdots \\ &\quad + h_n^2 \sum_{j=1}^{s} x_{ij} [f_n + h_n c_j {}^{x}\!f_n + h_n (v_j \theta_n' + (c_j - v_j)\kappa_n)\, {}^{y}\!f_n + \cdots] + \cdots \\ &= \theta_n + h_n (v_i \theta_n' + (c_i - v_i)\kappa_n) + h_n^2 \left(\frac{v_i}{2} \theta_n'' + w_i \kappa_n' + (X \cdot e)_i f_n \right) \\ &\quad + h_n^3 \left(\frac{v_i}{3!} \theta_n^{(3)} + \frac{w_i}{2} \kappa_n'' + (X \cdot c)_i {}^{x}\!f_n + \{(X \cdot v)_i \theta_n' + (X \cdot (c - v))_i \kappa_n\} {}^{y}\!f_n \right) + \mathrm{O}(h_n^4) \\ &= \theta_n + \alpha_{1,i} h_n + \alpha_{2,i} h_n^2 + \alpha_{3,i} h_n^3 + \mathrm{O}(h_n^4) \end{aligned}$$

whereby the superscript denotes derivatives and the subscript n indicates that all evaluations are at $x = x_n$. Further e (also denoted as e_s) is the s-vector with unit entries and $\alpha_j = (\alpha_{j,1}, \alpha_{j,2}, \ldots, \alpha_{j,s})^{\mathrm{T}}$, $(j \in \{1,2,3\})$:

$$\begin{aligned} f(x_n + c_i h_n, Y_{n,i}) = f_n &+ h_n (c_i {}^{x}f_n + \alpha_1 {}^{y}f_n) + h_n^2 \left(\frac{c_i^2}{2} {}^{xx}f_n + c_i \alpha_1 {}^{xy}f_n + \frac{\alpha_1^2}{2} {}^{yy}f_n + \alpha_2 {}^{y}f_n \right) \\ &+ h_n^3 \left(\frac{c_i^3}{6} {}^{xxx}f_n + \frac{c_i^2}{2} \alpha_1 {}^{xxy}f_n + \frac{c_i}{2} \alpha_1^2 {}^{xyy}f_n + \frac{\alpha_1^3}{6} {}^{yyy}f_n \right. \\ &\left. + \alpha_3 {}^{y}f_n + c_i \alpha_2 {}^{xy}f_n + \alpha_1 \alpha_2 {}^{yy}f_n \right) \\ &+ \mathrm{O}(h_n^4) \end{aligned}$$

where

$$\alpha_1 = c\kappa_n + v(\theta_n' - \kappa_n),$$

$$\alpha_2 = (A \cdot e) f_n + \frac{v}{2}(\theta_n'' - f_n) + w(\kappa_n' - f_n),$$

$$\begin{aligned} \alpha_3 = &(A \cdot c)({}^{x}f_n + \kappa_n {}^{y}f_n) + \frac{v}{6}(\theta_n^{(3)} - ({}^{x}f_n + \kappa_n {}^{y}f_n)) + \frac{w}{2}(\kappa_n'' - ({}^{x}f_n + \kappa_n {}^{y}f_n)) \\ &+ (X \cdot v)(\theta_n' - \kappa_n) {}^{y}f_n \end{aligned}$$

such that for $B_i \in \{b_i, \bar{b}_i\}$ and $B \in \{b, \bar{b}\}$,

$$\begin{aligned} &\sum_{i=1}^{s} B_i f(x_n + c_i h_n, Y_{n,i}) \\ &= (B^{\mathrm{T}} \cdot e) f_n + h_n [B^{\mathrm{T}} \cdot c ({}^{x}f_n + \kappa_n {}^{y}f_n) + B^{\mathrm{T}} \cdot v(\theta_n' - \kappa_n) {}^{y}f_n] \\ &\quad + h_n^2 \left[\frac{1}{2} B^{\mathrm{T}} \cdot c^2 ({}^{xx}f_n + 2\kappa_n {}^{xy}f_n + \kappa_n^2 {}^{yy}f_n) + B^{\mathrm{T}} \cdot (cv)(\theta_n' - \kappa_n)({}^{xy}f_n + \kappa_n {}^{yy}f_n) \right. \\ &\quad \left. + \frac{1}{2} B^{\mathrm{T}} \cdot v^2 (\theta_n' - \kappa_n)^2 {}^{yy}f_n + \left(B^{\mathrm{T}} \cdot A \cdot e f_n + \frac{1}{2} B^{\mathrm{T}} \cdot v(\theta_n'' - f_n) + B^{\mathrm{T}} \cdot w(\kappa_n' - f_n) \right) {}^{y}f_n \right] \\ &\quad + h_n^3 \left[\frac{1}{6} B^{\mathrm{T}} \cdot c^3 ({}^{xxx}f_n + 3\kappa_n {}^{xxy}f_n + 3\kappa_n^2 {}^{xyy}f_n + \kappa_n^3 {}^{yyy}f_n) \right. \\ &\quad + \frac{1}{2} B^{\mathrm{T}} \cdot (c^2 v)({}^{xxy}f_n + 2\kappa_n {}^{xyy}f_n + \kappa_n^2 {}^{yyy}f_n)(\theta_n' - \kappa_n) \\ &\quad + \frac{1}{2} B^{\mathrm{T}} \cdot (cv^2)({}^{xyy}f_n + \kappa_n {}^{yyy}f_n)(\theta_n' - \kappa_n)^2 + \frac{1}{6} B^{\mathrm{T}} \cdot v^3 {}^{yyy}f_n (\theta_n' - \kappa_n)^3 \\ &\quad \left. + \left(B^{\mathrm{T}} \cdot (cA \cdot e) f_n + \frac{1}{2} B^{\mathrm{T}} \cdot (cv)(\theta_n'' - f_n) + B^{\mathrm{T}} \cdot (cw)(\kappa_n' - f_n) \right) ({}^{xy}f_n + \kappa_n {}^{yy}f_n) \right. \end{aligned}$$

$$+\left[B^{\mathrm{T}}\cdot(vA\cdot e)\,f_n+\frac{1}{2}B^{\mathrm{T}}\cdot v^2\,(\theta_n''-f_n)+B^{\mathrm{T}}\cdot(v\,w)\,(\kappa_n'-f_n)\right](\theta_n'-\kappa_n)\,{}^{yy}f_n$$

$$+\left[(B^{\mathrm{T}}\cdot A\cdot c)({}^{x}f_n+\kappa_n\,{}^{y}f_n)+\frac{1}{6}B^{\mathrm{T}}\cdot v\,(\theta_n^{(3)}-({}^{x}f_n+\kappa_n\,{}^{y}f_n))\right.$$

$$\left.+\frac{1}{2}B^{\mathrm{T}}\cdot w\,(\kappa_n''-({}^{x}f_n+\kappa_n\,{}^{y}f_n))+B^{\mathrm{T}}\cdot X\cdot v\,(\theta_n'-\kappa_n)\,{}^{y}f_n\right]{}^{y}f_n\Bigg]$$

$$+\mathrm{O}(h_n^4). \tag{3.2}$$

If $y''(x)=f(x,\,y(x))$ (i.e. $\theta=y$ and $\kappa=y'$), then the components of $\phi(\Delta z)_n$ become

$${}^{1}\phi_n=h_n\left(\frac{1}{2!}-\bar{b}^{\mathrm{T}}\cdot e\right)f_n+h_n^2\left(\frac{1}{3!}-\bar{b}^{\mathrm{T}}\cdot c\right)({}^{x}f_n+{}^{y}f_n\,y_n')$$

$$+h_n^3\left[\left(\frac{1}{4!}-\frac{1}{2}\bar{b}^{\mathrm{T}}\cdot c^2\right)({}^{xx}f_n+2\,{}^{xy}f_n\,y_n'+{}^{yy}f_n\,y_n'^2)+\left(\frac{1}{4!}-\bar{b}^{\mathrm{T}}\cdot A\cdot e\right){}^{y}f_n\,f_n\right]+\mathrm{O}(h_n^4),$$

$${}^{2}\phi_n=(1-b^{\mathrm{T}}\cdot e)\,f_n+h_n\left(\frac{1}{2!}-b^{\mathrm{T}}\cdot c\right)({}^{x}f_n+{}^{y}f_n\,y_n')$$

$$+h_n^2\left(\left(\frac{1}{3!}-\frac{1}{2}\,b^{\mathrm{T}}\cdot c^2\right)({}^{xx}f_n+2\,{}^{xy}f_n\,y_n'+{}^{yy}f_n\,y_n'^2)+\left(\frac{1}{3!}-b^{\mathrm{T}}\cdot A\cdot e\right){}^{y}f_n\,f_n\right)$$

$$+h_n^3\left(\left(\frac{1}{4!}-\frac{1}{6}\,b^{\mathrm{T}}\cdot c^3\right)({}^{xxx}f_n+3\,{}^{xxy}f_n\,y_n'+3\,{}^{xyy}f_n\,y_n'^2+{}^{yyy}f_n\,y_n'^3)\right.$$

$$\left.+\left(\frac{1}{8}-b^{\mathrm{T}}\cdot(cA\cdot e)\right)({}^{xy}f_n+{}^{yy}f_n\,y_n')\,f_n+\left(\frac{1}{4!}-b^{\mathrm{T}}\cdot A\cdot c\right)({}^{x}f_n+{}^{y}f_n\,y_n')\,{}^{y}f_n\right)$$

$$+\mathrm{O}(h_n^4).$$

One notices that, if the series expansion is carried out as far as $\mathrm{O}(h_n^p)$, in this way all the order conditions to achieve order p can be recognised. It thus becomes clear that the term in h_n^i, $0\leqslant i\leqslant p-1$ becomes zero when the method is of order p. We thus have $\phi(\Delta z)_n=\mathrm{O}(h_n^p)$. So the assumption (i) is a representation of the global error of the method ϕ with p the order of the method.

In the same way the condition (ii) of Theorem 1 expresses the order of the residual with the higher-order method ϕ^*. Analogous to the previous derivation, $\phi^*(\Delta z)_n=\mathrm{O}(h_n^{p^*})$ can be deduced.

The value r from assumption (iii) follows from the expansion of

$$\psi(\Delta u)_n=\begin{pmatrix}-h_n\left(\sum_{i=1}^{s}\bar{b}_i\,f(x_n+c_i\,h_n,\,Y_{n,i})-\sum_{i=1}^{s^*}\bar{b}_i^*\,f(x_n+c_i^*\,h_n,\,Y_i^*)\right)\\ -\sum_{i=1}^{s}b_i\,f(x_n+c_i\,h_n,\,Y_{n,i})+\sum_{i=1}^{s^*}b_i^*\,f(x_n+c_i^*\,h_n,\,Y_i^*)\end{pmatrix}.$$

To analyse this expression, we start from (3.2). The coefficient of h_n^i in ${}^{j}\phi(\Delta u)_n$ ($j\in\{1,\,2\}$) is an expression which contains combinations of partial derivatives (elementary differentials) multiplied by coefficients, which are expressed in terms of the letters $\bar{b}$ (first component) and b (second component), and A, X, c, v, w and e. Some of these coefficients have fixed values since the order conditions up to order p are fulfilled. These conditions are also fulfilled for the higher-order method ϕ^*, such that these terms no longer appear in the expression $\psi(\Delta u)=\phi(\Delta u)-\phi^*(\Delta u)$. The other

terms in $\psi(\Delta u)$ (for which the coefficients contain at least one of the parameters) do not vanish automatically. However, they will vanish as well whenever they have the same value in ϕ and ϕ^*. Therefore, we can conclude that $\psi(\Delta u) = \mathrm{O}(h_n^r)$ where $r = \min(p, q)$, $q = \max\{i | E_i = 0\}$ and

$$E_1 := 0,$$

$$E_2 := |b^{\mathrm{T}} \cdot v - b^{*\mathrm{T}} \cdot v^*|,$$

$$E_3 := |b^{\mathrm{T}} \cdot (c\,v) - b^{*\mathrm{T}} \cdot (c^*\,v^*)| + |b^{\mathrm{T}} \cdot v^2 - b^{*\mathrm{T}} \cdot v^{*2}| + |b^{\mathrm{T}} \cdot w - b^{*\mathrm{T}} \cdot w^*| + |\bar{b}^{\mathrm{T}} \cdot v - \bar{b}^{*\mathrm{T}} \cdot v^*|,$$

$$\begin{aligned} E_4 := {} & |b^{\mathrm{T}} \cdot v^3 - b^{*\mathrm{T}} \cdot v^{*3}| + |b^{\mathrm{T}} \cdot (c\,v^2) - b^{*\mathrm{T}} \cdot (c^*\,v^{*2})| + |b^{\mathrm{T}} \cdot (c^2\,v) - b^{*\mathrm{T}} \cdot (c^{*2}\,v^*)| \\ & + |b^{\mathrm{T}} \cdot (c\,w) - b^{*\mathrm{T}} \cdot (c^*\,w^*)| + |b^{\mathrm{T}} \cdot (v\,w) - b^{*\mathrm{T}} \cdot (v^*\,w^*)| + |b^{\mathrm{T}} \cdot X \cdot v - b^{*\mathrm{T}} \cdot X^* \cdot v^*| \\ & + b^{\mathrm{T}} \cdot ((X \cdot e)\,v) - b^{*\mathrm{T}} \cdot ((X^* \cdot e^*)\,v^*)| \\ & + |\bar{b}^{\mathrm{T}} \cdot v - \bar{b}^{*\mathrm{T}} \cdot v^*| + |\bar{b}^{\mathrm{T}} \cdot w - \bar{b}^{*\mathrm{T}} \cdot w^*| + |\bar{b}^{\mathrm{T}} \cdot (c\,v) - \bar{b}^{*\mathrm{T}} \cdot (c^*\,v^*)| + |\bar{b}^{\mathrm{T}} \cdot v^2 - \bar{b}^{*\mathrm{T}} \cdot v^{*2}|, \end{aligned}$$

$$E_5 := \cdots.$$

The following rule of thumb can be used to construct E_i: this expression contains all meaningfull strings of weight i that start with b or $\bar{b}$ and end with either v or w. The weight of a string is defined as the sum of the weights of its letters where $\bar{b}$, X and w have weight two, b, c and v weight one and e weight zero. We remark that the weight of a string that expresses an order condition of order i is just i by this definition.

In the case of first-order systems E_i contains similar expressions [16]: it are all meaningful strings of weight i that start with b and end with v. The weight of a string is defined as the sum of the weights of its letters where b, X, c, v have weight 1 and e weight 0.

In view of Theorem 1, we can state that the order of the DC scheme (1.2) can be $2p$ at most. This is the case if $r = p = q$ and thus $p^* = p + r = 2p$. This is in agreement with Böhmer et al. [3], who state that a pth-order driver-operator and a p^*th-order target operator ($p < p^*$) make a numerical approximation with order of accuracy $\min(ip, p^*)$ at most in the ith iteration. The question now is how to construct such high-order convergent schemes. We now know that besides the (parameter independent) order conditions extra (parameter dependent) conditions $E_i = 0$ have to imposed. One may thus ask whether there exist parameter pairs (v, w) and (v^*, w^*) which guarantee the maximal increase of order. Due to the above rule which expresses the weight of strings, we have the following result.

Theorem 2. *The order of the DC scheme* (1.2) *will be* $\min\{p^*, 2p\}$ *if there exist constants* α_0, α_1, β_0, β_1 *and* β_2 *such that*

$$v = \alpha_0 e + \alpha_1 c, \quad w = \beta_0\, e + \beta_1\, c + \beta_2\, c^2,$$

$$v^* = \alpha_0 e^* + \alpha_1 c^*, \quad w^* = \beta_0\, e^* + \beta_1\, c^* + \beta_2\, c^{*2}.$$

If symmetry is required then $2\alpha_0 + \alpha_1 = 1$ and $\beta_0 + 2\beta_1 + 2\beta_2 = 1$. In the case of symmetric MIRKN methods and Lobatto methods, we want to identify the first two rows of the modified tableau with y_n and y_{n+1}. This will be realised when

$$v = c \quad \text{and} \quad w = \beta_2 c(c - e), \tag{3.3}$$

where the choice $v = c$ is the same as in the case of first-order systems.

The above results can also be extended to the case of iterated DC schemes of the form

$$\phi^{[1]}(\eta^{[1]}) = 0,$$

$$\phi^{[1]}(\eta^{[i]}) = \phi^{[1]}(\eta^{[i-1]}) - \phi^{[i]}(\eta^{[i-1]}), \quad i = 2, \ldots, m. \tag{3.4}$$

Suppose $\phi^{[i]}$ is of order $p_i = ip$, then the overall scheme will have order $p_m = m\,p$ if all parameters $v^{[i]}$ and $w^{[i]}$ are the same linear and quadratic functions of the corresponding $c^{[i]}$ vectors:

$$v^{[i]} = \alpha_0\, e^{[i]} + \alpha_1\, c^{[i]},$$

$$i = 1, \ldots, m.$$

$$w^{[i]} = \beta_0 e^{[i]} + \beta_1\, c^{[i]} + \beta_2 c^{[i]^2},$$

4. Linear stability of DC schemes

As Asher [1] already remarked, the stability of the deferred correction scheme when $h \to 0$ follows from the stability of the driver-operator. For stiff problems another type of stability where the product of the steplength and the eigenvalues of the problem tends to infinity, is of importance and will be examined further on.

When applying a PIRKN method with constant mesh size h to the scalar test equation $y'' = -\lambda^2 y$ one finds

$$\begin{pmatrix} 1 + H^2 \bar{b}^{\mathrm{T}} \cdot U \cdot v & H^2 \bar{b}^{\mathrm{T}} \cdot U \cdot w \\ H^2 b^{\mathrm{T}} \cdot U \cdot v & 1 + H^2 b^{\mathrm{T}} \cdot U \cdot w \end{pmatrix} \begin{pmatrix} y_{n+1} \\ h y'_{n+1} \end{pmatrix} \tag{4.1}$$

$$= \begin{pmatrix} 1 - H^2 \bar{b}^{\mathrm{T}} \cdot U \cdot (e - v) & 1 - H^2 \bar{b}^{\mathrm{T}} \cdot U \cdot (c - v - w) \\ -H^2 b^{\mathrm{T}} \cdot U \cdot (e - v) & 1 - H^2 b^{\mathrm{T}} \cdot U \cdot (c - v - w) \end{pmatrix} \begin{pmatrix} y_n \\ h y'_n \end{pmatrix}, \tag{4.2}$$

with $U = (I_s + H^2 X)^{-1}$ and $H = \lambda h$. This can be expressed in the form

$$D\eta_{n+1} = N\eta_n, \quad D, N \in \mathbb{R}^{2d \times 2d}, \quad \eta_{n+1},\ \eta_n \in \mathbb{R}^{2d}\,. \tag{4.3}$$

The DC scheme (3.4) applied to $y'' = -\lambda^2 y$ then gives

$$\eta_{n+1}^{[1]} = D_1^{-1} \cdot N_1 \eta_n^{[1]},$$

$$\eta_{n+1}^{[i]} = D_1^{-1} \cdot (D_1 - D_i)\eta_{n+1}^{[i-1]} + D_1^{-1} \cdot (N_i - N_1)\, \eta_n^{[i-1]} + D_1^{-1} \cdot N_1\, \eta_n^{[i]}$$

with $i = 2, \ldots, m$ and D_i and N_i as defined in (4.3). In particular, for $m = 2$ we find

$$\eta_n^{[2]} = R^n \eta_0^{[2]} + \sum_{i=0}^{n-1} R^i \cdot S \cdot R^{n-i-1} \eta_0^{[1]} \tag{4.4}$$

with $R = D_1^{-1} \cdot N_1$ and $S = D_1^{-1} \cdot (N_2 - D_2 \cdot D_1^{-1} \cdot N_1)$. When R and S commute, (4.4) simplifies to

$$\eta_n^{[2]} = R^n\, \eta_0^{[2]} + n\, R^{n-1} \cdot S\, \eta_0^{[1]}. \tag{4.5}$$

Analogous as in [8] we call a RKN method stable if the solutions (4.5) are bounded for each finite n and for each H. To satisfy this type of stability, it is necessary that R and S remain finite for $|H| \to \infty$. This will be taken into consideration when selecting a DC scheme in the following section.

5. A high-order convergent DC scheme based on MIRKN methods

In this section we will construct a pair of symmetric MIRKN methods of orders 4 and 8 for which the order of the overall DC scheme reaches its maximum value, which is 8 according to Theorem 2.

5.1. The fourth-order method

The symmetric fourth-order method with the minimal number of stages is the Gauss method with $s = 2$ stages. There is a unique parameterization of the method as a symmetric MIRKN method, if one chooses $\beta_2 = \frac{1}{2}$ in (3.3):

$$\begin{array}{c|cc|c|cc} \frac{1}{2}-\frac{\sqrt{3}}{6} & \frac{1}{2}-\frac{\sqrt{3}}{6} & & -\frac{1}{12} & 0 & 0 \\ \frac{1}{2}+\frac{\sqrt{3}}{6} & \frac{1}{2}+\frac{\sqrt{3}}{6} & & -\frac{1}{12} & 0 & 0 \\ \hline & & & & \frac{1}{4}+\frac{\sqrt{3}}{12} & \frac{1}{4}-\frac{\sqrt{3}}{12} \\ & & & & \frac{1}{2} & \frac{1}{2} \end{array} \tag{5.1}$$

5.2. The eighth-order method

Our aim is to construct a symmetric MIRKN method of order 8 with the minimum number of stages and which allows high-order convergence when combined with (5.1) in a DC framework. Based on (2.2) we propose the nonconfluent scheme

$$\begin{array}{c|cc|cccccccc} 0 & 0 & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ c_3 & c_3 & \frac{1}{2}c_3(c_3-1) & x_{31} & x_{32} & 0 & 0 & 0 & 0 & 0 & 0 \\ 1-c_3 & 1-c_3 & \frac{1}{2}c_3(c_3-1) & x_{32} & x_{31} & 0 & 0 & 0 & 0 & 0 & 0 \\ c_5 & c_5 & \frac{1}{2}c_5(c_5-1) & x_{51} & x_{52} & x_{53} & x_{54} & 0 & 0 & 0 & 0 \\ 1-c_5 & 1-c_5 & \frac{1}{2}c_5(c_5-1) & x_{52} & x_{51} & x_{54} & x_{53} & 0 & 0 & 0 & 0 \\ c_7 & c_7 & \frac{1}{2}c_7(c_7-1) & x_{71} & x_{72} & x_{73} & x_{74} & x_{75} & x_{76} & 0 & 0 \\ 1-c_7 & 1-c_7 & \frac{1}{2}c_7(c_7-1) & x_{72} & x_{71} & x_{74} & x_{73} & x_{76} & x_{75} & 0 & 0 \\ \hline & & & \bar{b}_1 & b_1-\bar{b}_1 & \bar{b}_3 & b_3-\bar{b}_3 & \bar{b}_5 & b_5-\bar{b}_5 & \bar{b}_7 & b_7-\bar{b}_7 \\ & & & b_1 & b_1 & b_3 & b_3 & b_5 & b_5 & b_7 & b_7 \end{array} \tag{5.2}$$

To obtain such a scheme, we will fully exploit the symmetry in the modified Butcher tableau. Therefore, we introduce the matrices

$$U := \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad V := \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},$$

and write $X = Y \otimes U + Z \otimes V$ where

$$Y = \begin{pmatrix} 0 & 0 & 0 & 0 \\ y_{31} & 0 & 0 & 0 \\ y_{51} & y_{53} & 0 & 0 \\ y_{71} & y_{73} & y_{75} & 0 \end{pmatrix} \quad \text{and} \quad Z = \begin{pmatrix} 0 & 0 & 0 & 0 \\ z_{31} & 0 & 0 & 0 \\ z_{51} & z_{53} & 0 & 0 \\ z_{71} & z_{73} & z_{75} & 0 \end{pmatrix}$$

with $y_{ij} = x_{ij} + x_{i,j+1}$ and $z_{ij} = x_{ij} - x_{i,j+1}$. Further, we also define the vector $d = e - 2c$ and the diagonal matrices

$$D_+^{[k]} = \operatorname{diag}\left(\frac{1}{2}, \frac{(1-d_3)^k + (1+d_3)^k}{2^{k+1}}, \frac{(1-d_5)^k + (1+d_5)^k}{2^{k+1}}, \frac{(1-d_7)^k + (1+d_7)^k}{2^{k+1}}\right),$$

$$D_-^{[k]} = \operatorname{diag}\left(-\frac{1}{2}, \frac{(1-d_3)^k - (1+d_3)^k}{2^{k+1}}, \frac{(1-d_5)^k - (1+d_5)^k}{2^{k+1}}, \frac{(1-d_7)^k - (1+d_7)^k}{2^{k+1}}\right).$$

For these matrices we have the relations

$$\sum_{i=0}^{n} (-1)^i \binom{n}{i} D_-^{[n+i]} = 0 \quad \text{and} \quad \sum_{i=0}^{n} (-1)^i \binom{n}{i} (n+i) D_+^{[n+i-1]} = 0. \tag{5.3}$$

The order conditions can now be solved as follows:

1. The quadrature conditions $b^{\mathrm{T}} \cdot c^{2i} = 1/(2i+1)$, $i = 0, 1, \ldots, 3$ determine the vector b in terms of c_3, c_5 and c_7.
2. Imposing the relation $\bar{b} = b(e - c)$ [11] and stage order 3, only 5 order conditions are left:

$$b^{\mathrm{T}} \cdot X \cdot c^2 = \tfrac{1}{360}\text{: (a) } B^{\mathrm{T}} \cdot Y \cdot D_+^{[2]} \cdot e_4 = \tfrac{1}{720},$$

$$b^{\mathrm{T}} \cdot (c^2 X \cdot c^2) = -\tfrac{1}{1680}\text{: (b) } B^{\mathrm{T}} \cdot D_+^{[2]} \cdot Y \cdot D_+^{[2]} \cdot e_4 + B^{\mathrm{T}} \cdot D_-^{[2]} \cdot Z \cdot D_-^{[2]} \cdot e_4 = -\tfrac{1}{3360},$$

$$b^{\mathrm{T}} \cdot (cX \cdot c^3) = \tfrac{1}{1120}\text{: (c) } B^{\mathrm{T}} \cdot D_+^{[1]} \cdot Y \cdot D_+^{[3]} \cdot e_4 + B^{\mathrm{T}} \cdot D_-^{[1]} \cdot Z \cdot D_-^{[3]} \cdot e_4 = \tfrac{1}{2240},$$

$$b^{\mathrm{T}} \cdot X \cdot c^4 = \tfrac{1}{210}\text{: (d) } B^{\mathrm{T}} \cdot Y \cdot D_+^{[4]} \cdot e_4 = \tfrac{1}{420},$$

$$b^{\mathrm{T}} \cdot X^2 \cdot c^2 = -\tfrac{1}{15120}\text{: (e) } B^{\mathrm{T}} \cdot Y^2 \cdot D_+^{[2]} \cdot e_4 = -\tfrac{1}{30240},$$

   where $B^{\mathrm{T}} = (b_1, b_3, b_5, b_7)$.
3. Stage order 2 implies $X \cdot e = 0$ and since

$$\begin{aligned} X \cdot e &= (Y \otimes U + Z \otimes V) \cdot (e_4 \otimes e_2) \\ &= Y \cdot e_4 \otimes e_2, \end{aligned} \tag{5.4}$$

   we have $Y \cdot e_4 = 0$. Eqs. (a) and (d) combined with $B^{\mathrm{T}} \cdot Y \cdot e_4 = 0$ which follows from (5.4), form a linear system which determines the components of $B^{\mathrm{T}} \cdot Y = (b_5 y_{51} + b_7 y_{71}, b_5 y_{53} + b_7 y_{73}, b_7 y_{75}, 0)$. Together with $Y \cdot e_4 = 0$, this determines the unknowns y_{51}, y_{53}, y_{71} and y_{73} in terms of y_{75}, d_3, d_5 and d_7.
4. One of the conditions that is identically satisfied if stage order 3 holds is

$$b^{\mathrm{T}} \cdot (cX \cdot c) = -\frac{1}{720}\text{: (f) } B^{\mathrm{T}} \cdot D_+^{[1]} \cdot Y \cdot D_+^{[1]} \cdot e_4 + B^{\mathrm{T}} \cdot D_-^{[1]} \cdot Z \cdot D_-^{[1]} \cdot e_4 = -\frac{1}{1440}.$$

Since $D_+^{[1]}$ is a scalar matrix and $Y \cdot e_4 = 0$ and $D_-^{[2]} = D_-^{[1]}$ one has

$$B^{\mathrm{T}} \cdot D_-^{[2]} \cdot Z \cdot D_-^{[2]} \cdot e_4 = -\frac{1}{1440}$$

such that condition (b) can be replaced by $B^{\mathrm{T}} \cdot D_+^{[2]} \cdot Y \cdot D_+^{[2]} \cdot e_4 = \frac{1}{2520}$. This yields the remaining y_{75} in terms of d_3, d_5 and d_7 and all y_{ij} coefficients are now fixed in terms of d_3, d_5 and d_7. Substituting the values in (e) one fixes the value $d_7 = 1/\sqrt{3}$.

5. Finally, we have to determine the z_{ij} coefficients. The only condition which has not been used so far is condition (c), which, due to $D_+^{[1]} - 3D_+^{[2]} + 2D_+^{[3]} = 0$ obtained from (5.3) can be reformulated as

$$B^{\mathrm{T}} \cdot D_-^{[1]} \cdot Z \cdot D_-^{[3]} \cdot e_4 = -\frac{1}{1680}.$$

Stage order 3 gives

$$X \cdot c = \frac{1}{48} d(e - d^2),$$

where $X \cdot c = Z \cdot D_-^{[1]} \cdot e_4$. This results in a system of four equations in six unknowns z_{ij}.

We thus find the following family of methods which depends on 4 free parameters d_3, d_5, z_{53} and either z_{73} or z_{75}:

$$b_1 = \frac{6 - 7d_3^2 - 7d_5^2}{105(d_3^2 - 1)(d_5^2 - 1)},$$

$$b_3 = \frac{2(7d_5^2 + 1)}{105(d_3^2 - 1)(3d_3^2 - 1)(d_5^2 - d_3^2)},$$

$$b_5 = \frac{2(7d_3^2 + 1)}{105(d_5^2 - 1)(3d_5^2 - 1)(d_3^2 - d_5^2)},$$

$$b_7 = \frac{9(3 - 7d_5^2 - 7d_3^2 + 35d_3^2 d_5^2)}{70(3d_3^2 - 1)(3d_5^2 - 1)} \neq 0,$$

$$\bar{b}_i = b_i \frac{1 + d_i}{2}, \quad i = 1, \ldots, 4,$$

$$y_{31} = 0, \quad y_{51} = -y_{53}, \quad y_{71} = -y_{73} - y_{75},$$

$$y_{53} = -\frac{(d_3^2 - d_5^2)(d_5^2 - 1)}{6(1 + 7d_3^2)(d_3^2 - 1)},$$

$$y_{73} = \frac{(1 - 3d_3^2)(1 + 4d_3^2 - 21d_5^4)}{162(d_3^2 - 1)(d_5^2 - d_3^2)(3 - 7d_5^2 + 7d_3^2(5d_5^2 - 1))},$$

$$y_{75} = \frac{(3d_3^2 - 1)(7d_3^2 + 1)(3d_5^2 - 1)}{162(d_5^2 - 1)(d_3^2 - d_5^2)(3 - 7d_5^2 + 7d_3^2(5d_5^2 - 1))},$$

$$z_{31} = \frac{d_3(d_3^2 - 1)}{24}, \quad z_{51} = \frac{d_5(d_5^2 - 1)}{24} - d_3 z_{53}, \quad z_{71} = -\frac{\sqrt{3}}{108} - (d_3 z_{73} + d_5 z_{75}),$$

and (if z_{75} is a free parameter)

$$z_{73} = \frac{3d_3^2 - 1}{9\sqrt{3}(3 - 7d_5^2 - 7d_3^2 + 35d_3^2d_5^2)} \left(\frac{3d_5^2 - 1}{3d_3(d_3^2 - 1)} - \frac{4(1 + 7d_3^2)d_5 z_{53}}{(d_3^2 - d_5^2)(d_5^2 - 1)} \right) + \frac{(d_5 - d_5^3)z_{75}}{d_3^3 - d_3},$$

or (if z_{53} is a free parameter)

$$z_{75} = \frac{3d_3^2 - 1}{9\sqrt{3}(3 - 7d_5^2 - 7d_3^2 + 35d_3^2d_5^2)} \left(\frac{3d_5^2 - 1}{3d_5(d_5^2 - 1)} + \frac{4(d_3^2 - 1)\,(1 + 7d_5^2)\,d_3 z_{53}}{(d_5^2 - d_3^2)(d_5^2 - 1)^2} \right) + \frac{(d_3 - d_3^3)z_{73}}{d_5^3 - d_5}.$$

We can however use the free parameters to reduce the number of stages: choosing either (a) $d_3 = 0 = z_{53} = z_{73}$, or (b) $d_5 = 0 = z_{53} = z_{75}$ results in a scheme where two rows coincide, such that the corresponding columns can be added together. We thus find two 1-parameter families of 7-stage order 8 MIRKN methods. Below, we present the method for which $d_3 = 0$, i.e., $c_3 = \frac{1}{2}$.

$$\begin{array}{c|c|c|cccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{2} & \frac{1}{2} & -\frac{1}{8} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1-d_5}{2} & \frac{1-d_5}{2} & \frac{d_5^2-1}{8} & x_{51} & x_{52} & \frac{d_5^2}{6}(1-d_5^2) & 0 & 0 & 0 & 0 \\
\frac{1+d_5}{2} & \frac{1+d_5}{2} & \frac{d_5^2-1}{8} & x_{52} & x_{51} & \frac{d_5^2}{6}(1-d_5^2) & 0 & 0 & 0 & 0 \\
\frac{3-\sqrt{3}}{6} & \frac{3-\sqrt{3}}{6} & -\frac{1}{12} & x_{71} & x_{72} & \frac{1-21d_5^4}{162d_5^2(7d_5^2-3)} & x_{74} & x_{75} & 0 & 0 \\
\frac{3+\sqrt{3}}{6} & \frac{3+\sqrt{3}}{6} & -\frac{1}{12} & x_{72} & x_{71} & \frac{1-21d_5^4}{162d_5^2(7d_5^2-3)} & x_{75} & x_{74} & 0 & 0 \\
\hline
 & & & \frac{7d_5^2-6}{105(d_5^2-1)} & 0 & \frac{2(7d_5^2+1)}{105d_5^2} & \bar{b}_4 & \bar{b}_5 & \bar{b}_6 & \bar{b}_7 \\
 & & & \frac{7d_5^2-6}{105(d_5^2-1)} & \frac{7d_5^2-6}{105(d_5^2-1)} & \frac{4(7d_5^2+1)}{105d_5^2} & b_4 & b_4 & b_6 & b_6
\end{array} \tag{5.5}$$

$$x_{51} = \frac{d_5(1 + 4d_5)(d_5^2 - 1)}{48}, \qquad x_{52} = \frac{d_5(1 - 4d_5)(1 - d_5^2)}{48},$$

$$x_{71} = \frac{4 - 5\sqrt{3} + 6(3\sqrt{3} - 7)d_5^2 - 21(\sqrt{3} - 2)d_5^4}{648(d_5^2 - 1)(7d_5^2 - 3)},$$

$$x_{72} = \frac{4 + 5\sqrt{3} - 6(3\sqrt{3} + 7)d_5^2 + 21(\sqrt{3} + 2)d_5^4}{648(d_5^2 - 1)(7d_5^2 - 3)},$$

$$x_{74} = \frac{(2\sqrt{3}d_5 - 1)(3d_5^2 - 1)}{324d_5^2(d_5^2 - 1)(7d_5^2 - 3)}, \qquad x_{75} = \frac{(2\sqrt{3}d_5 + 1)(1 - 3d_5^2)}{324d_5^2(d_5^2 - 1)(7d_5^2 - 3)}$$

$$\bar{b}_4 = \frac{1}{105(1 - d_5)d_5^2(3d_5^2 - 1)}, \qquad \bar{b}_5 = \frac{1}{105(1 + d_5)d_5^2(3d_5^2 - 1)},$$

$$\bar{b}_6 = \frac{3(3+\sqrt{3})(7d_5^2-3)}{140(3d_5^2-1)}, \qquad \bar{b}_7 = \frac{3(3-\sqrt{3})(7d_5^2-3)}{140(3d_5^2-1)},$$

$$b_4 = \frac{2}{105d_5^2(1-d_5^2)(3d_5^2-1)}, \qquad b_6 = \frac{9(7d_5^2-3)}{70(3d_5^2-1)}.$$

5.3. Linear stability of the DC scheme

The linear stability matrices D and N from (4.3) of the two-stage method (5.1) are given by

$$D_1 = \begin{pmatrix} 1+\frac{H^2}{6} & -\frac{H^2}{24} \\ \frac{H^2}{2} & 1-\frac{H^2}{12} \end{pmatrix}, \quad N_1 = \begin{pmatrix} 1-\frac{H^2}{3} & 1-\frac{H^2}{24} \\ -\frac{H^2}{2} & 1-\frac{H^2}{12} \end{pmatrix}.$$

For the 8-stage family of order 8 methods we have the following result:

$$D_2 = \begin{pmatrix} 1+\frac{H^2}{6}-\frac{H^4}{720}-\frac{H^6}{30240}-a_8H^8 & -\frac{H^2}{24}-\frac{H^4}{1440}-\frac{H^6}{60480} \\ \frac{H^2}{2} & 1-\frac{H^2}{12}-\frac{H^4}{720}-\frac{H^6}{30240} \end{pmatrix},$$

$$N_2 = \begin{pmatrix} 1-\frac{H^2}{3}-\frac{H^4}{720}-\frac{H^6}{30240}-a_8H^8 & 1-\frac{H^2}{24}-\frac{H^4}{1440}-\frac{H^6}{60480} \\ -\frac{H^2}{2} & 1-\frac{H^2}{12}-\frac{H^4}{720}-\frac{H^6}{30240} \end{pmatrix},$$

where $a_8 = \frac{1}{2}b_7d_7z_{75}z_{53}z_{31}$.

One finds that the matrices R and S in (4.4) commute if and only if $a_8 = 0$. In that case they are given by

$$R = \begin{pmatrix} \frac{144-60H^2+H^4}{144+12H^2+H^4} & \frac{12(12-H^2)}{144+12H^2+H^4} \\ \frac{12H^2(H^2-12)}{144+12H^2+H^4} & \frac{144-60H^2+H^4}{144+12H^2+H^4} \end{pmatrix},$$

$$S = \begin{pmatrix} \frac{2H^6(-504+30H^2+H^4)}{36(144+12H^2+H^4)^2} & \frac{H^4(6046-2376H^2-18H^4+H^6)}{210(144+12H^2+H^4)^2} \\ \frac{-H^6(6046-2376H^2-18H^4+H^6)}{210(144+12H^2+H^4)^2} & \frac{2H^6(-504+30H^2+H^4)}{35(144+12H^2+H^4)^2} \end{pmatrix}. \tag{5.6}$$

The matrix S is however unbounded for $H \to \infty$, such that for problems where H^2 becomes large severe errors may be expected.

6. A high-order convergent DC scheme based on Lobatto IIIA methods

In this section we briefly discuss a DC scheme based on the parameterized Lobatto IIIA methods of orders 4 and 8.

6.1. The fourth-order method

$$\begin{array}{c|cc|ccc} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{8} & 0 & 0 & 0 \\ \hline & & & \frac{1}{6} & \frac{1}{3} & 0 \\ & & & \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{array}. \tag{6.1}$$

6.2. The eighth-order method

$$\begin{array}{c|ccc|ccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{7-\sqrt{21}}{14} & \frac{7-\sqrt{21}}{14} & -\frac{1}{14} & \frac{1}{392} & \frac{1}{392} & -\frac{5}{504} & -\frac{4}{441} & \frac{1}{72} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{8} & \frac{1}{128} & \frac{1}{128} & \frac{7}{1152} & -\frac{1}{36} & \frac{7}{1152} \\ \frac{7+\sqrt{21}}{14} & \frac{7+\sqrt{21}}{14} & -\frac{1}{14} & \frac{1}{392} & \frac{1}{392} & \frac{1}{72} & -\frac{4}{441} & -\frac{5}{504} \\ \hline & & & \frac{1}{20} & 0 & \frac{49+7\sqrt{21}}{360} & \frac{8}{45} & \frac{49-7\sqrt{21}}{360} \\ & & & \frac{1}{20} & \frac{1}{20} & \frac{49}{180} & \frac{16}{45} & \frac{49}{180} \end{array} \tag{6.2}$$

6.3. Linear stability of the DC scheme

The stability matrix R is again given by (5.6) and S is given by

$$S = \begin{pmatrix} -\frac{504}{5}\dfrac{H^6(-12+H^2)}{(144+12H^2+H^4)^2(H^2-42)} & -\frac{42}{5}\dfrac{H^4(144-60H^2+H^4)}{(144+12H^2+H^4)^2(H^2-42)} \\ \frac{42}{5}\dfrac{H^6(144-60H^2+H^4)}{(144+12H^2+H^4)^2(H^2-42)} & -\frac{504}{5}\dfrac{H^6(-12+H^2)}{(144+12H^2+H^4)^2(H^2-42)} \end{pmatrix}.$$

Both R and S (which again commute) are bounded, thus we have constructed a stable DC scheme.

7. Numerical examples

We apply the high-order convergent DC schemes with fixed step size to the problem [12]

$$y'' = \lambda^2 y, \quad y(0) = 1, \quad y(1) = 0, \quad x \in [0,\ 1], \tag{7.1}$$

whose solution is given by $y(x) = (\exp(-\lambda x) - \exp(\lambda(x-2)))/(1-\exp(-2\lambda))$. For large values of λ, (7.1) is a singular perturbation problem and its solution has a boundary layer of width $O(\lambda^{-1})$ at $x = 0$.

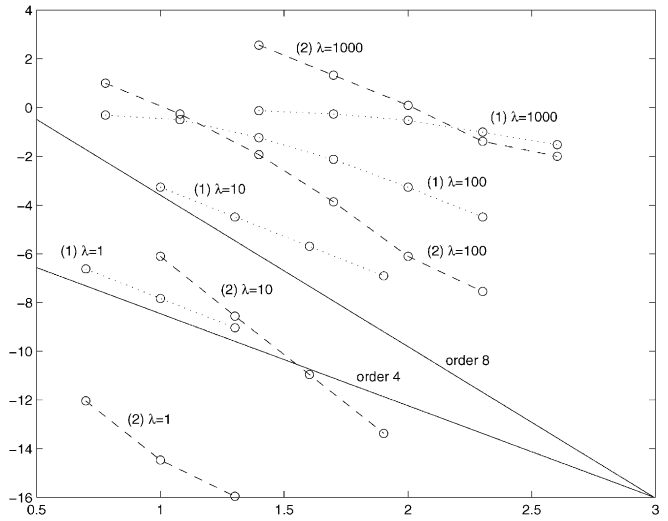

Fig. 1. $\log_{10}$ abs.error$_{\max}$ vs. $-\log_{10} h$ after one stage (1) or both stages (2) of the MIRKN high-order convergent order 8 DC scheme when applied to (7.1).

First we consider the MIRKN based DC scheme. The numerical solution clearly shows the increase of the algebraic order by four as can be seen in Fig. 1. Here the maximum of the absolute value of the absolute error over the integration interval is plotted as a function of the number of grid points in the interval [0, 1] in a ($\log_{10}$, $\log_{10}$) scale for several values of λ and for the two stages in our DC scheme marked respectively as (1) and (2). The solid lines indicate respectively orders 4 and 8. The instability of the DC scheme in case of large values of $|\lambda h|$ can be noticed in the case where $\lambda = 100$ or 1000. Since R (which is the stability matrix of the P-stable Gauss method of order 4) is bounded, bounded errors are generated in the first stage of the DC scheme, but the unboundedness of S causes unbounded errors in the second stage of the overall method.

Fig. 2 displays similar results for the DC scheme based on the Lobatto IIIA methods of orders 4 and 8, except for $\lambda = 1000$ where bounded errors are obtained due to the stability of the scheme.

One may think that, in the case of the MIRKN DC scheme, the unbounded behaviour for large values of H is related to the explicit computation of the deferred correction. To prove that this is not the case, we also applied three-other DC schemes, L24, L26 and L28, on (7.1) with $\lambda = 1000$. For each scheme the basic method is the second-order Lobatto IIIA method and the method ϕ^* is the fourth (respectively sixth and eighth) order Lobatto IIIA method. For all four methods the parameterization (3.3) with $\beta = \frac{1}{2}$ was used. The resulting schemes (which all have order four) have the following properties:

- L24 : explicit deferred corrections with a nonstable matrix S,
- L26 : implicit deferred corrections with a stable matrix S,
- L28 : implicit deferred corrections with a non-stable matrix S.

From Fig. 3 it is clear that the unstable behaviour is related to the unboundedness of S, not to the explicitness of the deferred corrections.

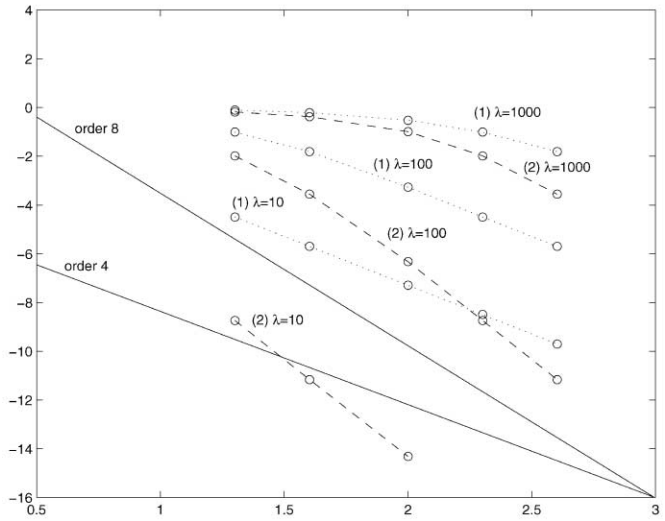

Fig. 2. $\log_{10}$ abs.error$_{\max}$ vs. $-\log_{10} h$ after one stage (1) or both stages (2) of the high-order convergent order 8 DC scheme of Lobatto IIIA methods when applied to (7.1).

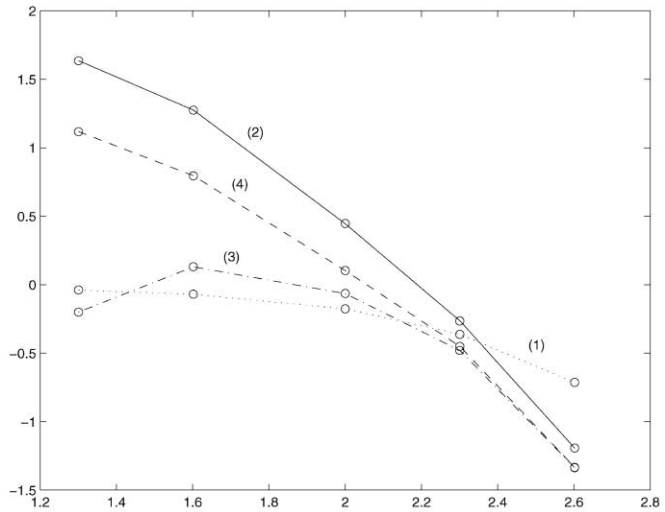

Fig. 3. $\log_{10}$ abs.error$_{\max}$ vs. $-\log_{10} h$ after one stage with ϕ the Lobatto IIIA method of order 2 (1) or both stages of the order 8 DC scheme with ϕ^* the Lobatto IIIA method of orders (2) 4, (3) 6, (4) 8 when applied to (7.1) with $\lambda = 1000$.

Finally, we briefly discuss a simple variable step size implementation. Table 1 contains the max. norm of the absolute errors of the numerical solution obtained by a variable step implementation of the DC scheme based on the constructed fourth and eighth order MIRKN methods. These results are compared with those of TWPBVP [6,7] which inspired us for the step size strategy. We omit the details because it was not our intention to build a competitive code. The total number of gridpoints and the individual size of the used grids are added between brackets.

Table 1
Abs. errors for the variable step size implementation with DC MIRKN48 and TWPBVP

	$\lambda = 10$	$\lambda = 100$	$\lambda = 1000$
DC MIRKN48	$6.1e-9$ $(30 = 5 + 25)$	$3.4e-10$ $(80 = 22 + 58)$	$4.6e-11$ $(274 = 82 + 192)$
TWPBVP	$2.1e-9$ $(37 = 7 + 12 + 18)$	$9.8e-10$ $(81 = 7 + 13 + 25 + 36)$	$1.8e-10$ $(358 = 7 + 13 + 25 + 49 + 97 + 167)$

8. Conclusion

In this paper, we have discussed the possibility of constructing high-order convergent (iterated) DC schemes based on two (or more) PIRKN methods. We found that the conditions to obtain high-order convergence can be expressed in terms of the parameters v and w of these PIRKN methods. If the parameters are chosen in an appropriate way, all conditions are identically fulfilled.

We focussed our attention mainly on the construction of an order 8 DC scheme which consists of two MIRKN methods of orders 4 and 8 since these MIRKN methods offer some computational advantages. A brief investigation of the linear stability properties of this scheme shows however that this particular scheme is not very well suited to solve problems with rapidly varying solutions. For such problems, the stable scheme based on Lobatto IIIA methods is more suited.

References

[1] U.M. Asher, R.M.M. Mattheij, R.D. Russell, Numerical Solution of Boundary Value Problems for Ordinary Differential Equations, Prentice-Hall, Englewood Cliffs, NJ, 1988.
[2] Z. Bashir-Ali, J.R. Cash, H.H.M. Silva, Lobatto deferred correction for stiff two-point boundary value problems, Comp. Math. Appl. 36 (1998) 59–69.
[3] K. Böhmer, P. Hemker, H.J. Stetter, The defect correction approach, Comput. Suppl. 5 (1984) 1–32.
[4] J.C. Butcher, The Numerical Analysis of Ordinary Differential Equations: Runge–Kutta and General Linear Methods, Wiley, Chichester, 1987.
[5] J.R. Cash, The numerical integration of nonlinear two-point boundary value problems using iterated deferred correction, Part 2: the development and analysis of highly stable deferred correction formulae, SIAM J. Numer. Anal. 25 (1988) 862–882.
[6] J.R. Cash, M.H. Wright, A deferred correction method for nonlinear two point boundary value problems: implementation and numerical evaluation, SIAM J. Sci. Statist. Comput. 12 (1991) 971–989.
[7] J.R. Cash, G. Moore, R.W. Wright, An automatic continuation strategy for the solution of singularly perturbed linear two point boundary value problems, J. Comput. Phys. 122 (1995) 266–279.
[8] J.R. Cash, H.H.M. Silva, Iterated deferred correction for linear two-point boundary value problems, Comput. Appl. Math. 15 (1) (1996) 55–75.
[9] J.R. Cash, R.W. Wright, An automatic continuation strategy for the solution of singularly perturbed linear two point boundary value problems, APNUM 28 (1998) 59–69.
[10] J.W. Daniel, A.J. Martin, Numerov's method with deferred corrections for two-point boundary-value problems, SIAM J. Numer. Anal. 14 (1977) 1033–1050.
[11] E. Hairer, S.P. Nørsett, G. Wanner, Solving Ordinary Differential Equations I, Nonstiff Problems, Springer, Heidelberg, 1987.
[12] P.W. Hemker, A Numerical Study of Stiff Two-Point Boundary Value Problems, Mathematisch Centrum, Amsterdam, 1977.

[13] R.D. Skeel, A theoretical framework for proving accuracy results for deferred correction, SIAM J. Numer. Anal. 19 (1981) 171–196.
[14] M. Van Daele, H. De Meyer, T. Van Hecke, G. Vanden Berghe, On a class of P-stable mono-implicit Runge–Kutta–Nyström methods, Appl. Numer. Math. 27 (1998) 69–82.
[15] M. Van Daele, T. Van Hecke, G. Vanden Berghe, H. De Meyer, Deferred correction with mono-implicit Runge–Kutta methods for first order IVPs, J. Comput. Appl. Math. 111 (1999) 37–47.
[16] M. Van Daele, J. Cash, Superconvergent deferred correction methods for first order systems of nonlinear two-point boundary value problems, SIAM J. Sci. Comput., to be published.
[17] T. Van Hecke, M. Van Daele, G. Vanden Berghe, H. De Meyer, P-stable Mono-implicit Runge–Kutta–Nyström modifications of the Numerov Method, in: Proceedings of the First International Workshop WNAA'96, Rousse, Bulgaria, Lecture Notes in Computer Science, vol. 1996, Springer, Berlin, 1997, pp. 537–545.
[18] T. Van Hecke, Mono-impliciete Runge–Kutta(–Nyström) methodes, Ph.D. Thesis, University of Ghent, 1998.