Summation by parts formula for noncentered finite differences

Author(s):
Bodenmann, R.

Publication Date:
1995

Permanent Link:
https://doi.org/10.3929/ethz-a-004284331

Rights / License:
In Copyright - Non-Commercial Use Permitted
Summation by Parts Formula for Noncentered Finite Differences

R. Bodenmann

Research Report No. 95-07
September 1995
Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland
Summation by Parts Formula for Noncentered Finite Differences

R. Bodenmann
Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Research Report No. 95-07    September 1995

ABSTRACT. Boundary conditions for general high order finite difference approximations to $\frac{d}{dx}$ are constructed. It is proved that it is always possible to find boundary conditions and a weighted norm in such a way that a summation by parts formula holds.

Keywords: high order finite differences, summation by parts formula, energy method, linear stability, initial-boundary value problem
1. Introduction

For analysing numerical methods applied to partial differential equations, a priori estimates of the growth rate of the solution are necessary. Such stability estimates ensure that small discretisation errors cannot grow arbitrarily.

Analysing convergence behaviour of finite difference approximations to nonlinear problems, it turns out that stability inequalities of the linearisation of the methods have to be proved [1], [8]. Since schemes which are obtained by linearisation have variable coefficients, the energy method turns out to be one of the most effective devices to derive bounds of the growth rate [2], [5], [7].

One of the most important components of the energy method in the continuous case is the integration by parts formula. To get an energy estimate for the discretised problem, the integration by parts formula has to be replaced by its discrete analog, the summation by parts formula. This is accomplished by constructing the difference operator including numerical boundary conditions in an appropriate manner.

Kreiss and Scherer proved that in general new scalar products and norms have to be introduced to make possible the existence of difference approximations which fulfil a summation by parts formula. They constructed weighted diagonal or even more general types of norms [3] and proved the existence of difference approximations to \( \frac{d}{dx} \) of almost antisymmetric type such that the summation by parts formula holds. These results, which are reviewed in the next section, can also be found in [9] together with some examples.

In Section 3, the theory is developed for more general difference operators, which do not need to be antisymmetric. It is pointed out in Section 4 that the results from Section 3 are optimal, even though they are weaker than the ones which can be proved in the antisymmetric case.

An algorithm to compute proper boundary conditions for the diagonal norm case is presented in Section 5 and some of its results are displayed in Section 6.

2. Review of the Antisymmetric Case

The goal is to construct difference approximations \( Q \) to \( \frac{d}{dx} \) on the interval \([0,1]\), which are of arbitrarily high order in the interior as well as on the boundaries of the interval and which fulfil a summation by parts formula:

\[
(u, Qv)_h = -(Qu, v)_h + u_J v_J - u_0 v_0
\]

for any two grid functions \( u \) and \( v \) and for a certain discrete scalar product \((\cdot, \cdot)_h\).

Obviously it is enough to investigate the halfline problem. So let us consider the halfline \([0, \infty)\), which we divide into intervals of length \( h > 0 \). Denote the gridpoints by \( x_j = jh, \ j = 0,1, \ldots \) and let \( u_j = u(x_j) \) be a scalar grid function in \( l_2 \), i.e.
A discrete scalar product and a norm are defined by

\[(u, v)_h = <u^l, Hv^l>_h + h \sum_{j=r}^\infty u_j v_j\]

(1)

\[\|u\|^2_h = (u, u)_h,\]

Here \(u^l = (u_0, u_1, \ldots, u_{r-1})^T\) and \(H = H^T > 0\) is a symmetric, positive definite \((r \times r)\)-matrix. We want to construct matrices \(H\) and a difference approximation \(Q\) to \(\frac{d}{dx}\), such that the summation by parts formula

\[(u, Qv)_h = -(Qu, v)_h - u_0 v_0\]

(2)

holds.

The following theorems are proved in [3], [6], and [9].

**Theorem 2.1 (Diagonal Norms).** There exist scalar products (1), with \(H\) diagonal, and difference operators \(Q\) of accuracy \(p\) at the boundary and \(2p\) in the interior, \(1 \leq p \leq 4\), such that the summation by parts formula (2) holds.

**Theorem 2.2 (Full Norms).** There exist scalar products (1) and difference operators \(Q\) of accuracy \(2p - 1\) at the boundary and \(2p\) in the interior, \(p > 0\), such that the summation by parts formula (2) holds.

The proofs are constructive and examples are computed in [9]. The simplest example for the diagonal norm case is given by

\[\frac{1}{h}Q = \begin{pmatrix}
-1 & 1 \\
-0.5 & 0 & 0.5 \\
-0.5 & 0 & 0.5 \\
& \ddots & \ddots & \ddots
\end{pmatrix}\]

and \(H = \text{diag}(0.5, 1, 1, \ldots)\).

### 3. Noncentered Differences

We now want to construct pairs of finite difference operators \((Q_-, Q_+)\) and scalar products (1) such that the summation by parts formula

\[(u, Q_-v)_h = -(Q_+u, v)_h - u_0 v_0\]

holds for any scalar grid functions \(u\) and \(v\).
Since $u$ and $v$ are considered as infinite vectors, $Q_\mp$ are represented as infinite matrices. Assume that $Q_\mp$ have the form

\[ hQ_\mp = \begin{pmatrix} \left(Q_{\mp}^1\right)_{11} & \left(Q_{\mp}^1\right)_{12} \\ C_\mp^T & D_\mp \end{pmatrix}, \]

with

\[ \left(Q_{\mp}^1\right)_{11} = \begin{pmatrix} q_{0,0} & q_{0,1} & \cdots & q_{0,r-1} \\ \vdots & \vdots & & \vdots \\ q_{r-1,0} & q_{r-1,1} & \cdots & q_{r-1,r-1} \end{pmatrix}, \]

\[ \left(Q_{\mp}^1\right)_{12} = \begin{pmatrix} q_{0,r} & \cdots & q_{0,m} & 0 & \cdots \\ \vdots & \vdots & & \vdots \end{pmatrix}, \]

\[ \begin{pmatrix} a_0 & a_1 & \cdots & a_R & 0 & 0 & \cdots & 0 & 0 & \cdots \\ a_{-1} & a_0 & a_1 & \cdots & a_R & 0 & \cdots & 0 & 0 & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \cdots & \vdots & \vdots & \ddots \\ a_{-L} & a_{-1} & a_0 & a_1 & \cdots & a_R & 0 & \cdots & 0 & \cdots \\ 0 & a_{-L} & a_{-1} & a_0 & a_1 & \cdots & a_R & 0 & \cdots & \cdots \end{pmatrix} \]

\[ D_- = \begin{pmatrix} \alpha_{-L} & 0 & 0 & \cdots & 0 \\ \alpha_{-L+1} & \alpha_{-L} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \alpha_{-2} & \alpha_{-3} & \cdots & \alpha_{-L} & 0 \\ \alpha_{-1} & \alpha_{-2} & \cdots & \alpha_{-L+1} & \alpha_{-L} \end{pmatrix}, \]

\[ D_+ = -D_-^T, \]

\[ C_\mp = \begin{pmatrix} C_\mp^0 & 0 & \cdots \\ C_\mp^0 & 0 & \cdots \end{pmatrix}, \]

\[ \tilde{C}_- = \begin{pmatrix} a_{-L} & 0 & 0 & \cdots & 0 \\ a_{-L+1} & a_{-L} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ a_{-2} & a_{-3} & \cdots & a_{-L} & 0 \\ a_{-1} & a_{-2} & \cdots & a_{-L+1} & a_{-L} \end{pmatrix}, \]

\[ \tilde{C}_+ = -\begin{pmatrix} a_R & 0 & 0 & \cdots & 0 \\ a_{R-1} & a_R & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ a_2 & a_3 & \cdots & a_R & 0 \\ a_1 & a_2 & \cdots & a_{R-1} & a_R \end{pmatrix}, \]

$C_-^0$ and $C_+^0$ are $(r - L) \times L$ and $(r - R) \times R$ matrices respectively with only zeroes.
Theorem 3.1. The operators $Q_-$ and $Q_+$ satisfy the relation (3) if and only if they can be written as

\[
\begin{align*}
\h Q_- &= \left( \frac{H^{-1}A}{C^-_T} \right), \\
\h Q_+ &= \left( \frac{H^{-1}B}{D^-_T} \right),
\end{align*}
\]

where $A$ and $B$ are $(r \times r)$-matrices of the form $A = A_1 + A_2$, $B = B_1 + B_2$ with

\[
A_1 = \begin{pmatrix}
a_{11} & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0
\end{pmatrix}, \quad B_1 = \begin{pmatrix}
b_{11} & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0
\end{pmatrix}, \quad a_{11} + b_{11} = -1,
\]

and $A_2 + B_2^T = 0$.

**Proof.** We can formally write (3) as

\[
-u_0v_0 = <u^I, H(Q_-)_{11}v^I> + <u^I, H(Q_-)_{12}v^{II}> + <u^{II}, C^-_Tv^I> + <u^{II}, D^-Tv^{II}> + <(Q_+)_{11}u^I, Hv^I> + <(Q_+)_{12}u^{II}, Hv^I> + <C^T_+u^I, v^{II}> + <D^+_u u^{II}, v^{II}>
\]

and by making use of the symmetry of $H$

\[
-u_0v_0 = <u^I, H(Q_+)_1v^I> + <u^I, H(Q_+)_{12}v^{II}> + <u^{II}, C^+_Tv^I> + <u^{II}, D^+_Tv^{II}> + <(Q_-)_{11}u^I, Hv^I> + <(Q_-)_{12}u^{II}, Hv^I> + <C^T_+u^I, v^{II}> + <D^-_u u^{II}, v^{II}>,
\]

where $u^I = (u_0, u_1, \ldots, u_{r-1})^T$ and $u^{II} = (u_r, u_{r+1}, \ldots)^T$. For $D_\pi$ we have the relation $<u^{II}, D^-Tv^{II}> + <D^+_u u^{II}, v^{II}> = 0$. Letting $v^I = 0$, we get

\[
<u^I, H(Q_-)_{12}v^{II}> + <C^T_+u^I, v^{II}> = <u^I, (H(Q_-)_{12} + C^+_+)v^{II}> = 0
\]

and

\[
<u^I, H(Q_+)_1v^{II}> + <C^T_+u^I, v^{II}> = <u^I, (H(Q_+)_{12} + C^-_+)v^{II}> = 0
\]

for all vectors $u^I$ and $v^{II}$. Therefore we have

\[
(Q_-)_{12} = -H^{-1}C^+_+ \quad \text{and} \quad (Q_+)_{12} = -H^{-1}C^-_+.
\]

Only the relation

\[
<u^I, (H(Q_-)_{11} + (Q^+_T)_{11}H)v^I> = -u_0v_0
\]

remains. Therefore $H(Q_-)_{11} = A$ and $H(Q_+)_{11} = B$ must be of the required form. \(\Box\)
Let $h = 1$ and denote by

$$w_l = \begin{pmatrix} e_l & f_l \end{pmatrix}, \quad e_l = (-1)^l \begin{pmatrix} r^l \\ (r-1)^l \\ \vdots \\ 1 \end{pmatrix}, \quad f_l = \begin{pmatrix} 0^l \\ 1 \end{pmatrix}, \quad l = 0, 1, \ldots$$

the discretisation of $(x - r)^l$, with the conventions $0^0 = 1$, and $e_{-1} = 0$.

The following lemma characterises the accuracy of $Q_x$ in the interior.

**Lemma 3.2.** The operators $h^{-1} \langle C^T, D\rangle$ approximate $\frac{d}{dx}$ with order $q$ at the points $x_j$, $j = r, r + 1, \ldots$ if and only if

$$\sum_{\nu = -L}^{R} \alpha_{\nu} \nu^p = \begin{cases} 1, & p = 1, \\ 0, & p = 0, 2, 3, \ldots, q. \end{cases}$$

(5)

Accuracy conditions for $Q_x$ at boundary points are given by the next lemma.

**Lemma 3.3.** The operators $h^{-1} \langle (Q_x)_{11}, (Q_x)_{12} \rangle$ approximate $\frac{d}{dx}$ with order $\tau$ at the points $x_j$, $j = 0, 1, \ldots, r - 1$ if and only if

$$le_{l-1} = (Q_x)_{11} e_l + (Q_x)_{12} f_l, \quad l = 0, 1, \ldots, \tau.$$

(6)

We now want to calculate the coefficients of $Q_x$ at the boundary points. To do so, we first rewrite (6). By (4) we get

$$A_2 e_i = g_i^-, \quad g_i^- = lHe_{l-1} - A_1 e_i + C_+ f_i, \quad l = 0, 1, \ldots, \tau,$$

$$B_2 e_i = g_i^+, \quad g_i^+ = lHe_{l-1} - B_1 e_i + C_- f_i, \quad l = 0, 1, \ldots, \tau.$$

Since $A_2 + B_2^T = 0$, the following compatibility conditions must hold

$$< e_i, g_m^- > + < e_m, g_i^+ > = 0, \quad 0 \leq m, l \leq \tau.$$  

(8)

If these conditions are fulfilled, the system (7) can be resolved as it is expressed in Lemma 3.4.

**Lemma 3.4.** Assume that $r \geq \tau + 1$ and that the relations (8) hold. Then there are matrices $A_2$ and $B_2$ with $A_2 + B_2^T = 0$ such that (7) is valid.

We have to show that there exists a symmetric, positive definite matrix $H$ such that the compatibility conditions (8) hold. By (7) we know that such a matrix $H$ has to fulfil

$$m < e_i, He_{m-1} > + l < e_m, He_{l-1} > = M_{i,m}, \quad 0 \leq l, m \leq \tau,$$

(9)
where
\[
M_{l,m} = <\epsilon_i, A_1 \epsilon_m > - <\epsilon_i, C_+ f_m > + <\epsilon_m, B_1 \epsilon_i > - <\epsilon_m, C_- f_i >
\]
\[
= <\epsilon_i, (A_1 + B_1) \epsilon_m > - <\epsilon_i, C_+ f_m > - <\epsilon_m, C_- f_i >.
\]

Lemma 3.5. \(M_{l,m}\) can be written as
\[
M_{l,m} = -(-r)^{l+m} + J_{l,l+m}
\]
\[
J_{l,\sigma} = - \sum_{\nu=1}^{L} \alpha_{-\nu} \sum_{\mu=0}^{\nu-1} (\mu - \nu)^{\sigma-1} \mu^l
\]
\[
+ \sum_{\nu=1}^{R} \alpha_{\nu} \sum_{\mu=0}^{\nu-1} (\mu - \nu)^{\sigma-1} \mu^{\sigma-l}, \quad \sigma = l + m.
\]

Proof. The scalar products can be calculated as
\[
<\epsilon_i, (A_1 + B_1) \epsilon_m > = (-1)^{l+m+1} r^{l+m},
\]
\[
<\epsilon_m, C_- f_i > = (-1)^m \sum_{\nu=1}^{L} \alpha_{-\nu} \sum_{\mu=0}^{\nu-1} (\mu - \nu)^m \mu^l
\]
\[
<\epsilon_i, C_+ f_m > = (-1)^l \sum_{\nu=1}^{R} \alpha_{\nu} \sum_{\mu=0}^{\nu-1} (\mu - \nu)^l \mu^{\sigma-l}
\]
The proof is concluded by using these expressions in (9) and by introducing \(\sigma = l + m\).

If we introduce the notation \(\rho_{l,m} = <\epsilon_i, H \epsilon_m >\), then we can write (9) as
\[
m \rho_{l,m-1} + l \rho_{m,l-1} = M_{l,m}, \quad 0 \leq l, m \leq \tau.
\]

Here \(\rho_{l,-1} = 0\) by the convention for \(\epsilon_{-1}\) and \(\rho_{l,m} = \rho_{m,l}\) by the symmetry of \(H\). From the solution of this system in the \(\rho_{l,m}\) we obtain the elements of \(H\). The matrix \(H\) has to be positive definite in order to be used as a norm matrix. The equivalence between this condition and the positive definiteness of the matrix defined by the \(\rho_{l,m}\) is stated by Lemma 3.9 below.

In the next step we derive conditions for \(M_{l,m}\) such that the system (11) has a solution, and therefore the compatibility conditions hold. These conditions are resumed in Lemma 3.7.

For \(l = m = 0\) we obtain from (11) and (10)
\[
0 = 0 \rho_{0,-1} + 0 \rho_{0,-1} = M_{0,0} = -1 + \sum_{\nu=-L}^{R} \nu \alpha_{\nu}, \quad \text{i.e.} \quad \sum_{\nu=-L}^{R} \nu \alpha_{\nu} = 1.
\]

By Lemma 3.2 this is fulfilled if the approximation in the interior is at least of order 1.

Since the left hand side of (11) is symmetric, the same has to be true for the right hand side. The following lemma gives conditions for \(M_{l,m}\) to be symmetric.
Lemma 3.6. If the approximation in the interior is at least of order \(2\sigma\), then
\[ M_{l,m} = M_{m,l}, \quad 0 \leq l, m \leq \tau. \]

Proof. We have to show that \(J_{l,\sigma} = J_{m,\sigma}\) for \(0 \leq l, m \leq \tau\), \(l + m = \sigma\). We introduce the following abbreviations:
\[ N_{l,\sigma}(\nu) = \sum_{\mu=0}^{\nu-1} \mu^\sigma(-\nu)^l + \sum_{\mu=1}^{\nu} (\mu - \nu)^l \mu^{\sigma-1}, \quad 0 \leq l \leq \sigma \]
and
\[ \widetilde{N}_{l,\sigma}(\nu) = \sum_{\mu=0}^{\nu-1} \mu^\sigma(-\nu)^l - \sum_{\mu=1}^{\nu} (\mu - \nu)^l \mu^{\sigma-1}, \quad 0 \leq l \leq \sigma. \]

With
\[ \sum_{\mu=0}^{\nu-1} \mu^l (-\nu)^{\sigma-l} = (-1)^\sigma \sum_{\mu=1}^{\nu} (\mu - \nu)^l \mu^{\sigma-1} \]
and letting \(s = \max(L, R)\), \(\alpha_+ = 0\), \(\nu > L\), and \(\alpha_+ = 0\), \(\nu > R\), we get
\begin{equation}
J_{l,\sigma} - J_{m,\sigma} = \begin{cases} 
\sum_{\nu=1}^{s} (\alpha_+ + \alpha_-) \widetilde{N}_{l,\sigma}(\nu), & \text{if } \sigma \text{ even}, \\
\sum_{\nu=1}^{s} (\alpha_+ + \alpha_-) N_{l,\sigma}(\nu), & \text{if } \sigma \text{ odd}.
\end{cases}
\end{equation}

In [3] and [9] it is proved that \(N_{l,\sigma}(\nu)\) is of the following form:
\[ N_{l,\sigma}(\nu) = \gamma_1^{(l,\sigma)} \nu^{\sigma+1} + \gamma_3^{(l,\sigma)} \nu^{\sigma-1} + \ldots, \quad \gamma_1^{(l,\sigma)} \neq 0. \]
For \(\widetilde{N}_{l,\sigma}(\nu)\) we get
\[ \widetilde{N}_{l,\sigma}(\nu) = \sigma^\sigma(-\nu)^l - \sigma^l \nu^{\sigma-1} = \begin{cases} 
0, & l \neq 0 \text{ and } l \neq \sigma, \\
\nu^\sigma, & l = 0, \\
(-\nu)^\sigma, & l = \sigma.
\end{cases} \]

The proof is completed by using these expressions for \(N_{l,\sigma}(\nu)\) and \(\widetilde{N}_{l,\sigma}(\nu)\) in (13) and by Lemma 3.2. \(\square\)

We can proceed now as in [3] and [9] for the antisymmetric case. Suppose that the conditions of Lemma 3.6 are fulfilled, then \(M_{l,m} = M_{m,l}\), \(l, m \leq \tau\) and therefore (11) has to be considered only for \(l \leq m \leq \tau\). For \(l = 0\) we get the following condition from (11)
\[ \rho_{0,m-1} = \frac{1}{m} M_{0,m}, \quad m = 1, 2, \ldots, \tau. \]

If \(l > 0\), then \(\rho_{l,m-1}\) can be calculated explicitly
\[ \rho_{l,m-1} = \frac{1}{m} M_{l,m} - \frac{l}{m} \rho_{l-1,m}. \]
Now (11) can be written recursively as

\[
\rho_{l,m-1} = \frac{\rho_{l,m}}{m} M_{l,m} + \sum_{k=1}^{\alpha} (-1)^k \frac{l!(m-1)!}{(l-k)!(m+k)!} M_{l-k,m+k} \\
+ (-1)^{\alpha+1} \frac{l!(m-1)!}{(l-\alpha-1)!(m+\alpha)!} \rho_{l-\alpha-1,m+\alpha},
\]

where \( \alpha = \min(l-1, \tau - m) \).

If \( l - 1 < \tau - m \), i.e. \( l + m \leq \tau \), then \( \alpha = l - 1 \) and the following condition is obtained

\[
\rho_{l-\alpha-1,m+\alpha} = \rho_{0,l+m-1} = \frac{1}{l+m} M_{0,l+m}.
\]

Thus (12) and (14) imply that \( \rho_{l,m-1} \) is completely determined by the \( M_{l,m} \) provided \( l + m \leq \tau \).

If \( l + m > \tau \) then \( \alpha = \tau - m \) and

\[
\rho_{l-\alpha-1,m+\alpha} = \rho_{l+m-\tau,\tau} = \rho_{\nu,\tau}, \quad \nu = l + m - 1 - \tau.
\]

There are no further relations that \( \rho_{\nu,\tau} \) needs to satisfy.

If we use equation (14) with \( l = n, m = n \) and \( l = n - 1, m = n + 1 \), for \( n \leq \tau \), then we obtain representations for \( \rho_{n,n-1} \) and \( \rho_{n-1,n} \) but by (11) \( \rho_{n,n-1} = \rho_{n-1,n} \). If \( l + m = 2n > \tau \), then by (16) these two relations determine \( \rho_{\nu,\tau} \), \( \nu = 2n - 1 - \tau \) and no conditions for the \( M_{l,m} \) result. If \( l + m = 2n \leq \tau \) we get from (14) and (15)

\[
\sum_{k=0}^{n} (-1)^k \frac{n!(n-1)!}{(n+k)!(n-k)!} M_{n-k,n+k} = \rho_{n,n-1} = \rho_{n-1,n}
\]

\[
= \sum_{k=1}^{n} (-1)^{k-1} \frac{(n-1)!n!}{(n+k)!(n-k)!} M_{n-k,n+k}
\]

This relation can be written as

\[
\frac{1}{2n} M_{n,n} + \sum_{k=1}^{n} (-1)^{k} \frac{(n-1)!n!}{(n+k)!(n-k)!} M_{n-k,n+k} = 0
\]

for \( n = 1, 2, \ldots \), with \( 2n \leq \tau \).

We obtain

**Lemma 3.7.** If the difference approximation in the interior is accurate at least to the order of \( 2\tau \) and the \( M_{l,m} \) satisfy the relations (17), then the system (11) has a solution. \( \square \)
If the system (11) has a solution, then the \( p_{l,m} \) are determined for \( 0 \leq l, m \leq \tau \) if one specifies those \( p_{v,\tau}, \; v = 0, 1, \ldots, \tau \), which are not determined by the system. They can be used to define a symmetric matrix \( R_\tau = (p_{l,m})_{0 \leq l, m \leq \tau} \). The following lemma states that it is always possible to choose the parameters \( p_{v,\tau} \) so that the matrix \( R_\tau \) becomes positive definite.

**Lemma 3.8.** If \( R_{\tau-1} = (p_{l,m})_{0 \leq l, m \leq \tau-1} = R_{\tau-1}^T \) is positive definite then one can choose \( p_{\tau,\tau} \) such that \( R_\tau \) is symmetric and positive definite as well, independently of the values of \( p_{v,\tau}, \; v = 0, 1, \ldots, \tau - 1 \).

We now have to determine a positive definite \((r \times r)\)-matrix \( H \) such that

\[
< e_i, H e_m > = p_{l,m}, \quad 0 \leq l, m \leq \tau.
\]

In fact the following lemma holds.

**Lemma 3.9.** If \( r \geq \tau + 1 \) and the matrix \( R_\tau \) is positive definite, then there are \( H = H^T > 0 \) such that (18) holds. In particular if \( r = \tau + 1 \), then \( H \) is uniquely defined by

\[
E^T H E = R_\tau, \quad E = (e_0, e_1, \ldots, e_{r-1}).
\]

Now we can state the main result.

**Theorem 3.10.** For every \( q = 2, 3, \ldots \) and for a sufficiently large \( r \) there is a symmetric, positive definite matrix \( H \), a scalar product of the form (1), and approximations \( Q_\tau \) of \( \frac{d^q}{dx^q} \) that are accurate with order \( \tau = \lfloor \frac{q}{2} \rfloor \) for \( x = x_j, \; j = 0, 1, \ldots, r - 1 \), and accurate with order \( q \) for \( x = x_j, \; j \geq r \) such that (3) holds.

To prove stability of difference approximations to systems of partial differential equations, it is more convenient to have so-called restricted full norms [4], defined by a symmetric and positive definite matrix of the form

\[
H = \begin{pmatrix}
\lambda_0 & 0 & \cdots & 0 \\
0 & h_{1,1} & \cdots & h_{1,r-1} \\
\vdots & \vdots & \ddots & \vdots \\
0 & h_{r-1,1} & \cdots & h_{r-1,r-1}
\end{pmatrix}.
\]

The following lemma holds.

**Lemma 3.11.** If \( R_\tau \) is positive definite, then one can choose \( H \) in the form (19) such that (18) holds. In general one has to choose \( r \geq \tau + 2 \).

Finally we can rewrite Theorem 3.10 for restricted full norms.
Theorem 3.12. For every $q = 2, 3, \ldots$ there is a symmetric, positive definite matrix $H$ of the form (19), a scalar product of the form (1), and approximations $Q_\pm$ of $\frac{d}{dx}$, which are accurate with order $\tau = \left\lfloor \frac{q}{2} \right\rfloor$ for $x = x_j$, $j = 0, 1, \ldots, r - 1$, and accurate with order $q$ for $x = x_j$, $j \geq r$ such that (3) holds.

The proofs of the above lemmata and theorems exactly follow those in [3] and in [9].

4. Concluding Remarks

Comparing the results of the antisymmetric case in Theorem 2.2 with those for noncentered schemes – Theorem 3.10 and Theorem 3.12 – we see that for the antisymmetric case we only need a scheme of order $\tau + 1$ in the interior to get boundary conditions of order $\tau$, whereas for noncentered schemes, we need order $2\tau$ (respectively $2\tau + 1$) in the interior to get order $\tau$ at the boundary. The reason is that for an antisymmetric scheme we always have $a_\nu + a_{-\nu} = 0$ by definition. Therefore (13) reads $J_{i,\sigma} - J_{m,\sigma} = 0$ and the statement of Lemma 3.6 is always true without any accuracy assumptions on the interior scheme and the only interior accuracy assumptions have to be made in order to fulfil the conditions (17), cf. [9].

5. Algorithms

We now want to present algorithms to compute the norm matrix $H$ and the boundary conditions $A$ respectively $B$. For the sake of simplicity we assume $H$ to be diagonal. From Theorem 2.1 we know that in the antisymmetric case such norms exist for low order schemes. It is still possible to construct diagonal norms for higher order schemes but then one has to take a scheme with order $q \geq 2\tau$ in the interior to get boundary conditions of order $\tau$. As we have seen in Section 3, the theory for antisymmetric schemes can be generalised to noncentered schemes and similar results can be proved. That is why one can expect results for noncentered schemes that are similar to Theorem 2.1, but we should reckon with the same problems with schemes of higher order.

The following notation is used in the algorithms, which are written in Maple:

$$a_{ij} = (Q_-)_{ij}, \quad b_{ij} = (Q_+)_{ij}, \quad 1 \leq i, j \leq r + s, \quad s = \max(L, R).$$

The conditions $A_{11} + B_{11} = -1$ and $A_2 + B_2^T = 0$ are designated as ‘structure assumptions’ and the accuracy conditions are the ones from formula (5). The matrices $C_\mp$ are called ‘interfaces’. This is where the interior scheme enters the computations.
5.1. Third order scheme. The first algorithm computes boundary conditions of first order for the scheme with \( \alpha_{-2} = \frac{5}{6}, \, \alpha_{-1} = -1, \, \alpha_0 = \frac{5}{2}, \, \alpha_1 = \frac{1}{3} \), which is accurate of third order.

\[
\begin{align*}
& r := 2; \\
& s := 2; \\
& \textbf{# norm} \\
& \text{normout} := \{ \text{seq (lambda[i]=1, i=r+1..r+s) } \}; \\
& \text{eqns} := \text{normout}; \\
& \textbf{# interfaces} \\
& \text{inta} := \{ \text{seq (a[i,j]=0,j=1..r-s), i=r+1..r+s) } \}; \\
& \text{intb} := \{ \text{seq (b[i,j]=0,j=1..r-s), i=r+1..r+s) } \}; \\
& \text{eqns} := \text{eqns union inta union intb}; \\
& \textbf{# essential interfaces} \\
& \text{intea} := \{ a[r+1,r-1]=1/6, a[r+1,r]=-1, a[r+2,r-1]=0, a[r+2,r]=1/6 \}; \\
& \text{inteb} := \{ b[r+1,r-1]=0, b[r+1,r]=1/3, b[r+2,r-1]=0, b[r+2,r]=0 \}; \\
& \text{eqns} := \text{eqns union intea union inteb}; \\
& \textbf{# structure assumptions} \\
& \text{bdterm} := \{ \text{lambda[i]*a[i,j]=lambda[j]*b[i,j], i=1..r, j=2..r+s) } \}; \\
& \text{adj1} := \{ \text{seq( lambda[i]*a[i,j]=-lambda[j]*b[i,j], i=1..r, j=2..r+s) } \}; \\
& \text{adj2} := \{ \text{seq( lambda[i]*a[i,j]=-lambda[j]*b[i,j], i=2..r+s, j=1..r) } \}; \\
& \text{eqns} := \text{eqns union bdterm union adj1 union adj2}; \\
& \textbf{# accuracy conditions} \\
& \text{ord0a} := \{ \text{seq (value(Sum(a[i,j], l=1..r+s)=0), i=1..r) } \}; \\
& \text{ord0b} := \{ \text{seq (value(Sum(b[i,j], l=1..r+s)=0), i=1..r) } \}; \\
& \text{ord1a} := \{ \text{seq (value(Sum((l-1)*a[i,j], l=1..r+s)=1), i=1..r) } \}; \\
& \text{ord1b} := \{ \text{seq (value(Sum((l-1)*b[i,j], l=1..r+s)=1), i=1..r) } \}; \\
& \text{eqns} := \text{eqns union ord0a union ord0b union ord1a union ord1b}; \\
& \text{solve(eqns)}; \\
\end{align*}
\]

5.2. Fourth order scheme. Now we are going to take a fourth order scheme in the interior in order to get boundary conditions of second order. The interior scheme is defined by \( \alpha_{-3} = -\frac{1}{12}, \, \alpha_{-2} = \frac{1}{2}, \, \alpha_{-1} = -\frac{2}{3}, \, \alpha_0 = \frac{5}{6}, \, \alpha_1 = \frac{1}{4} \). Therefore we have to change the first two lines in the algorithm to

\[
\begin{align*}
& r := 4; \\
& s := 3;
\end{align*}
\]

The interfaces now read
# Interfaces
inta := { seq (a[i,j]=0,j=1..r-s), i=r+1..r+s };
intb := { seq (b[i,j]=0,j=1..r-s), i=r+1..r+s };

# Essential Interfaces
intea := { a[r+1,r-2]=-1/12, a[r+1,r-1]=1/2, a[r+1,r]=3/2, a[r+2,r-2]=0,
a[r+2,r-1]=-1/12, a[r+2,r]=1/2, a[r+3,r-2]=0, a[r+3,r-1]=0, a[r+3,r]=-1/12 };
inteb := { b[r+1,r-2]=0, b[r+1,r-1]=0, b[r+1,r]=-1/4,
seq (b[r+i,r-i]=0,j=0..s-1), i=2..s };

Finally the accuracy conditions are changed to

# Accuracy conditions
ord0a := { seq (value(Sum(a[i,l], l=1..r+s)=0), i=1..r) };
ord0b := { seq (value(Sum(b[i,l], l=1..r+s)=0), i=1..r) };
ord1a := { seq (value(Sum((1-l)*a[i,l], l=1..r+s)=1), i=1..r) };
ord1b := { seq (value(Sum((1-l)*b[i,l], l=1..r+s)=1), i=1..r) };
ord2a := { seq (value(Sum((1-l)^2*a[i,l], l=1..r+s)=0), i=1..r) };
ord2b := { seq (value(Sum((1-l)^2*b[i,l], l=1..r+s)=0), i=1..r) };

6. Numerical Results

For the third order scheme from Section 5 we get the following normmatrix $H$ and
boundary conditions $A$ and $B$, which are of first order

$$H = \text{diag}(\frac{5}{12}, \frac{13}{12}), \quad H^{−1}A = \begin{pmatrix} -1 & 1 \\ -\frac{9}{13} & \frac{5}{13} \end{pmatrix}, \quad H^{−1}B = \begin{pmatrix} -\frac{7}{5} & \frac{2}{5} \\ -\frac{5}{13} & -\frac{5}{13} \end{pmatrix}.$$  

The whole difference approximations $Q_-$ and $Q_+$ are

$$hQ_- = \begin{pmatrix} -1 & 1 \\ -\frac{9}{13} & \frac{5}{13} & \frac{4}{13} \\ \frac{1}{6} & -1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{6} & -1 & \frac{1}{2} & \frac{1}{3} \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}. $$
and

\[ hQ_+ = \begin{pmatrix}
\frac{7}{6} & \frac{2}{3} & -\frac{2}{9} \\
\frac{5}{13} & \frac{5}{13} & \frac{12}{13} & -\frac{2}{13} \\
-\frac{1}{3} & -\frac{1}{2} & 1 & -\frac{1}{6} \\
-\frac{1}{3} & -\frac{1}{2} & 1 & -\frac{1}{6} \\
\ddots & \ddots & \ddots & \ddots
\end{pmatrix}. \]

The algorithm for the fourth order difference approximations from Section 5 gives a one parameter family of pairs of difference operators. For stability reasons, we choose the free parameter in such a way that the symmetric part \( Q_- - Q_+ \) of the difference approximations becomes positive semidefinite. Hence, we get the following norm and finite difference approximations

\[ H = \text{diag}(\frac{49}{144}, \frac{61}{48}, \frac{41}{48}, \frac{149}{144}), \]

\[ hQ_- = \begin{pmatrix}
\frac{223}{20} & \frac{3}{2} & \frac{5}{6} & \frac{1}{4} & 0 & 0 \\
-\frac{7}{20} & \frac{1}{2} & -\frac{3}{2} & \frac{5}{6} & \frac{1}{4} & 0 \\
-\frac{1}{12} & \frac{1}{2} & -\frac{3}{2} & \frac{5}{6} & \frac{1}{4} & 0 \\
\ddots & \ddots & \ddots & \ddots & \ddots & \ddots
\end{pmatrix}, \]

\[ hQ_+ = \begin{pmatrix}
\frac{78}{49} & \frac{223}{61} & \frac{38}{122} & \frac{2}{123} \\
\frac{151}{123} & \frac{20}{122} & \frac{117}{246} & \frac{4}{149} \\
\frac{2}{123} & \frac{20}{122} & \frac{407}{41} & \frac{4}{41} \\
\frac{15}{208} & \frac{20}{149} & \frac{47}{288} & \frac{120}{149} & \frac{72}{149} & \frac{12}{149} \\
0 & 0 & -\frac{1}{4} & -\frac{5}{6} & \frac{3}{2} & -\frac{1}{12} & \frac{1}{12} \\
0 & 0 & -\frac{1}{4} & -\frac{5}{6} & \frac{3}{2} & -\frac{1}{12} & \frac{1}{12} \\
\ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots
\end{pmatrix}. \]
References


