Banded, stable, skew-symmetric differentiation matrices of high order

Ernst Hairer Arieh Iserles

November 9, 2015

Abstract

Differentiation matrices play an important role in the space discretisation of partial differential equations. The present work considers grids on a finite interval and treats homogeneous Dirichlet boundary conditions. Differentiation matrices of order 3 and 4 are derived that are banded, stable, and skew symmetric. To achieve these desirable properties, non-equidistant grids have to be considered. The impact of the use of skew-symmetric differentiation matrices is illustrated at examples of advection and diffusion equations.

1 Introduction

sect:intro

This paper deals with the approximation of the first derivative $u'(\xi)$ on a grid over a finite interval

$$a = \xi_0 < \xi_1 < \cdots < \xi_n < \xi_{n+1} = b. \qquad (1)$$ grid-xi

To keep the notation simple we suppress the dependence of ξ_k on n, the number of internal grid points. We always assume that the grid is dense in $[a, b]$, meaning that $\max_{k=0,\ldots,n} |\xi_{k+1} - \xi_k| = \mathcal{O}(n^{-1})$. In view of applications to the space discretisation of partial differential equations with Dirichlet boundary conditions, our main interest is in functions $u(\xi)$ vanishing at both endpoints. We assume that the derivative approximation is obtained by a linear combination of function values over the same grid. This leads to the consideration of

$$u'(\xi_m) \approx \sum_{k=1}^{n} \widehat{\mathcal{D}}_{m,k} u(\xi_k) \qquad (2)$$ differenxi

for $m = 1, \dots, n$. The matrix $(\widehat{\mathcal{D}}_{m,k})$ is called a *differentiation matrix.* Our interest is in such matrices that have the following features:

- the matrix $(\widehat{\mathcal{D}}_{m,k})_{m,k=1}^n$ is banded, i.e., there exists an integer $r \geq 1$ such that $\widehat{\mathcal{D}}_{m,k} = 0$ for $|m-k| \geq r+1$. This assumption brings an evident computational advantage, and guarantees that the derivative approximation is local.

- the differentiation formula is stable, which means that the entries of the differentiation matrix satisfy $|\widehat{\mathcal{D}}_{m,k}| = \mathcal{O}(n)$.

- the approximation is of order $P \geq 2$. This means that for smooth functions $u(\xi)$ the defect in (2) is $\mathcal{O}(h^P)$, where $h = \max_{k=0}^n |\xi_{k+1} - \xi_k|$. For banded, stable grids this is equivalent to requiring that the formula (2) is exact for all polynomials of degree $\leq P$ that vanish at the endpoints.

- the matrix $(\widehat{\mathcal{D}}_{m,k})_{m,k=1}^n$ is skew symmetric. Since the first derivative is a skew-symmetric operator, this is a natural assumption in the spirit of geometric numerical integration. It has several interesting implications for the discretisation of partial differential equations (see [1]).

For an equidistant grid $\{\xi_k = a + kh\}$ with $h = (b-a)/(n+1)$ the discretisation

$$u'(\xi) \approx \frac{1}{2h}\Big(u(\xi+h) - u(\xi-h)\Big)$$

leads to a differentiation matrix that satisfies all four properties, but which is only of second order. The nonzero entries are $\widehat{\mathcal{D}}_{k,k+1} = 1/(2h)$ and $\widehat{\mathcal{D}}_{k,k-1} = -1/(2h)$. Our focus is on differentiation matrices of order higher than 2.

On an equidistant grid the fourth-order approximation

$$u'(\xi) \approx \frac{1}{12h}\Big(u(\xi-2h) - 8u(\xi-h) + 8u(\xi+h) - u(\xi+2h)\Big) \qquad (3)$$ diff-formula-4

can be used at all grid points with the exception of ξ_1 and ξ_n. For these two grid points one-sided approximations can be used. This, however, destroys the skew-symmetry of the matrix.

The difficulty in the construction of skew-symmetric differentiation matrices lies in the fact that we are dealing with Dirichlet boundary conditions on a finite interval. For equidistant infinite grids on the whole real line or for equidistant finite grids and periodic boundary conditions a differentiation formula like (3) can be used everywhere, retaining skew-symmetry.

Finite difference approximations satisfying a "summation by parts rule" (see [5, 8, 7]) are closely connected to the present work. Their approach consists in using an equidistant grid and a standard differentiation formula (like (3)) in the interior of the interval. Close to the end points, one-sided finite difference approximations are considered such that the differentiation matrix becomes skew-symmetric with respect to a *modified scalar product*. Under the assumption that the modified scalar product is defined by a diagonal (positive definite) matrix, approximations of order τ close to the boundary, and of order 2τ in the interior are proposed for $\tau = 1, \dots, 4$. For scalar products with a full matrix at the end points of the interval, difference approximations with order $\tau = 3$ and $\tau = 5$ at the boundary, and order $\tau + 1$ in the interior have been computed. Related work for arbitrary fixed grids can be found in [6] and [4].

In the present work we do not modify the Euclidean inner product, but we consider a grid that is equidistant with the exception of a few grid points at both ends of the interval. The main advantage of keeping the original Euclidean inner product is that it allows methods that preserve unitarity. On parts where the grid is equidistant, one and the same differentiation formula is used (Section 2). This makes an implementation very efficient. It is known from [1] that skew-symmetry of the differentiation matrix and equidistant grids are not compatible with order higher than 2. Section 3 is devoted to conditions on the grid that are necessary for the existence of higher-order skew-symmetric differentiation matrices. The construction of such matrices is outlined in Section 4 for order 3, in Section 5 for order 4, and a short discussion for higher orders is given in Section 6. We finally present numerical experiments in Section 7. For pure advection equations the absence of skew symmetry can lead to unstable space discretisations, whereas the use of skew-symmetric matrices guarantees stability. For space discretisation of diffusion equations, the use of non-skew-symmetric differentiation matrices may result in linear differential equations with a matrix whose eigenvalues are not all on the negative real line. For skew-symmetric differentiation matrices, like for the original linear differential operator, all eigenvalues stay on the negative real line.

2 Structure of considered differentiation matrices

sect:structure

The differentiation matrix originating in a differentiation formula such as (3) is skew symmetric everywhere with the exception of its left upper and right lower corners. The idea is to modify the coefficients in these parts to achieve skew symmetry and retain high order, stability and the banded structure. It follows from the results of [3, 1] that this is not possible for an equidistant grid. We therefore also have to modify the grid, and we do this only at the endpoints of the interval in question.

2.1 The choice of the grid

sect:grid

We fix positive integers N and L and consider the symmetric grid

$$x_k = \begin{cases} -a_{-k}h, & k = -L, \ldots, -1, \\ kh, & k = 0, 1, \ldots, N, \\ 1 + a_{k-N}h, & k = N+1, \ldots, N+L, \end{cases} \tag{4}$$

grid-a

where $h = 1/N$ and $a_1, \ldots, a_L$ are parameters ($0 < a_1 < \cdots < a_L$), which may depend on h. This grid corresponds to the interval $[-a_L h, 1 + a_L h]$ and has $n = N + 2L - 1$ interior grid points. It is equidistant except for a few subintervals at the endpoints.

2.2 The pattern of the differentiation matrix

sect:pattern

Associated to the grid (4) we consider differentiation matrices $(\mathcal{D}_{m,k})$ yielding approximations

$$u'(x_m) \approx \sum_{k=-L+1}^{N+L-1} \mathcal{D}_{m,k} u(x_k) \tag{5}$$

differentiation

for functions $u(x)$ vanishing at the endpoints x_{-L} and x_{N+L}. For the definition of the matrix $(\mathcal{D}_{m,k})$ we consider a basic differentiation rule for an equidistant grid

$$u'(x) \approx \sum_{k=-R}^{R} \delta_k u(x + kh) \tag{6}$$

basic

satisfying $\delta_0 = 0$ and $\delta_{-k} = -\delta_k$. For example, the differentiation rule (3) has $R = 2$, $\delta_1 = 8/(12h)$ and $\delta_2 = -1/(12h)$. In general,

$$\delta_k = \frac{(-1)^{k-1}}{kh} \frac{R!^2}{(R-k)!(R+k)!}, \qquad k = 1, \ldots, R,$$

gives raise to an $(2R)$-order differentiation rule.

$$\begin{bmatrix}
\circ & \times & \times & \times & & & & & & & & \\
\times & \circ & \times & \times & & & & & & & & \\
\times & \times & \circ & \times & \bullet & & & & & & & \\
\times & \times & \times & \circ & \bullet & \bullet & & & & & & \\
 & & \bullet & \bullet & \circ & \bullet & \bullet & & & & & \\
 & & & \bullet & \bullet & \circ & \bullet & \bullet & & & & \\
 & & & & \bullet & \bullet & \circ & \bullet & \bullet & & & \\
 & & & & & \bullet & \bullet & \circ & \bullet & \bullet & & \\
 & & & & & & \bullet & \bullet & \circ & \times & \times & \times \\
 & & & & & & & \bullet & \times & \circ & \times & \times \\
 & & & & & & & & \times & \times & \circ & \times \\
 & & & & & & & & \times & \times & \times & \circ
\end{bmatrix}$$

Figure 1: Pattern of the differentiation matrix with $L = 3$, $M = 2$, $R = 2$, and $N = 7$. fig:pattern

In addition to the integers N and L of Section 2.1, we fix a positive integer M. We then assume that

$$\mathcal{D}_{m,k} = \delta_{k-m} \tag{7}$$ coef-delta

for indices m, k satisfying $|k - m| \le R$, and $M \le m \le N - M$ or $M \le k \le N - M$. The remaining entries are zero, with the exception of the two $(M + L - 1) \times (M + L - 1)$ matrices on the upper left and lower right corners: $(\mathcal{D}_{m,k})_{m,k=-L+1}^{M-1}$ and $(\mathcal{D}_{m,k})_{m,k=N-M+1}^{N+L-1}$. We assume them to be skew symmetric. Moreover, we assume that the whole matrix is skew persymmetric, so that the lower right sub-matrix is determined by the upper left one via

$$\mathcal{D}_{N-m,N-k} = -\mathcal{D}_{m,k} \qquad \text{for} \quad -L + 1 \le m, k \le M - 1.$$

The whole situation is illustrated in Figure 1 for $L = 3$, $M = 2$, $R = 2$, and $N = 7$. The symbol $\bullet$ indicates the entries given by (7), and $\times$ indicates non-zero entries of the two blocks.

For example, if we consider the differentiation formula (3) as basic rule, we have

$$\mathcal{D}_{m,k} = \begin{cases} 1/(12h) & \text{if } k = m - 2 \\ -8/(12h) & \text{if } k = m - 1 \\ 0 & \text{if } k = m \\ 8/(12h) & \text{if } k = m + 1 \\ -1/(12h) & \text{if } k = m + 2 \end{cases} \tag{8}$$ coef-order4

for indices m, k satisfying $M \le m \le N - M$ or $M \le k \le N - M$. For smooth functions $u(x)$ the defect in (5) is $\mathcal{O}(h^4)$ for $M \le m \le N - M$. Our aim is to complete the matrix $\mathcal{D}$ in the upper left and lower right corners such that $\mathcal{D}$ is skew-symmetric and for all indices the defect in (5) is at least $\mathcal{O}(h^3)$.

We note that the differentiation matrices considered here are computationally very efficient. For a given basic differentiation rule only a finite, fairly small number of coefficients (independent of the number of grid points) need be computed.

2.3 Transformation to an arbitrary grid

sect:gengrid

The grid (4) is convenient for theoretical investigations. It has $n = N + 2L - 1$ interior grid points. In practice we have to work with a grid (1) for an interval $[a, b]$. We connect both grids by

$$\xi_k = a + \eta\big(x_{k-L} + a_L h\big), \quad k = 0, 1, \ldots, n+1, \qquad \text{where} \qquad \eta = \frac{b-a}{1 + 2a_L h}.$$

Note that η is the ratio between the lengths of the corresponding intervals. The grid (ξ_k) is also essentially equidistant and only a few subintervals at the end points have a different length.

The coefficients of the differentiation matrix (2) have to be scaled by η:

$$\widehat{\mathcal{D}}_{m,k} = \eta^{-1} \mathcal{D}_{m,k}.$$

3 Order conditions

sect:ordercond

A banded, stable differentiation matrix is of order P if, for all m, the relation (5) is satisfied exactly for all polynomials $u(x)$ of degree P vanishing at the end points.

3.1 The linear system for the order conditions

sect:linearsyst

For differentiation matrices as defined in Section 2.2 the order conditions for order P lead to a linear system with:

- $K(K-1)/2$ unknowns, where $K = M + L - 1$: these are the elements below the diagonal of the skew-symmetric sub-matrix $(\mathcal{D}_{m,k})_{m,k=-L+1}^{M-1}$.

- $(P-1)K$ equations: these are obtained by imposing equality in (5) for $m = -L+1, \dots, M-1$ and for polynomials $u(x)$ vanishing at $-a_L h$ and at $1 + a_L h$. Such polynomials are of the form $u(x) = \big(a_L h(1 + a_L h) + x(1-x)\big)p(x)$, where p is of degree $P-2$.

The resulting linear system can be written in the form

$$\mathcal{A}_{L,M}\mathcal{D}_{L,M} = \mathbf{b}_{L,M}, \qquad (9)$$ linearsystem

where $\mathcal{D}_{L,M}$ is a linear arrangement of the subdiagonal elements of the $K \times K$ block $(\mathcal{D}_{m,k})_{m,k=-L+1}^{M-1}$.

For example, for $L = 3$ and $M = 2$ we have

$$\mathcal{D}_{L,M} = \big(\mathcal{D}_{-1,-2}, \mathcal{D}_{0,-2}, \mathcal{D}_{0,-1}, \mathcal{D}_{1,-2}, \mathcal{D}_{1,-1}, \mathcal{D}_{1,0}\big)$$

and the matrix $\mathcal{A}_{L,M}$ is given (for $P = 3$) by

$$\mathcal{A}_{L,M} = \begin{pmatrix} -u^1(x_{-1}) & -u^1(0) & 0 & -u^1(h) & 0 & 0 \\ u^1(x_{-2}) & 0 & -u^1(0) & 0 & -u^1(h) & 0 \\ 0 & u^1(x_{-2}) & u^1(x_{-1}) & 0 & 0 & -u^1(h) \\ 0 & 0 & 0 & u^1(x_{-2}) & u^1(x_{-1}) & u^1(0) \\ -u^2(x_{-1}) & -u^2(0) & 0 & -u^2(h) & 0 & 0 \\ u^2(x_{-2}) & 0 & -u^2(0) & 0 & -u^2(h) & 0 \\ 0 & u^2(x_{-2}) & u^2(x_{-1}) & 0 & 0 & -u^2(h) \\ 0 & 0 & 0 & u^2(x_{-2}) & u^2(x_{-1}) & u^2(0) \end{pmatrix}$$

where $u^j(x) = \big(a_L h(1 + a_L h) + x(1-x)\big)p^j(x)$, while p^1 and p^2 are linearly independent polynomials of degree $P - 2 = 1$.

The vector $\mathbf{b}_{L,M}$ consists of $P-1$ sub-vectors $\mathbf{b}^j_{L,M}$ (for $j = 1, \dots, P-1$), each of which is formed by the K elements ($m = -L+1, \dots, M-1$)

$$(u^j)'(x_m) - \sum_{k \ge M} \mathcal{D}_{m,k} u^j(x_k). \qquad (10)$$ defbLMj

The coefficients $\mathcal{D}_{m,k} = \delta_{k-m}$ are those inherited from the basic differentiation rule (6). Since $\delta_{k-m} = 0$ for $k - m > R$, the sum in (10) is nonzero only for $m \ge M - R$.

lem:rowrank **Lemma 1.** *For a differentiation matrix of order $P \ge 2$ the row rank of the matrix $\mathcal{A}_{L,M}$ is at most*

$$(P-1)K - \binom{P}{2}.$$

In particular, it is $2K-3$ for $P = 3$ and $3K-6$ for $P = 4$.

Proof. We let $\mathbf{u}^j = \big(u^j(x_{-L+1}), \dots, u^j(x_{M-1})\big)$, and set $\mathbf{0} = (0, \dots, 0)$.

Case $P = 3$. Multiplication of $\mathcal{A}_{L,M}$ from the left with one of the $2K$-dimensional row vectors $(\mathbf{u}^1, \mathbf{0})$, $(\mathbf{0}, \mathbf{u}^2)$ and $(\mathbf{u}^2, \mathbf{u}^1)$ yields the zero vector. Since these vectors are linearly independent, this proves the statement for $P = 3$.

Case $P = 4$. In this case we multiply $\mathcal{A}_{L,M}$ from the left with one of the $3K$-dimensional row vectors $(\mathbf{u}^1, \mathbf{0}, \mathbf{0})$, $(\mathbf{0}, \mathbf{u}^2, \mathbf{0})$, $(\mathbf{0}, \mathbf{0}, \mathbf{u}^3)$, and $(\mathbf{u}^2, \mathbf{u}^1, \mathbf{0})$, $(\mathbf{u}^3, \mathbf{0}, \mathbf{u}^1)$, $(\mathbf{0}, \mathbf{u}^3, \mathbf{u}^2)$ to obtain the zero vector. This proves the case $P = 4$.

A straightforward extension yields the statement for the general case. □

3.2 Necessary conditions for the grid

Lemma 1 leads to necessary conditions for achieving order P. We denote by $\mathbf{u}^j_{L,M} = \big(u^j(x_{-L+1}), \dots, u^j(x_{M-1})\big)$ the vectors used in the proof of Lemma 1, and we split the vector $\mathbf{b}_{L,M}$ into $\big(\mathbf{b}^1_{L,M}, \dots, \mathbf{b}^{P-1}_{L,M}\big)$. We then have the following necessary order conditions for the existence of a solution for (9):

$$\mathbf{u}^j_{L,M} \cdot \mathbf{b}^k_{L,M} + \mathbf{u}^k_{L,M} \cdot \mathbf{b}^j_{L,M} = 0 \qquad \text{for} \quad 1 \le j \le k \le P-1. \tag{11}$$

necessaryc

The dot denotes the scalar product of two vectors. These relations represent nonlinear conditions for the parameters $a_1, \dots, a_L$.

lem:M2 **Lemma 2.** *Assume that the basic differentiation rule (6) is exact for all polynomials of degree P, then the necessary conditions (11) are the same for all $M \ge R$.*

Proof. We consider the first expression of the left-hand side of (11) and compute the difference for two consecutive values of M. With the convention $\delta_j = 0$ for $|j| > R$ this yields

$$\begin{aligned}
&\mathbf{u}^j_{L,M+1} \cdot \mathbf{b}^k_{L,M+1} - \mathbf{u}^j_{L,M} \cdot \mathbf{b}^k_{L,M} \\
&= \sum_{m=-L+1}^{M} u^j(x_m)\Big((u^k)'(x_m) - \sum_{l \ge M+1} \delta_{l-m} u^k(x_l)\Big) \\
&- \sum_{m=-L+1}^{M-1} u^j(x_m)\Big((u^k)'(x_m) - \sum_{l \ge M} \delta_{l-m} u^k(x_l)\Big) \\
&= u^j(x_M)\Big((u^k)'(x_M) - \sum_{l \ge M+1} \delta_{l-M} u^k(x_l)\Big) + \sum_{m=-L+1}^{M-1} u^j(x_m)\delta_{M-m} u^k(x_M).
\end{aligned}$$

Adding the same expression with exchanged values of j and k und using the skew-symmetry relation $\delta_{-j} = -\delta_j$ gives

$$
\begin{aligned}
&\left(\mathbf{u}^j_{L,M+1}\cdot\mathbf{b}^k_{L,M+1} + \mathbf{u}^k_{L,M+1}\cdot\mathbf{b}^j_{L,M+1}\right) - \left(\mathbf{u}^j_{L,M}\cdot\mathbf{b}^k_{L,M} + \mathbf{u}^k_{L,M}\cdot\mathbf{b}^j_{L,M}\right)\\
&\quad = u^j(x_M)\Big((u^k)'(x_M) - \sum_{|l-M|\le R}\delta_{l-M}u^k(x_l)\Big)\\
&\qquad + u^k(x_M)\Big((u^j)'(x_M) - \sum_{|l-M|\le R}\delta_{l-M}u^j(x_l)\Big).
\end{aligned}
$$

For $M \ge R$ the grid points x_l for $l = M-R,\dots,M+R$ all belong to the equidistant sub-grid of (4). Since the degree of the polynomials u^j and u^k is $\le P$, our assumption implies that the right-hand side of the above equation vanishes identically. □

As a consequence of Lemma 2, there is no advantage in considering the nonlinear system (11) for M larger than R. We therefore always assume $M = R$.

lem:indep **Lemma 3.** *The solutions of the nonlinear system (11) are independent of the choice of the functions $p^j(x)$ in the definition of $u^j(x)$.*

Proof. If $\hat u^j$ (for $j = 1,\dots,P-1$), is a linear combination of the functions u^j, $j = 1,\dots,P-1$, the expressions in (11) for the hat quantities are a linear combination of those for the original functions. Consequently, the set of solutions remains unchanged. □

3.3 An explicit form of the order conditions

t:orderexplicit

With the functions $u^j(x) = \big(a_L h(1+a_L h) + x(1-x)\big)x^{j-1}$ and the notation $A = a_L(1+a_L h)$ the condition (11) for order P can be written as

$$
\sum_{m=-L+1}^{M-1}\Big(u^j(x_m)(u^k)'(x_m) + u^k(x_m)(u^j)'(x_m)\Big) = Q_{j,k}(A) \qquad (12)
$$
oc

for $1 \le j \le k \le P-1$, where $Q_{j,k}(A) = P_{j,k}(A) + P_{k,j}(A)$ and

$$
P_{j,k}(A) = \sum_{m=M-R}^{M-1} u^j(mh)\sum_{l=M}^{m+R}\delta_{l-m}u^k(lh).
$$

We have exploited the fact that $\delta_{l-m} = 0$ for $|l - m| > R$ and that $x_m = mh$ for $m \geq 0$. The derivative of $u^j(x) = (hA + x(1-x))x^{j-1}$ is

$$(u^j)'(x) = \big((j-1)hA + jx - (j+1)x^2\big)x^{j-2}.$$

Using the abbreviation (with the convention $0^0 = 1$)

$$S_k = \sum_{l=-L+1}^{M-1} x_l^k = h^k\Big((-a_{L-1})^k + \ldots + (-a_1)^k + 0^k + 1^k + \ldots + (M-1)^k\Big),$$

the order condition (12) thus becomes

$$\begin{aligned} h^2A^2(j+k-2)S_{j+k-3} + 2hA\Big((j+k-1)S_{j+k-2} - (j+k)S_{j+k-1}\Big) \\ +\Big((j+k)S_{j+k-1} - 2(j+k+1)S_{j+k} + (j+k+2)S_{j+k+1}\Big) = Q_{j,k}(A). \end{aligned} \tag{13}$$

We note that not only S_k contains the factor h^k, but also $u^j(x_m)$ and $u^k(x_l)$ contain the factor h^j and h^k, respectively. Therefore, we can divide the equation by h^{j+k-1}. Doing this and using the notation

$$s_k = (-a_{L-1})^k + \ldots + (-a_1)^k + 0^k + 1^k + (M-1)^k \tag{14}$$ skdef

the order conditions finally become (for $1 \leq j \leq k \leq P-1$)

$$\begin{aligned} A^2(j+k-2)s_{j+k-3} + 2A\Big((j+k-1)s_{j+k-2} - (j+k)hs_{j+k-1}\Big) \\ +\Big((j+k)s_{j+k-1} - 2(j+k+1)hs_{j+k} + (j+k+2)h^2s_{j+k+1}\Big) = q^h_{j,k}(A). \end{aligned} \tag{15}$$

where $q^h_{j,k}(A) = h^{-j-k+1}Q_{j,k}(A)$.

rem:cond **Remark 1.** The left-hand expression of (15) is seen to depend only on the sum $j+k$. Therefore, it is necessary that $q^h_{j,k}(A) = q^h_{\hat{j},\hat{k}}(A)$ whenever $j + k = \hat{j} + \hat{k}$. For $P = 3$ this does not give any restriction. For order $P = 4$ we get a single condition, $q^h_{2,2}(A) = q^h_{1,3}(A)$. For order $P = 5$ there are two additional conditions $q^h_{2,3}(A) = q^h_{1,4}(A)$ and $q^h_{3,3}(A) = q^h_{2,4}(A)$, and for general P there is a total of $\binom{P-2}{2}$ conditions.

3.4 Connection with the necessary conditions of [1]

For curiosity we relate the order conditions (11) to those obtained in [1]. Without any stability restriction on the differentiation matrix it is shown in [1] that the grid of a Pth order skew-symmetric differentiation matrix has to satisfy

$$\sum_{k=-L+1}^{N+L-1} f'(x_k) = 0 \qquad \text{for} \qquad f(x) = (x - x_{-L})^2 (x_{N+L} - x)^2 p(x) \tag{16}$$

necessary1

for all polynomials $p(x)$ of degree $2P-4$. For the case of a symmetric grid, the condition (16) is automatically fulfilled for polynomials satisfying $p(1-x) = p(x)$, so that only the polynomials $p(x) = (x - 1/2)^{2j-1}$, $j = 1, \ldots, P-2$ have to be considered. For $P = 3$, this is one equation in contrast to Condition (11) which consists of three equations for $P = 3$. This shows that additional order conditions have to be satisfied for grids of the form (4).

lem:nec1

Lemma 4. *If a grid (4) satisfies the order conditions (11) for order P, then it also satisfies the condition (16) for the same order P.*

Proof. We put $u(x) = (x - x_{-L})(x_{N+L} - x)(x - 1/2)^{j-1} x$ for $j \in \{1, \ldots, P-2\}$ and let $v(x) = u(1-x)$, so that for the function $f(x)$ of (16) we have

$$f(x) = (x - x_{-L})^2 (x_{N+L} - x)^2 \Big(x - \frac{1}{2}\Big)^{2j-1} = \frac{1}{2}\Big(u(x)^2 - v(x)^2\Big). \tag{17}$$

fuv

Condition (11) with $j = k$ and $u(x)$, respectively $v(x)$, in place of $u^j(x)$ (this is justified by Lemma 3) thus reads

$$\sum_{k=-L+1}^{M-1} u(x_k)\Big(u'(x_k) - \sum_{j \ge M} \mathcal{D}_{k,j} u(x_j)\Big) = 0$$

$$\sum_{k=-L+1}^{\widehat{M}-1} v(x_k)\Big(v'(x_k) - \sum_{j \ge \widehat{M}} \mathcal{D}_{k,j} v(x_j)\Big) = 0,$$

where $M \ge R$ and $\widehat{M} \ge R$ can be arbitrarily fixed (Lemma 2). Using the symmetry of the grid and the relation $v(x) = u(1-x)$, the condition for $v(x)$ can be written in terms of $u(x)$ as:

$$\sum_{k=N-\widehat{M}+1}^{N+L-1} u(x_k)\Big(-u'(x_k) - \sum_{j \le N-\widehat{M}} \mathcal{D}_{N-k,N-j} u(x_j)\Big) = 0.$$

By the central skew-symmetry $\mathcal{D}_{N-k,N-j} = -\mathcal{D}_{k,j}$ and with the choice of $\widehat{M}$ such that $M + \widehat{M} = N + 1$, a subtraction of both relations for $u(x)$ yields

$$\sum_{k=-L+1}^{N+L-1} u(x_k)u'(x_k) = 0, \tag{18}$$

uup

because the two double sums cancel by the skew-symmetry of $\mathcal{D}_{k,j}$. An analogous relation is obtained for $v(x)$. The statement now follows from (17) and the fact that the relation (16) is satisfied for $u(x)^2$ and $v(x)^2$ by (18). □

This is a point to emphasise a subtle issue. In [1] the condition (16) is also sufficient for the existence of a skew-symmetric matrix of order P on a given grid, except that such a matrix, which is banded, need not be stable. Once we construct differentiation matrices like in this paper, (16) is clearly insufficient. It is an open problem whether (16) is necessary *and* sufficient for the construction of a stable, skew-symmetric differentiation matrix for every $P \geq 3$.

4 Differentiation matrices of order 3

t:constr-order3

In this section we construct skew-symmetric differentiation matrices of the form introduced in Section 2 and based on the differentiation formula (3).

4.1 Order conditions for order 3

sect:order3

For the choice (8) of the basic differentiation formula we have $R = 2$. We put $M = 2$ and get $q^h_{j,k}(A) = p^h_{j,k}(A) + p^h_{k,j}(A)$ with (see Section 3.3)

$$p^h_{j,k}(A) = \frac{1}{12\,h^{j+k}}\Big(-u^j(0)u^k(2h) + u^j(h)\big(8u^k(2h) - u^k(3h)\big)\Big).$$

Inserting $u^j(x) = (hA + x(1-x))x^{j-1}$ into (15) results in the system

$$\begin{aligned} 2A(s_0 - 2hs_1) + (2s_1 - 6hs_2 + 4h^2s_3) &= q^h_{1,1}(A) \\ &= \tfrac{1}{6}\big(6A^2 + 18A + 13 - 36h - 26hA + 23h^2\big) \\ A^2s_0 + 2A(2s_1 - 3hs_2) + (3s_2 - 8hs_3 + 5h^2s_4) &= q^h_{1,2}(A) \\ &= \tfrac{1}{6}\big(9A^2 + 26A + 18 - 48h - 36hA + 30h^2\big) \\ 2A^2s_1 + 2A(3s_2 - 4hs_3) + (4s_3 - 10hs_4 + 6h^2s_5) &= q^h_{2,2}(A) \\ &= \tfrac{1}{6}\big(13A^2 + 36A + 23 - 60h - 50hA + 37h^2\big) \end{aligned} \tag{19}$$

linsyst3h

which is necessary for order $P = 3$. We choose $L = 3$, so that three parameters a_1, a_2, a_3 are available. We solve this system numerically and obtain the following values:

N	a_1	a_2	a_3
10	1.003781632659	2.007632287816	2.752472830434
100	1.003023528871	2.004132100779	2.762836922708
1000	1.002922802068	2.003722837122	2.764655844052
10000	1.002912404585	2.003681363399	2.764850531480
100000	1.002911361492	2.003677210713	2.764870136085
∞	1.002911245556	2.003676749244	2.764872315908

The coefficients of the corresponding differentiation matrix are obtained from the solution of the linear system (9). Their values, multiplied by h, are as follows:

N	$\mathcal{D}_{-1,-2}$ / $\mathcal{D}_{1,-2}$	$\mathcal{D}_{0,-2}$ / $\mathcal{D}_{1,-1}$	$\mathcal{D}_{0,-1}$ / $\mathcal{D}_{1,0}$
10	- 0.3387467173411	- 0.4225146995560	- 0.3192512422566
	0.2032627246264	- 0.10262241060990	- 0.6027828536519
100	- 0.3075432630829	- 0.4243406234660	- 0.3351340348980
	0.1896736705106	- 0.08174504699602	- 0.6135039576113
1000	- 0.3034984572193	- 0.4242812535380	- 0.3374683924251
	0.1879789457053	0.07914665533779	- 0.6147792562106
10000	- 0.3030820398568	- 0.4242713002458	- 0.3377122271803
	0.1878052095303	0.07888039374437	- 0.6149091785254
∞	- 0.3030356334373	- 0.4242701465998	- 0.3377394411464
	0.1877858563381	0.07885073516130	- 0.6149236416578

All coefficients $\mathcal{D}_{m,k}$ are seen to be bounded by $\mathcal{O}(h^{-1})$ and thus lead to stable discretizations.

sect:limitcase

4.2 The limit case $h \to 0$

We set $h = 0$ in the order conditions (19) and we note that $A = a_L(1 + a_L h)$ becomes a_L. The conditions (for order $P = 3$) are thus (we write a instead of a_L)

$$\begin{aligned} 2as_0 + 2s_1 &= q_{1,1}(a) = \tfrac{1}{6}\big(6a^2 + 18a + 13\big) \\ a^2 s_0 + 4as_1 + 3s_2 &= q_{1,2}(a) = \tfrac{1}{6}\big(9a^2 + 26a + 18\big) \\ 2a^2 s_1 + 6as_2 + 4s_3 &= q_{2,2}(a) = \tfrac{1}{6}\big(13a^2 + 36a + 23\big). \end{aligned} \quad (20)$$

linsyst3

Since $s_0 = L + 1$, this system permits us to compute s_1, s_2, s_3 as functions of $a = a_L$. For $L \geq 4$ we can arbitrarily fix a and then we can obtain $a_1, \ldots, a_{L-1}$ from the values for s_1, s_2, s_3. Here, we are interested in the case $L = 3$. In this situation we obtain

$$\begin{aligned} -a_1 - a_2 + 1 &= s_1 = \tfrac{1}{12}\big(6a^2 - 30a + 13\big) \\ a_1^2 + a_2^2 + 1 &= s_2 = \tfrac{1}{6}\big(-4a^3 + 15a^2 + 6\big) \\ -a_1^3 - a_2^3 + 1 &= s_3 = \tfrac{1}{24}\big(18a^4 - 60a^3 + 23\big). \end{aligned}$$

With the aim of obtaining a polynomial equation for $a = a_3$, we compute

$$\begin{aligned} s_1^2 &= a_1^2 + a_2^2 + 1 + 2a_1a_2 - 2a_1 - 2a_2 = 2a_1a_2 + s_2 + 2(s_1 - 1) \\ s_1s_2 &= \ldots = s_3 + (a_1a_2 + 1)(s_1 - 1) + s_2 - 1 \end{aligned}$$

Elimination of the term a_1a_2 gives the relation

$$s_1^3 + 2s_3 - 3s_1s_2 - 3s_1^2 + 6s_1 + 3s_2 - 6 = 0$$

and leads to the polynomial equation for $a = a_3$

$$\chi(a) = 216a^6 - 1512a^5 + 3780a^4 - 4032a^3 + 1638a^2 - 90a - 143 = 0. \quad (21)$$ polynom

A plot of this polynomial function is given in Figure 2. There are two real zeros. One is close to -0.2144 and the other is (up to 20 digits)

$$a_3 = 2.7648723159082823436\,6.$$

This zero is in agreement with the numerical computations of Section 4.1.

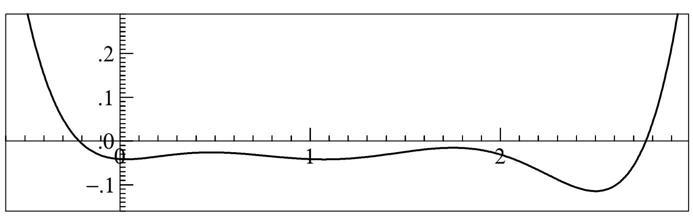

Figure 2: The graph of the polynomial $\chi(a)$ of (21). fig:polynom

4.3 Order 4 is not possible

sect:not4

Let us try to construct a differentiation matrix of the form (8) which has order $P = 4$. By choosing $L \geq 4$ we can introduce more parameters. In addition to (19) there are three more order conditions for $P = 4$. The critical one is for $(j, k) = (1, 3)$ and it is given by

$$\begin{aligned} 2A^2 s_1 + 2A(3s_2 - 4hs_3) + (4s_3 - 10hs_4 + 6h^2 s_5) &= q^h_{1,3}(A) \\ &= \tfrac{1}{6}\big(13A^2 + 36A + 25 - 60h - 42hA + 35h^2\big), \end{aligned}$$

where $A = a_L(1+a_L h)$. The left-hand side is the same as that for the last equation of (19). Therefore, we can have a differentiation matrix of order $P = 4$ only if

$$0 = q^h_{1,3}(A) - q^h_{2,2}(A) = \tfrac{1}{6}\big(2 + 8hA - 2h^2\big)$$

(see Remark 1 of Section 3.3). This relation cannot be satisfied with $a_L = a_L(h)$ uniformly bounded for $h \to 0$.

5 An alternative approach for order 4

t:constr-order4

Up to an error of size $\mathcal{O}(h^6)$, we have

$$h^5 u^{(5)}(0) \approx \frac{1}{240}\Big(-u(-3h) + 4u(-2h) - 5u(-h) + 5u(h) - 4u(2h) + u(3h)\Big)$$

We keep the grid (4) unchanged, but consider differentiation matrices as in Section 2.2, where

$$\mathcal{D}_{m,k} = \begin{cases} -\ \gamma/(240h) & \text{if } k = m-3 \\ 1/(12h) + 4\,\gamma/(240h) & \text{if } k = m-2 \\ -8/(12h) - 5\,\gamma/(240h) & \text{if } k = m-1 \\ 0 & \text{if } k = m \\ 8/(12h) + 5\,\gamma/(240h) & \text{if } k = m+1 \\ -1/(12h) - 4\,\gamma/(240h) & \text{if } k = m+2 \\ \gamma/(240h) & \text{if } k = m+3 \end{cases} \qquad (22)$$

coef-order4a

for indices m, k satisfying $M \leq m \leq N - M$ or $M \leq k \leq N - M$ and for a given parameter γ. The bandwidth is increased by one, so that $R = 3$. The underlying differentiation rule is of order 4 for all values of γ. For $\gamma = 0$ we obtain the matrix of Section 4.1, and for $\gamma = 4 + \mathcal{O}(h^2)$ the coefficients of (22) represent a derivative approximation of order 6.

5.1 Order conditions

The matrix of the linear system (9) representing the order conditions remains the same, only the inhomogeneity vector $\mathbf{b}_{L,M}$ has to be adapted. Since the bandwidth of the matrix is increased, we work with $M = 3$. With s_k given by (14) and the notation $A = a_L(1 + a_L h)$ in place, the condition (15) for order $P = 4$ becomes

$$\begin{aligned}
&2A(s_0 - 2hs_1) + (2s_1 - 6hs_2 + 4h^2 s_3) = q^h_{1,1}(A) \\
&\quad = \tfrac{1}{6}\big(6A^2 + 30A + 37 - 180h - 74hA + 215h^2\big) + \tfrac{\gamma}{30}h^2, \\
&A^2 s_0 + 2A(2s_1 - 3hs_2) + (3s_2 - 8hs_3 + 5h^2 s_4) = q^h_{1,2}(A) \\
&\quad = \tfrac{1}{6}\big(15A^2 + 74A + 90 - 432h - 180hA + 510h^2\big) + \tfrac{\gamma}{60}h, \\
&2A^2 s_1 + 2A(3s_2 - 4hs_3) + (4s_3 - 10hs_4 + 6h^2 s_5) = q^h_{2,2}(A) \\
&\quad = \tfrac{1}{6}\big(37A^2 + 180A + 215 - 1020h - 434hA + 1189h^2\big) \\
&\qquad + \tfrac{\gamma}{30}\big(1 + 3hA - 3h^2\big), \\
&2A^2 s_1 + 2A(3s_2 - 4hs_3) + (4s_3 - 10hs_4 + 6h^2 s_5) = q^h_{1,3}(A) \\
&\quad = \tfrac{1}{6}\big(37A^2 + 180A + 217 - 1020h - 426hA + 1187h^2\big) \\
&\qquad - \tfrac{\gamma}{60}\big(3 + 14hA - 8h^2\big), \\
&3A^2 s_2 + 2A(4s_3 - 5hs_4) + (5s_4 - 12hs_5 + 7h^2 s_6) = q^h_{2,3}(A) \\
&\quad = \big(15A^2 + 72A + 85 - 396h - 170hA + 455h^2\big) - \tfrac{\gamma}{60}\big(A + 2h\big), \\
&4A^2 s_3 + 2A(5s_4 - 6hs_5) + (6s_5 - 14hs_6 + 8h^2 s_7) = q^h_{3,3}(A) \\
&\quad = \tfrac{1}{6}\big(215A^2 + 1020A + 1189 - 5460h - 2374hA + 6191h^2\big) \\
&\qquad - \tfrac{\gamma}{30}\big(3 - A^2 + 8hA - 7h^2\big).
\end{aligned} \tag{23}$$

linsyst4h

Since the left-hand expression is the same for the 3rd and 4th equations, a necessary condition for order $P = 4$ is that $q^h_{1,3}(A) = q^h_{2,2}(A)$. This condition yields a linear equation for γ, which gives

$$\gamma = 4 \cdot \frac{1 + 4hA - h^2}{1 + 4hA - \frac{14}{5}h^2} = 4 + \frac{36}{5}h^2 - \frac{144A}{5}h^3 + \mathcal{O}(h^4). \tag{24}$$

gamma

It is interesting to note that with this choice of γ we have, away from the end points, an order 6 approximation of the derivative.

Before passing to concrete computations, let us consider the limit case $h \to 0$.

5.2 The limit case: $h \to 0$

In the limit $h \to 0$ we see from (24) that $\gamma = 4$. With this value of γ the system (23) becomes (again we write a for a_L)

$$\begin{aligned} 2as_0 + 2s_1 &= q_{1,1}(a) = \tfrac{1}{6}\big(6a^2 + 30a + 37\big) \\ a^2s_0 + 4as_1 + 3s_2 &= q_{1,2}(a) = \tfrac{1}{6}\big(15a^2 + 74a + 90\big) \\ 2a^2s_1 + 6as_2 + 4s_3 &= q_{2,2}(a) = \tfrac{1}{30}\big(185a^2 + 900a + 1079\big) \\ 2a^2s_1 + 6as_2 + 4s_3 &= q_{1,3}(a) = \tfrac{1}{30}\big(185a^2 + 900a + 1079\big) \\ 3a^2s_2 + 8as_3 + 5s_4 &= q_{2,3}(a) = \tfrac{1}{15}\big(225a^2 + 1079a + 1275\big) \\ 4a^2s_3 + 10as_4 + 6s_5 &= q_{3,3}(a) = \tfrac{1}{30}\big(1079a^2 + 5100a + 5933\big). \end{aligned} \tag{25}$$

linsyst4

This system permits us to compute explicitly the expressions $s_1, \dots, s_5$. We fix $L = 6$, $M = 3$ (so that $s_0 = L + M - 1 = 8$), choose arbitrarily $a_L = a_6$, and compute $a_1, \dots, a_5$ from the nonlinear system given by $s_1, \dots, s_5$. Figure 3 shows the numerically computed values $a_1, \dots, a_6$ as function of a_6. For the parameter a_6, values between 4 and 6 are considered. If a_6 is too close to 6, we have bad convergence of the Newton iterations. For $a_6 \le 5$, the grid points a_5 and a_6 are very close. The best choice of a_6 seems to be in the range between 5.7 and 5.8.

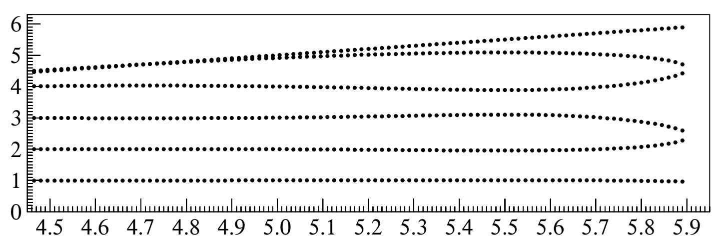

Figure 3: The grid points $a_1, \dots, a_6$ (top to bottom) as a function of a_6. fig:gridcomp

5.3 Numerical computations for the general case

Encouraged by the successful computations for the limit case $h \to 0$ (or $N \to \infty$), we attack the general case with $h > 0$. We fix $a_L = a_6 = 5.75$ and consider many different values of N (recall that $h = 1/N$). The solution can be found numerically by Newton's method. Figure 4 shows the solution $a_1, \dots, a_6$ as a function of N. It comes as a welcome surprise that the grid remains nearly unchanged for all values of N. For $N = 1000$ we obtain

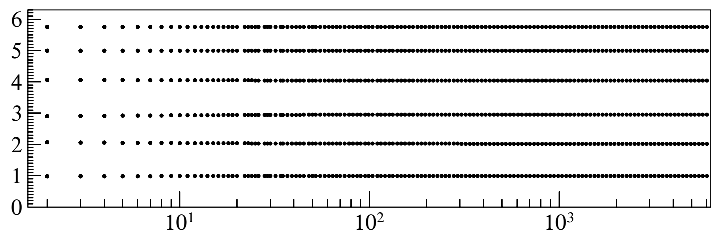

Figure 4: The grid points $a_1, \ldots, a_6$ as a function of the parameter N for the choice $a_6 = 5.75$.

fig:gridcompfix

N	a_1 / a_4	a_2 / a_5	a_3 / a_6
1000	0.99533846829173	2.02375529528507	2.95622105838923
	4.03687073420997	4.99823111049117	5.75000000000000

We see that the values $a_1, \ldots, a_5$ are very close to the continuation of the grid with constant step size. Similarly to the situation in Section 4.1 only the final subinterval is smaller.

The corresponding entries $\mathcal{D}_{j,k}$ of the differentiation matrix are obtained from the solution of a linear system (see Section 3.1). We have $(P-1)K = 24$ equations for $K(K-1)/2 = 28$ unknowns. This gives the freedom to fix some of the entries. To get a small bandwidth also in the left upper part (and the right lower part), we arbitrarily require

$$\mathcal{D}_{j,k} = 0 \qquad \text{for} \quad |k-j| \geq 6.$$

There remain 25 unknowns for the 24 equations. We compute a least-squares solution from the QR decomposition of the relevant matrix. For $N = 1000$ the leading digits of the coefficients $\mathcal{D}_{j,k}$, multiplied by h, are displayed in the following table:

$h\mathcal{D}_{j,k}$	$k=-5$	$k=-4$	$k=-3$	$k=-2$	$k=-1$	$k=0$	$k=1$
$j=-4$	-0.385						
$j=-3$	-0.323	-0.255					
$j=-2$	0.013	-0.201	-0.202				
$j=-1$	0.210	-0.058	-0.281	-0.174			
$j=0$	-0.084	0.099	-0.094	-0.180	-0.167		
$j=1$	0	-0.019	0.238	-0.082	-0.343	-0.324	
$j=2$	0	0	-0.085	0.067	0.116	-0.037	-0.686

6 Higher order differentiation matrices

onstr-orderhigh

There does not seem to be serious difficulty in constructing skew-symmetric differentiation matrices of order higher than four in a similar manner.

Order 5. In addition to the condition $q^h_{1,3}(A) = q^h_{2,2}(A)$ we have to require that $q^h_{1,4}(A) = q^h_{2,3}(A)$ and $q^h_{2,4}(A) = q^h_{3,3}(A)$. A solution of the arising nonlinear system should be possible with three parameters in the basic differentiation formula. The questions are: (a) does the solution of the nonlinear system for $h = 0$ lead to a grid with monotonically increasing and well separated grid points? (b) If such a grid exists, we have by continuity a nearby grid for sufficiently large N. Does it exist for all values of N?

7 Applications to partial differential equations

sect:numerical

This section is devoted to a few examples that demonstrate the advantage of using skew-symmetric differentiation matrices.

7.1 Advection equation

sect:advection

For $t \geq 0$ and $x \in [0, 1]$ we consider the one-dimensional advection equation

$$\partial_t u + c(x)\partial_x u = 0 \tag{26}$$

advection1

where the wave speed $c(x)$ depends on the spatial coordinate. We assume zero speed at the boundary, $c(0) = c(1) = 0$, so that homogeneous Dirichlet boundary conditions, $u(t, 0) = u(t, 1) = 0$ for all $t \geq 0$, make sense. The initial condition $u(0, x) = u_0(x)$ is assumed to be compatible with the boundary conditions.

Under these assumptions the advection equation has for all sufficiently smooth functions $F(u)$ the expression

$$I_F(t) = \int_0^1 \frac{F\big(u(t,x)\big)}{c(x)}\,\mathrm{d}x \tag{27}$$

conserved

as a conserved quantity. This follows from differentiation with respect to time t. In particular, this is the case for $F(u) = |u|^2$.

For a grid $0 = \xi_0 < \xi_1 < \cdots < \xi_n < \xi_{n+1} = 1$ (as the one of Section 2.3) we consider the following space discretisation of (26),

$$\dot{U} + \mathcal{C}\widehat{\mathcal{D}}U = 0, \qquad U(0) = U_0. \tag{28}$$

advectionnum

The elements of the vector $U(t) = \big(U_1(t),\dots,U_n(t)\big)^{\mathsf T}$ are approximations to $u(t,\xi_j)$ (for $j=1,\dots,n$), $\mathcal{C}$ is a diagonal matrix with entries $c(\xi_j)$, and $\widehat{\mathcal{D}}$ is a differentiation matrix as introduced in Section 1.

prop:conserv **Proposition 1.** *Assume that $c(x) > 0$ for $x \in (0,1)$, and that the differentiation matrix $\widehat{\mathcal{D}}$ is skew-symmetric. Then, the expression*

$$\mathcal{I}_2(t) = \frac{1}{n+1}\sum_{j=1}^{n}\frac{1}{c(\xi_j)}|U_j(t)|^2$$

is preserved along solutions of (28). Moreover, the differential equation (28) is stable and all the eigenvalues of $\mathcal{C}\widehat{\mathcal{D}}$ reside on the imaginary axis.

Proof. The preservation of $\mathcal{I}_2(t)$ follows from

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{I}_2(t) &= \frac{2}{n+1}\sum_{j=1}^{n}\frac{1}{c(\xi_j)}U_j(t)\dot U_j(t)\\ &= \frac{2}{n+1}U(t)^{\mathsf T}\mathcal{C}^{-1}\dot U(t) = -\frac{2}{n+1}U(t)^{\mathsf T}\widehat{\mathcal{D}}U(t),\end{aligned}$$

which vanishes identically due to the skew-symmetry of $\widehat{\mathcal{D}}$.

The stability of (28) can be seen from the fact that the eigenvalues of $\mathcal{C}\widehat{\mathcal{D}}$ are the same as those of the skew-symmetric matrix $\mathcal{C}^{1/2}\widehat{\mathcal{D}}\mathcal{C}^{1/2}$. □

To demonstrate that the properties of Proposition 2 are in general not satisfied, we consider the equidistant grid $\xi_j = jh$ ($j=0,1,\dots,n+1$) with $h = 1/(n+1)$. For $j = 2,\dots,n-1$ we use the differentiation formula (3), for $j=1$ the formula

$$u'(\xi) \approx \frac{1}{6h}\Big(-2u(\xi-h) - 3u(\xi) + 6u(\xi+h) - u(\xi+2h)\Big)$$

and for $j = n$ the reflected formula. This yields a banded, 3rd order differentiation matrix, which is not skew-symmetric at the upper left and lower right corners. We call it Method (N). For comparison we consider the skew-symmetric differentiation matrix of Section 4.1, and we call it Method (S).

We apply both discretisations to the advection equation (26) with wave speed $c(x) = x(1-x)^2$ and initial condition $u_0(x) = 3x(1-x)$. We use $N = 100$ for method (S) and $n = 105$ for method (N), so that both methods have 105 grid points in the interior of the interval $(0,1)$. The corresponding ordinary differential equation (28) is solved with high precision by an explicit Runge-Kutta

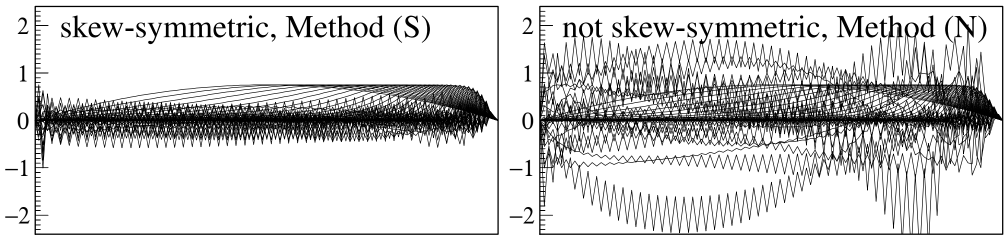

Figure 5: Solution $U_j(t)$ of the semi-discretised differential equation (28) as a function of $\xi_j \in [0, 1]$, for various values of t ranging from $t = 0$ to $t = 100$. g:conserv-dopri

code. The result is shown in Figure 5. For many different time instances ranging from $t = 0$ to $t = 100$ we plot the solution $U_j(t)$ as a function of ξ_j (for better visibility the values are connected by a polygon). For values of t between 0 and 30, the functions are nearly the same for both methods. For larger values of t, spurious oscillations can be observed. Whereas they remain bounded and of moderate size for the skew-symmetric Method (S), they grow exponentially fast for Method (N). This can be explained by computing the eigenvalues of the matrix $\mathcal{C}\widehat{\mathcal{D}}$ of (28). By Proposition 2 they lie exactly on the imaginary axis for Method (S). For method (N), numerical computations show that the matrix has eigenvalues with positive real part up to a size of ≈ 0.025.

7.2 Inhomogenous Dirichlet boundary conditions

The stability result of the previous section is not restricted to homogeneous boundary conditions. Consider again the advection equation (26) with wave speed satisfying $c(0) = c(1) = 0$. Here, we consider inhomogeneous boundary conditions

$$u(t, 0) = a(t), \qquad u(t, 1) = b(t), \qquad t \geq 0. \tag{29}$$ inhomo

It is common to consider the linear interpolation $\overline{u}(t, x) = (1 - x)a(t) + xb(t)$ and to work with the difference $v(t, x) = u(t, x) - \overline{u}(t, x)$. If $u(t, x)$ is solution of (26), then $v(t, x)$ solves

$$\partial_t v + c(x)\partial_x v = f(t, x), \tag{30}$$ advection1v

where $f(t, x) = c(x)\big(b(t) - a(t)\big) - (1 - x)\dot{a}(t) - x\dot{b}(t)$. Space discretisation of this equation yields

$$\dot{V} + \mathcal{C}\widehat{\mathcal{D}}V = F(t), \qquad U(0) = U_0, \tag{31}$$ advectionnumv

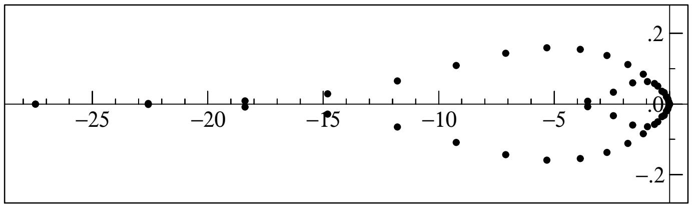

fig:diff-valpr Figure 6: Eigenvalues of $\widehat{\mathcal{D}}\,\mathcal{C}\widehat{\mathcal{D}}$, where $\mathcal{D}$ is the differentiation matrix of Method (N) with $n = 50$, and $\mathcal{C}$ is given by $c(x) = x(1-x)\big((3x-1)^2 + 0.001\big)$.

where the elements of the vector $V(t) = \big(V_1(t), \ldots, V_n(t)\big)^{\mathsf{T}}$ approximate $v(t, \xi_j)$. Since the homogeneous part of (31) is the same as that of (28), Proposition 2 guarantees the stability of (31) if a skew-symmetric differentiation matrix is used.

7.3 Diffusion equation

As a further application we consider the diffusion equation

$$\partial_t u = \partial_x \big(c(x)\partial_x u\big), \qquad (32)$$ diffusion

where $c(x) > 0$ in $(0, 1)$. We assume homogeneous Dirichlet boundary conditions $u(t, 0) = u(t, 1) = 0$ for all $t \geq 0$, and an initial condition $u(0, x) = u_0(x)$. Moreover, we let $c(0) = c(1) = 0$ so that not only $u(t, x)$, but also $c(x)\partial_x u(t, x)$ vanish at the endpoints of $[0, 1]$. Similarly to Section 7.1, an approximation of the space derivatives by finite differences leads to an ordinary differential equation

$$\dot U = \widehat{\mathcal{D}}\,\mathcal{C}\widehat{\mathcal{D}}\,U, \qquad U(0) = U_0. \qquad (33)$$ diffusionnum

prop:conserv **Proposition 2.** *Assume that the differentiation matrix $\widehat{\mathcal{D}}$ is skew-symmetric. Then the matrix of the linear system (33) is symmetric and negative semi-definite.*

Proof. The statement is obvious. □

As a consequence, the linear differential equation (33) and also its inhomogeneous analogue can be numerically solved by any integrator that has the negative real axis (or a large part thereof) in its stability region.

If $\widehat{\mathcal{D}}$ is not skew symmetric, the eigenvalues of $\widehat{\mathcal{D}}\,\mathcal{C}\widehat{\mathcal{D}}$ are in general not on the negative real axis. To justify this claim, we consider Method (N) of Section 7.1 and apply it to the diffusion equation (32). Figure 6 shows the eigenvalues of $\widehat{\mathcal{D}}\,\mathcal{C}\widehat{\mathcal{D}}$ for $n = 50$ and for the choice $c(x) = x(1-x)\big((3x-1)^2 + 0.001\big)$.

Only a few of the eigenvalues lie on the real axis. For numerical integrators that have only a narrow band around the negative real axis in the stability region (e.g., Runge–Kutta–Chebyshev methods, see [2, Section IV.2]), this may lead to severe step-size restrictions.

References

hairer15nsi [1] E. Hairer and A. Iserles. Numerical stability in the presence of variable coefficients. *Found. Comput. Math.*, pages 1–27, 2015.

hairer96sod [2] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*. Springer Series in Computational Mathematics 14. Springer-Verlag, Berlin, 2nd edition, 1996.

iserles14osd [3] A. Iserles. On skew-symmetric differentiation matrices. *IMA J. Numer. Anal.*, 34(2):435–451, 2014.

kitson03saf [4] A. Kitson, R. I. McLachlan, and N. Robidoux. Skew-adjoint finite difference methods on nonuniform grids. *New Zealand J. Math.*, 32(2):139–159, 2003.

kreiss74fea [5] H.-O. Kreiss and G. Scherer. Finite element and finite difference methods for hyperbolic partial differential equations. In C. de Boor, editor, *Proceedings of a Symposium conducted by the Mathematics Research Center, University of Wisconsin, Madison, Wis., April 1–3, 1974*, pages 195–212. Academic Press, New York-London, 1974.

mclachlan03sdo [6] R. I. McLachlan. Spatial discretization of partial differential equations with integrals. *IMA J. Numer. Anal.*, 23(4):645–664, 2003.

olsson95sbp [7] P. Olsson. Summation by parts, projections, and stability. I. *Math. Comp.*, 64(211):1035–1065, S23–S26, 1995.

strand94sbp [8] B. Strand. Summation by parts for finite difference approximations for d/dx. *J. Comput. Phys.*, 110(1):47–67, 1994.