Delft University of Technology
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft Institute of Applied Mathematics

Limiting and shock detection for discontinuous Galerkin solutions using multiwavelets

A thesis submitted to the
Delft Institute of Applied Mathematics
in partial fulfillment of the requirements

for the degree

MASTER OF SCIENCE
in
APPLIED MATHEMATICS

by

M.J. Vuik

Delft, The Netherlands
July, 2012

TU Delft

MSc thesis APPLIED MATHEMATICS

Limiting and shock detection for discontinuous Galerkin solutions using multiwavelets

M.J. VUIK

Delft University of Technology

Daily supervisor

Dr. J.K. Ryan

Responsible professor

Prof.dr.ir. A.W. Heemink

Other thesis committee members

Dr.ir. F.J. Vermolen

July, 2012

Delft, The Netherlands

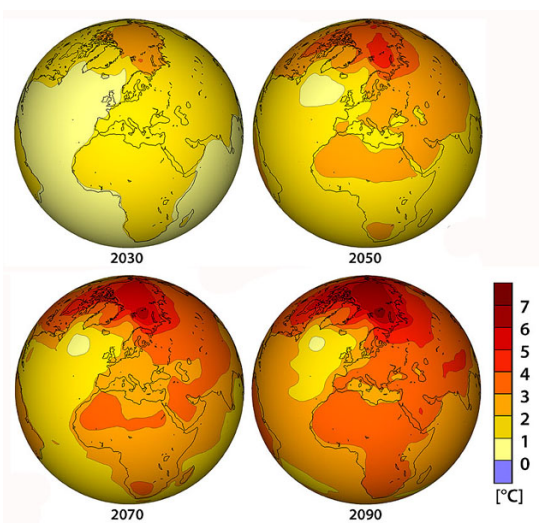
Contents

Preface

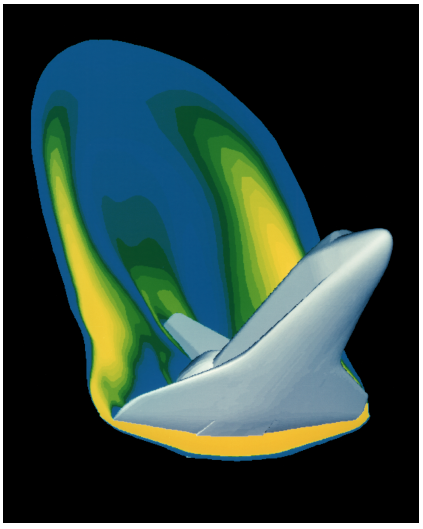
This master thesis is written as a part of my master programme, Applied Mathematics, at Delft University of Technology, and contains the results after a nine month during research.
I am very grateful to my daily supervisor, Jennifer Ryan. She is one of the most enthusiastic scientists that I met during my education at university. We had nice and fruitful discussions, and we are close friends.
I would like to thank Arnold Heemink and Fred Vermolen, who are part of my thesis committee, and attended the presentations about my work.
Furthermore, I owe much gratitude to my parents, family, and fellow students, for their constant encouragement and support.

Introduction

Many areas such as climate modelling, shallow water equations, and computational fluid dynamics use hyperbolic partial differential equations (PDE's) to describe the behaviour of some unknown quantity.
Examples of results that are computed by solving partial differential equations, are given in Figure 1.



(a) Climate modelling: simulation of the mean temperature change



(b) A computer simulation of high velocity air flow around the Space Shuttle during re-entry

Figure 1: Simulations that are computed by solving PDE's

Climate models are used for various purposes, such as computing weather forecasts, predicting hurricanes, and to look at climate changes, ocean circulations, etc. Clearly, it is very important to find accurate results, that are computed fast (we do not want to perform a computation that takes three weeks, to predict the weather for tomorrow).
The air flow around the space shuttle is computed using Computational Fluid Dynamics (CFD). We want to find reliable results, to avoid the use of costly wind tunnels. CFD is used for all problems that involve fluid flows. Most applications (for example the climate models and CFD problems) use nonlinear PDE models. In general, the solutions of these equations contain shocks, or develop discontinuities.

In order to create the simulations, as visualized in Figure 1, a model based on a PDE and numerical method is constructed. To solve these equations, various types of numerical methods can be used, such as finite difference, finite volume, and finite element methods. In this master thesis, the discontinuous Galerkin method (DG) is used, which is defined in Section 1.1. We concentrate on this method due to its increasing use in the CFD community, as well as many others. As we only use DG in space, this method transforms

the PDE into an ordinary differential equation in time. The time stepping is done using the third order total variation diminishing Runge Kutta method (Section 1.2).
Section 1.3 gives necessary information about multiwavelets. In this report, we use multiwavelets to transform the DG coefficients of an approximate solution into multiwavelet coefficients. Several examples of nonlinear PDE's are given in Section 1.4.
In practical applications, initial conditions may contain discontinuities, or the solution of a nonlinear equation may develop a discontinuity at a certain time. To efficiently apply DG in case of discontinuous solutions, limiting techniques are used to reduce the spurious oscillations, that are developed in the discontinuous regions. Information about limiters is given in Section 1.5. Unfortunately, most of the limiters do not work well for higher order approximations, or multidimensional cases.

Originally, this project focused on limiting DG solutions using multiwavelets. We hoped this would be a way to reduce spurious oscillations. This work is described in Section 2, where new multiwavelet limiters are introduced.
In Section 2.2 for example, we define limiters that use the multiwavelet contribution of the approximate solution. In Section 2.3, we construct limiters that adapt the multiwavelet coefficients to limit the approximation.

The work on limiting demonstrates that it is very difficult to limit the solution using the multiwavelet decomposition. However, we discovered that the multiwavelet expansion is quite practical for shock detection. This is the reason for new work presented in Sections 3 and 4, which deal with shock detection methods for DG. It turns out that both DG coefficients and multiwavelet coefficients can be used to detect discontinuities in the DG approximation. This is explained in Section 3, and we give some examples.
In Section 4, we combine the shock detectors with existing limiters. The limiter is applied only in the elements that are marked by the shock detector. Section 4.1 uses test problems where the exact solution is known. The results that are found in Sections 4.2 (Sod's shock tube) and 4.3 (Lax's shock tube) are promising, and can be used for further research. Sections 2 to 4 contain new work, that is not described in literature.
Finally, we state our conclusions and proposal for new work.

1 Background and motivation

In this section, some background information is given, that is found in literature. We start with an explanation of the discontinuous Galerkin method in one dimension.

1.1 The discontinuous Galerkin method

In this section, the discontinuous Galerkin method is explained using the following partial differential equation on $[a, b]$:

$$u_t + f(u)_x = 0, \quad x \in [a, b], t \geq 0, \tag{1a}$$
$$u(x, 0) = u^0(x), \quad x \in [a, b], \tag{1b}$$

where $u = u(x, t)$.
To discretize in space a mesh is defined. Therefore, N is chosen such that the number of elements in $[a, b]$ is $N + 1$, and $\Delta x = \frac{b-a}{N+1}$. The cell centers are given by,

$$x_j = a + \left(j + \frac{1}{2}\right)\Delta x, \quad j = 0, \ldots, N, \tag{2}$$

with corresponding elements,

$$I_j = (x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}] = \left(x_j - \frac{\Delta x}{2}, x_j + \frac{\Delta x}{2}\right], \quad j = 0, \ldots, N.$$

In general, the (unscaled) Legendre polynomials are used to approximate the solution, but we use the scaled Legendre polynomials for this purpose (due to the multiwavelets, which use the same basis, Section 1.3). For degree 0 to 4, the scaled Legendre polynomials on I_j are given by,

$$\phi_0(\xi) = \sqrt{\frac{1}{2}}, \tag{3a}$$
$$\phi_1(\xi) = \sqrt{\frac{3}{2}}\xi, \tag{3b}$$
$$\phi_2(\xi) = \frac{1}{2}\sqrt{\frac{5}{2}}(3\xi^2 - 1), \tag{3c}$$
$$\phi_3(\xi) = \frac{1}{2}\sqrt{\frac{7}{2}}(5\xi^3 - 3\xi), \tag{3d}$$
$$\phi_4(\xi) = \frac{1}{8}\sqrt{\frac{9}{2}}(35\xi^4 - 30\xi^2 + 3), \tag{3e}$$

where $\xi = \frac{2}{\Delta x}(x - x_j)$, such that $[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$ (global coordinates) transforms into $[-1, 1]$ (local coordinates). If $\xi \notin [-1, 1]$, then $\phi_\ell(\xi) = 0, \ell = 0, 1, \ldots$, and the functions ϕ_ℓ are pairwise orthogonal, which means that,

$$\langle \phi_\ell, \phi_m \rangle = \int_{-1}^{1} \phi_\ell(\xi)\phi_m(\xi)d\xi = \delta_{\ell m}. \tag{4}$$

The (unscaled) Legendre polynomials are given by $P^{(\ell)}(\xi) = \sqrt{\frac{2}{2\ell+1}}\phi_\ell(\xi), \ell = 0, 1, \ldots$.

To apply DG, an approximation space, $V_h(I_j)$, is used, given by,

$$V_h(I_j) = \{v : v \in \Phi^k(I_j)\}, j = 0, \ldots, N, \tag{5}$$

where $\Phi^k(I_j)$ is the space of scaled Legendre polynomials on element I_j of degree $k \in \mathbb{N}$. The weak formulation (Cockburn, [8]), is computed by multiplying equation (1a) by an arbitrary, smooth function $v \in C^1[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, and integrating over I_j, $j \in \{0, \ldots, N\}$. Using integration by parts, this gives,

$$0 = \int_{I_j} (u_t + f(u)_x)v dx = \int_{I_j} u_t v dx + f(u)v \,|_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} - \int_{I_j} f(u)v_x dx. \tag{6}$$

Next, v is replaced by a test function $v_h \in V_h(I_j)$, and the exact solution u by the approximate solution $u_h \in V_h(I_j)$. Using local coordinates $\xi = \frac{2}{\Delta x}(x - x_j)$, and an approximation space of degree k, these functions are given by,

$$v_h(x) = \phi_m(\xi), \qquad m \in \{0, \ldots, k\}, \tag{7a}$$

$$u_h(x, t) = \sum_{\ell=0}^{k} u_j^{(\ell)}(t)\phi_\ell(\xi), \text{ on element } I_j, \tag{7b}$$

where $u_j^{(\ell)}(t), \ell = 0, \ldots, k$, are the unknown DG coefficients. Using equations (7) in the weak formulation as given in (6), gives,

$$\int_{I_j} u_{h,t}(\xi)\phi_m(\xi)dx = \int_{I_j} f(u_h(\xi))\frac{d}{dx}\phi_m(\xi)dx - f(u_h)v_h \,|_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}}. \tag{8}$$

Furthermore, a change of coordinates is performed, such that $\xi = \frac{2}{\Delta x}(x - x_j)$. It follows that $\frac{\Delta x}{2}d\xi = dx$ and $\frac{d}{d\xi}\frac{d\xi}{dx} = \frac{2}{\Delta x}\frac{d}{d\xi}$, and equation (8) transforms into (using the orthogonality property of the scaled Legendre polynomials):

$$\frac{\Delta x}{2}\frac{du_j^{(m)}}{dt} = \int_{-1}^{1} f(u_h(\xi))\phi_m'(\xi)d\xi + \hat{F}_{j-\frac{1}{2}}v_{h,j-\frac{1}{2}}^{+} - \hat{F}_{j+\frac{1}{2}}v_{h,j+\frac{1}{2}}^{-}, \tag{9}$$

where $\hat{F}_{j\pm\frac{1}{2}}$ denote the fluxes through the boundaries $x_{j\pm\frac{1}{2}}$.

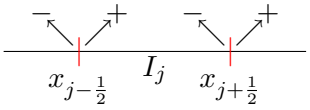


Figure 2: Boundaries of interval $I_j, j = 0, \ldots, N$

The test function v_h on the boundaries $x_{j\pm\frac{1}{2}}$ is taken from inside the cell, given by $v^+_{h,j-\frac{1}{2}}$ and $v^-_{h,j+\frac{1}{2}}$. The minus sign refers to the left side of the boundary, the plus sign belongs to the right side, as can be seen in Figure 2.
This means we use,

$$v^+_{h,j-\frac{1}{2}} = \phi_m(-1) \equiv (-1)^m \sqrt{m+\frac{1}{2}}, \tag{10a}$$

$$v^-_{h,j+\frac{1}{2}} = \phi_m(1) \equiv \sqrt{m+\frac{1}{2}}. \tag{10b}$$

In the weak form, that is given in equation (9), the fluxes at the left and right boundary of element I_j appear. This presents an ambiguity as the choice of the basis and test function is discontinuous at the points $x_{j\pm\frac{1}{2}}$. In general, the local Lax Friedrichs flux is used [18], which is defined as,

$$\hat{F}_{j-\frac{1}{2}} = \frac{1}{2}\left(f(u^-_{h,j-\frac{1}{2}}) + f(u^+_{h,j-\frac{1}{2}}) - a_{j-\frac{1}{2}}(u^+_{h,j-\frac{1}{2}} - u^-_{h,j-\frac{1}{2}}) \right), \tag{11a}$$

where,

$$a_{j-\frac{1}{2}} = \max(|f'(u_h)|) \text{ over all } u_h \text{ between } u^-_{h,j-\frac{1}{2}} \text{ and } u^+_{h,j-\frac{1}{2}}. \tag{11b}$$

If f is convex, this reduces to,

$$a_{j-\frac{1}{2}} = \max(|f'(u^-_{h,j-\frac{1}{2}})|, |f'(u^+_{h,j-\frac{1}{2}})|), \tag{11c}$$

which makes it easy to approximate the fluxes through the boundaries.
Note that from the definition for $\hat{F}_{j-\frac{1}{2}}$, it follows that the flux through $x_{j-\frac{1}{2}}$ depends on $u^{(\ell)}_{j-1}$ and $u^{(\ell)}_j, \ell = 0, \ldots, k$. Likewise, $\hat{F}_{j+\frac{1}{2}}$ uses $u^{(\ell)}_j$ and $u^{(\ell)}_{j+1}, \ell = 0, \ldots, k$.
The weak formulation in equation (9) is an ordinary differential equation in time for the DG coefficients $u^{(m)}_j, j = 0, \ldots, N, m = 0, \ldots, k$. If we define

$$\mathbf{u}_j = \left(u^{(0)}_j, u^{(1)}_j, \ldots, u^{(k)}_j\right)^\top, \quad j = 0, \ldots, N,$$

then we can rewrite equation (9) as $\frac{d}{dt}\mathbf{u}_j = L(\mathbf{u}_{j-1}, \mathbf{u}_j, \mathbf{u}_{j+1})$.
In Section 1.2, the time stepping method is explained.

1.2 Time stepping using TVD Runge Kutta

In Section 1.1, a system of ordinary differential equations in time was formed. These equations are solved using a third order total variation diminishing (TVD) Runge Kutta method (Gottlieb, [11]). If the approximation at time $t^s = s\Delta t, s = 0, 1, 2, \ldots,$ for element I_j is given by $\mathbf{w}_j^s, j \in \{0, \ldots, N\}$ then the approximation $\mathbf{w}_j^{s+1}$ at time t^{s+1} is computed using:

$$\begin{array}{rcl} \mathbf{w}_j^* & = & \mathbf{w}_j^s + \Delta t L(\mathbf{w}_{j-1}^s, \mathbf{w}_j^s, \mathbf{w}_{j+1}^s), \\ \mathbf{w}_j^{**} & = & \frac{3}{4}\mathbf{w}_j^s + \frac{1}{4}\mathbf{w}_j^* + \frac{1}{4}\Delta t L(\mathbf{w}_{j-1}^*, \mathbf{w}_j^*, \mathbf{w}_{j+1}^*), \\ \mathbf{w}_j^{s+1} & = & \frac{1}{3}\mathbf{w}_j^s + \frac{2}{3}\mathbf{w}_j^{**} + \frac{2}{3}\Delta t L(\mathbf{w}_{j-1}^{**}, \mathbf{w}_j^{**}, \mathbf{w}_{j+1}^{**}). \end{array} \tag{12}$$

Examples of the DG method, together with the TVD RK time stepping, are given in Appendix A.

1.3 Multiwavelets

It is possible to use multiwavelets while computing discontinuous Galerkin solutions. Multiwavelets are defined on $(-1, 1]$, but we can apply this theory for DG on a general interval $[a, b]$, which will be described later on.
The use of multiwavelets for DG is described by Alpert and Archibald [3, 4]. For $p = 0, 1, \ldots$, and $n = 0, 1, \ldots$, the space of piecewise polynomial functions, V_n^p, is defined as,

$$V_n^p = \{f : f \in \mathbb{P}^{p+1}(I_j^n), j = 0, \ldots, 2^n - 1\}, \tag{13}$$

where,

$$I_j^n = (-1 + 2^{-n+1}j, -1 + 2^{-n+1}(j+1)], \tag{14}$$

and $\mathbb{P}^{p+1}(I_j^n)$ is the space of all polynomials of degree less than $p+1$ on interval I_j^n.
A visualization of the intervals in $V_0^p, V_1^p, \ldots$ can be seen in Figure 3.

The space V_n^p has dimension $2^n(p+1)$ and the following nested property holds,

$$V_0^p \subset V_1^p \subset \cdots \subset V_n^p \subset \cdots .$$

The scale of the basis functions for V_j^p is too coarse for the details of the functions in V_{j+1}^p: the supported intervals in V_{j+1}^p are smaller than the intervals in V_j^p, such that more resolution is possible[1].

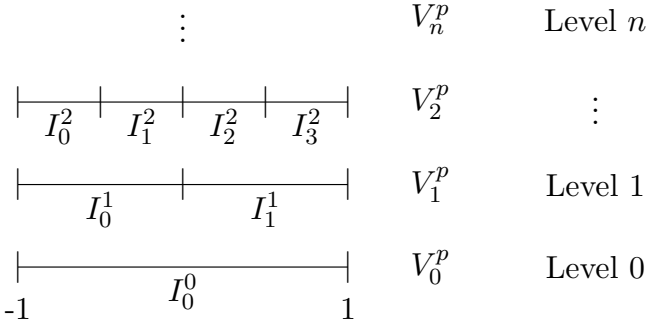


Figure 3: Visualization of the intervals in $V_0^p, V_1^p, \ldots$

In particular,

$$V_0^p = \{f \text{ is a polynomial of degree less than } p+1 \text{ on the interval } (-1, 1]\}.$$

In this report, we choose the scaled Legendre polynomials $\phi_0, \ldots, \phi_p$ from Section 1.1 to be the orthonormal basis for V_0^p. The space V_n^p is spanned by $2^n(p+1)$ functions which are obtained from $\phi_0, \ldots, \phi_p$ by dilation and translation,

$$\phi_{\ell j}^n(x) = 2^{n/2}\phi_\ell(2^n(x+1) - 2j - 1), \ell = 0, \ldots, p,\ j = 0, \ldots, 2^n - 1, x \in I_j^n, \tag{15}$$

[1]See the website `http://fourier.eng.hmc.edu/e161/lectures/wavelets/wavelets.html`

where the coefficient j belongs to the various intervals I_j^n, see also Keinert, [14]. The factor $2^{n/2}$ makes this an orthonormal basis for V_n^p. This can be proven using that ϕ_ℓ vanishes outside the interval $(-1,1]$: $2^n(x+1)-2j-1 \in (-1,1]$ for

$$x \in (-1+2^{-n+1}j, -1+2^{-n+1}(j+1)].$$

This means that,

$$\begin{aligned}\langle \phi_{\ell j}^n, \phi_{\ell j}^n \rangle &= \int_{-1}^{1} (\phi_{\ell j}^n(x))^2 dx = 2^n \int_{-1+2^{-n+1}j}^{-1+2^{-n+1}(j+1)} (\phi_\ell(2^n(x+1)-2j-1))^2 dx \\ &= \int_{-1}^{1} (\phi_\ell(y))^2 dy = 1.\end{aligned}$$

The functions $\phi_\ell, \ell = 0, \ldots, p$, are called scaling functions.

The multiwavelet subspace $W_n^p, n = 0, 1, 2, \ldots$, is defined as the orthogonal complement of V_n^p in V_{n+1}^p:

$$V_n^p \oplus W_n^p = V_{n+1}^p,\ W_n^p \perp V_n^p,\ W_n^p \subset V_{n+1}^p. \tag{16}$$

Hereby, the idea of the orthogonal decomposition theorem from Linear Algebra is used, see Strang, [23].
Due to definition (16), the orthonormal basis for W_0^p is given by $p+1$ piecewise polynomials, $\psi_0, \ldots, \psi_p$ (polynomials on $(-1,0]$ and $(0,1]$).
Similar to the basis for V_n^p, the space W_n^p is spanned by the functions,

$$\psi_{\ell j}^n(x) = 2^{n/2}\psi_\ell(2^n(x+1)-2j-1), \ell = 0, \ldots, p,\ j = 0, \ldots, 2^n-1, x \in I_j^n.$$

The functions $\psi_0, \ldots, \psi_p$ are the so-called multiwavelets. Appendix B contains an example of these multiwavelets, which are deduced from the scaling functions (the scaled Legendre polynomials).
Multiwavelets form a set of functions which, together with the scaling functions, can be used to approximate a function. The term multiwavelet refers to the fact that the bases for V_0^p and W_0^p contain multiple elements.

From definition (16) it follows that V_n^p can be split into $n+1$ orthogonal subspaces as,

$$V_n^p = V_0^p \oplus W_0^p \oplus W_1^p \oplus \cdots \oplus W_{n-1}^p.$$

Note that following the definition given in (15), for $n \in \mathbb{N}, j \in \{0, \ldots, 2^{n+1}-1\}$, $\ell \in \{0, \ldots, p\}$, the function $\phi_{\ell j}^{n+1}$ is narrower than $\phi_{\ell j}^n, j \in \{0, \ldots, 2^n-1\}$, and is translated in smaller steps. Therefore, it is possible to represent finer detail.
The same discussion holds for the function $\psi_{\ell j}^{n+1}$ with respect to $\psi_{\ell j}^n$, see Burrus [7].

The orthogonal projection of an arbitrary function $f \in L^2(-1,1)$ onto $V_n^p, n \in \mathbb{N}$ is given by,

$$P_n^p f(x) = \sum_{j=0}^{2^n-1} \sum_{\ell=0}^{p} s_{\ell j}^n \phi_{\ell j}^n(x), \tag{17}$$

which is called the single-scale decomposition of the approximate solution on level n. The coefficients $s_{\ell j}^n$ are given by,

$$s_{\ell j}^n = \langle f, \phi_{\ell j}^n \rangle = \int_{-1+2^{-n+1}j}^{-1+2^{-n+1}(j+1)} f(x)\phi_{\ell j}^n(x)dx \tag{18a}$$

$$= 2^{\frac{n}{2}} \int_{-1+2^{-n+1}j}^{-1+2^{-n+1}(j+1)} f(x)\phi_\ell(2^n(x+1)-2j-1)dx \tag{18b}$$

$$= 2^{-\frac{n}{2}} \int_{-1}^{1} f(-1+2^{-n}(y+2j+1))\phi_\ell(y)dy, \tag{18c}$$

which is the standard orthogonal projection onto an orthonormal basis.
In Keinert [14] it is proven that for any $f \in L^2(-1,1), P_n^p f \to f$ in $L^2(-1,1)$ as $n \to \infty$. Note that if $f \in V_n^p$, then $P_n^p f = f$.

Similarly, the function f has a multiwavelet expansion, given by (using equation (16)),

$$Q_n^p f(x) = P_{n+1}^p f(x) - P_n^p f(x) = \sum_{j=0}^{2^n-1} \sum_{\ell=0}^{p} d_{\ell j}^n \psi_{\ell j}^n(x), \tag{19}$$

which uses the multiwavelets $\psi_{\ell j}^n, \ell = 0,\ldots,p, j = 0,\ldots,2^n-1$. The coefficients are given by,

$$d_{\ell j}^n = \langle f, \psi_{\ell j}^n \rangle = \int_{-1+2^{-n+1}j}^{-1+2^{-n+1}(j+1)} f(x)\psi_{\ell j}^n(x)dx. \tag{20}$$

Using equation (19) recursively, a relation between expansions at different levels can be found:

$$P_{n+1}^p f(x) = P_n^p f(x) + Q_n^p f(x) = P_{n-1}^p f(x) + Q_{n-1}^p f(x) + Q_n^p f(x) \tag{21a}$$

$$= \ldots = P_0^p f(x) + \sum_{m=0}^{n} Q_m^p f(x) \tag{21b}$$

$$= \sum_{\ell=0}^{p} s_{\ell 0}^0 \phi_\ell(x) + \sum_{m=0}^{n} \sum_{j=0}^{2^m-1} \sum_{\ell=0}^{p} d_{\ell j}^m \psi_{\ell j}^m(x) \tag{21c}$$

$$= S^0(x) + \sum_{m=0}^{n} \sum_{j=0}^{2^m-1} D_j^m(x) \tag{21d}$$

$$= S^0(x) + \sum_{m=0}^{n} D^m(x). \tag{21e}$$

This representation of $P_{n+1}^p f(x)$ is called the multiscale decomposition. The coefficients $\{s_{\ell 0}^0\}_{\ell=0}^p$ represent the approximate solution on the coarsest level $n=0$, and $\{d_{\ell j}^m\}$ carry the multiscale information. The detail coefficients can be seen as carriers of individual

fluctuations of the solution, which, if added to the lowest-resolution information, enrich it up to the level $n+1$ of resolution (see Iacono, [13]).

Starting with the scaling function coefficients, $s^n_{\ell j}$, on level $n \in \mathbb{N}$, where $\ell = 0, \ldots, p$, $j = 0, \ldots, 2^n - 1$, it is possible to compute the coefficients $s^0_{\ell 0}, d^m_{\ell j}$, for $m = 0, \ldots, n-1$, $j = 0, \ldots, 2^m - 1$, using the multiwavelet decomposition, which is visualized in Figure 4. Decomposition is going top-down: starting in level n, all coefficients in the levels $n-1, \ldots, 0$ can be computed.

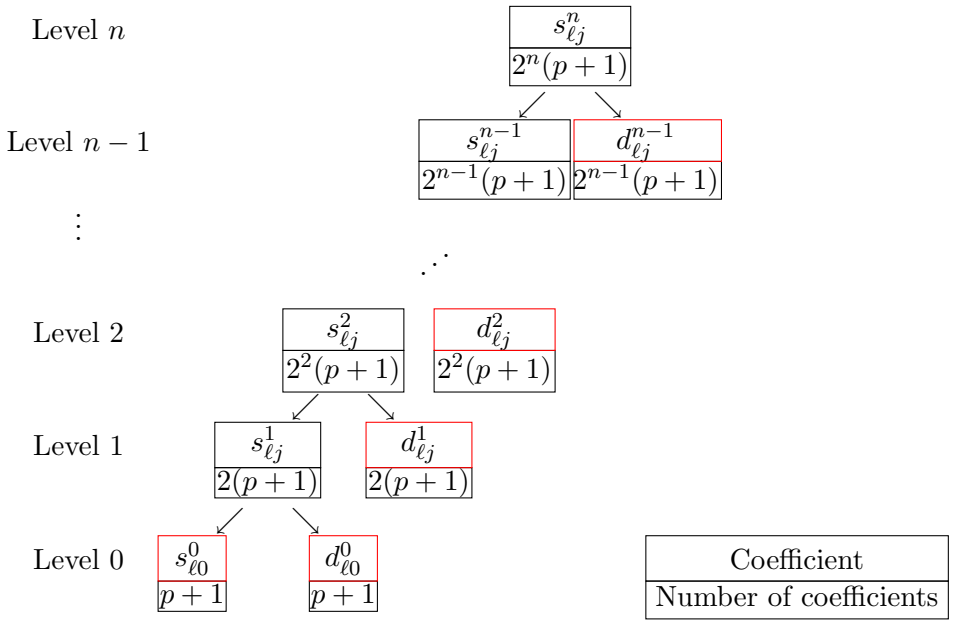


Figure 4: Visualization of multiwavelet decomposition. Red colored coefficients together carry the same information as $s^n_{\ell j}, \ell = 0, \ldots, p, j = 0, \ldots, 2^n - 1$

The red colored coefficients in Figure 4 together carry the same information as the scaling function coefficients on level n, given by $s^n_{\ell j}, \ell = 0, \ldots, p, j = 0, \ldots, 2^n - 1$. The decomposition and reconstruction methods are worked out in Appendix B.

Because the DG approximation and the single-scale decomposition of equation (17) are composed of the same basis functions, it is interesting to look at the relation between DG and multiwavelets. Therefore, the number of elements in $[a, b]$ is chosen such that $N + 1 = 2^n$, with boundaries (using definition (14))

$$\left(a + 2^{-n+1} j, a + 2^{-n+1}(j+1)\right], j = 0, \ldots, 2^n - 1.$$

The DG approximation of the solution is given by,

$$u_h(x,t) = \sum_{\ell=0}^{k} u_j^{(\ell)} \phi_\ell \left(\frac{2}{\Delta x}(x - x_j) \right), \tag{22}$$

for $x \in I_j, j = 0, \ldots, N$. Note that $\phi_\ell \left(\frac{2}{\Delta x}(x - x_j) \right)$ is only nonzero if $x \in I_j$, such that equation (22) equals,

$$u_h(x,t) = \sum_{j=0}^{N} \sum_{\ell=0}^{k} u_j^{(\ell)} \phi_\ell \left(\frac{2}{\Delta x}(x - x_j) \right), \tag{23}$$

for every $x \in [a, b]$.
Note that $\Delta x = \frac{b-a}{N+1} = \frac{b-a}{2^n}$, such that (using definition (2)) equation (23) can be written as,

$$\begin{aligned} u_h(x,t) &= \sum_{j=0}^{N} \sum_{\ell=0}^{k} u_j^{(\ell)} \phi_\ell \left(\frac{2}{\Delta x} \left(x - \left(a + \left(j + \frac{1}{2} \right) \Delta x \right) \right) \right) \\ &= \sum_{j=0}^{N} \sum_{\ell=0}^{k} u_j^{(\ell)} \phi_\ell \left(\frac{2^{n+1}}{b-a}(x-a) - 2j - 1 \right) \\ &= \sum_{j=0}^{N} \sum_{\ell=0}^{k} u_j^{(\ell)} \phi_\ell (2^n(y+1) - 2j - 1) \\ &= 2^{-\frac{n}{2}} \sum_{j=0}^{N} \sum_{\ell=0}^{k} u_j^{(\ell)} \phi_{\ell j}^n(y), \end{aligned} \tag{24}$$

where $y = -1 + 2\frac{x-a}{b-a}$ and definition (15) is used in line (24).
Note that $u_h(x,t)$ is a piecewise polynomial of degree k. If we transform the interval $[a, b]$ to $[-1, 1]$, using $y = -1 + 2\frac{x-a}{b-a}$, then we find,

$$u_h(x,t) = P_n^k u_h(x,t) = \sum_{j=0}^{2^n-1} \sum_{\ell=0}^{k} s_{\ell j}^n \phi_{\ell j}^n(y). \tag{25}$$

From equation (24) and (25) it follows that for every $\ell = 0, \ldots, k, j = 0, \ldots, 2^n - 1$,

$$2^{-\frac{n}{2}} u_j^{(\ell)} = s_{\ell j}^n. \tag{26}$$

The scaling function coefficients on level n can therefore be computed using the DG coefficients.

1.4 Nonlinear partial differential equations in 1D

In this section, several examples of nonlinear partial differential equations are given. In general, the exact solution of a nonlinear equation contains discontinuities, like shocks, contact discontinuities, and rarefaction waves. This is the reason why it is not easy to solve nonlinear equations numerically.
In Section 1.4.1, the inviscid Burgers' equation is taken into account, and the characteristics of the exact solution are considered. In Section 1.4.2, the Euler equations are given, together with two different initial conditions (Sod, Lax).

1.4.1 Inviscid Burgers' equation

The inviscid Burgers' equation on $[-1,1]$, is given by,

$$u_t + (f(u))_x = u_t + \left(\frac{u^2}{2}\right)_x = 0, \quad x \in [-1,1], t \geq 0, \tag{27a}$$

$$u(x,0) = u^0(x), \qquad x \in [-1,1], \tag{27b}$$

where $u = u(x,t)$ and periodic boundary conditions are assumed. The initial condition is given by Krivodonova, [15]:

$$u^0(x) = \frac{1}{2} + \frac{1}{2}\sin(\pi x). \tag{28}$$

An analysis of the exact solution of the inviscid Burgers' equation is given by LeVeque, [17], and Smoller, [21]. The characteristics of Burgers' equation satisfy,

$$x'(t) = u(x(t),t),$$

where u is constant along these characteristics. This can be seen by looking at,

$$\begin{aligned}\frac{d}{dt}u(x(t),t) &= \frac{\partial}{\partial t}u(x(t),t) + \frac{\partial}{\partial x}u(x(t),t)x'(t)\\ &= u_t + uu_x = 0.\end{aligned}$$

The solution u being constant, means that the slope $x'(t)$ is constant and that the characteristics are straight lines in the (x,t)-plane. Therefore,

$$\begin{aligned}x(t) &= x(0) + u(x(t),t)t\\ &= x(0) + u(x(0),0)t,\end{aligned} \tag{29}$$

where $x(0) \in [-1,1]$ can be chosen.
For nonlinear equations (such as the inviscid Burgers' equation), it is possible that the characteristics intersect. This would happen if for some $x_1 < x_2$, it holds that $x_1 + u(x_1,0)t = x_2 + u(x_2,0)t$. If the initial data $u(x,0)$ is smooth, this implies,

$$t = -\frac{x_2 - x_1}{u(x_2,0) - u(x_1,0)} \to -\frac{1}{u'(x_1,0)} \text{ if } x_2 \to x_1.$$

It is clear that characteristics intersect only if $u'(x,0)$ is negative for some $x \in [-1,1]$, and the first time this happens is at,

$$T_b = -\frac{1}{\min\limits_{x\in[-1,1]} u'(x,0)}. \tag{30}$$

For the initial condition given in (28), the first time characteristics intersect is at,

$$T_b = -\frac{1}{\min_{x\in[-1,1]} \frac{\pi}{2}\cos(\pi x)} = \frac{2}{\pi} \approx 0.6366.$$

A plot of the characteristics is given in Figure 5a. Following Thomas, [24], this discontinuity is a shock, because the characteristics on both sides of the discontinuity impinge on the discontinuity curve in the direction of increasing t.
Note that $\{y : u'(y,0) = \min_{x\in[-1,1]} \frac{\pi}{2}\cos(\pi x)\} = \{-1,1\}$. The periodic boundary conditions imply that $x(0) = -1$ and $x(0) = 1$ lead to the same shock location, given by,

$$\begin{aligned} x(T_b) = x\left(\frac{2}{\pi}\right) &= x(0) + \left(\frac{1}{2} + \frac{1}{2}\sin(\pi x(0))\right)\frac{2}{\pi}, \\ &= -1 + \left(\frac{1}{2} + \frac{1}{2}\sin(-\pi)\right)\frac{2}{\pi} = -1 + \frac{1}{\pi} \approx -0.6817, \end{aligned}$$

which is indicated by the red star in Figure 5a.
The shock will propagate with a speed $s(t)$ that changes with time. Using the Rankine-Hugoniot jump condition (LeVeque, [17]), the shock speed can be determined in terms of the values $u_l(t)$ and $u_r(t)$ immediately to the left and right of the shock, and is given by,

$$s(t) = \frac{f(u_r(t)) - f(u_l(t))}{u_r(t) - u_l(t)}.$$

Note that, for the inviscid Burgers' equation, this means,

$$s(t) = \frac{\frac{1}{2}(u_r(t)^2 - u_l(t)^2)}{u_r(t) - u_l(t)} = \frac{u_l(t) + u_r(t)}{2}.$$

The exact solution of equations (27) is given by,

$$u(x,t) = u^0(x - u(x,t)t), \tag{31}$$

and is computed using the bisection method, described by Burden [6]. Therefore, equation (31) is written as,

$$g(u) = u - u^0(x - ut) = 0, \tag{32}$$

for any $x \in [-1,1], t > 0$. The bisection method starts with an interval $[a_1, b_1]$, where $g(a_1)$ and $g(b_1)$ are of opposite sign. Following the Intermediate Value Theorem, there exists a $u \in [a_1, b_1]$ such that $g(u) = 0$. The first try is $u_1 = \frac{a_1+b_1}{2}$ (the midpoint of $[a_1, b_1]$). If $g(u_1) = 0$, then $u = u_1$ is the solution of equation (32). If this is not the case, then there are two possibilities:

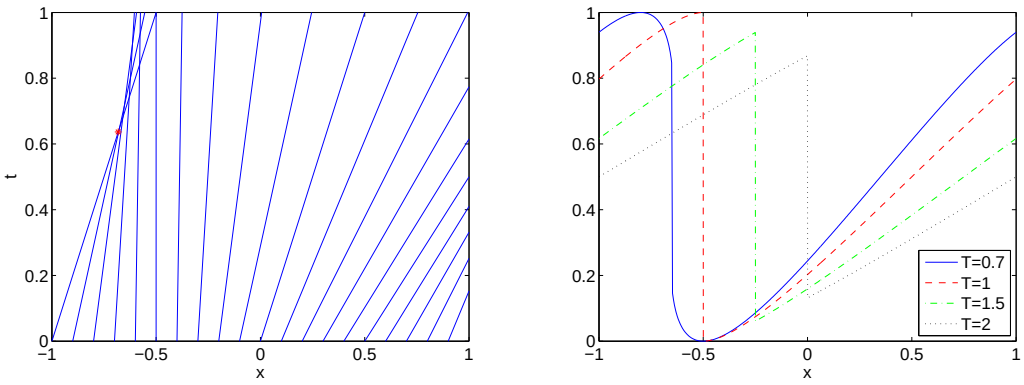
1. $g(u_1)$ and $g(a_1)$ have the same sign: in that case the root u should be in $[u_1, b_1]$. Therefore, $a_2 = u_1$ and $b_2 = b_1$;

2. $g(u_1)$ and $g(b_1)$ have the same sign: u should be in $[a_1, u_1]$, and $a_2 = a_1, b_2 = u_1$;

These steps are performed recursively, until for some $M \in \mathbb{N}$, and $\epsilon > 0$, it holds that $|a_M - b_M| < \epsilon$ or the maximum number of iterations is reached. In Figure 5b the exact solution can be seen for various times.

For times $t > T_b$, there is a point x for which relation (31) has two different solutions (the shock, see Figure 5b). This does not make sense in physical situations. The correct physical behaviour is determined by using the viscid Burgers' equation, which is given by $u_t + uu_x = \varepsilon u_{xx}$, where $\varepsilon > 0$ is the viscosity coefficient. This extra term avoids from getting a multivalued solution, and approaches the solution of the inviscid equation as $\varepsilon \to 0$.

The numerical solution is derived using the discontinuous Galerkin method (Section 1.1), combined with the third order TVD RK method (Section 1.2). In Appendix A.3, we define this method for the inviscid Burgers' equation, and give some results.



(a) Characteristics. Red star: location where shock is generated

(b) Exact solution using bisection

Figure 5: Characteristics and exact solution of inciscid Burgers' equation, initial condition (28)

1.4.2 Euler equations

In this section, the Euler equations are introduced. A nice introduction to gas dynamics is given in LeVeque, [17]. Here, three equations play a role:

1. The continuity equation, which models conservation of mass, given by,

$$\rho_t + (\rho u)_x = 0, \tag{33a}$$

   where $\rho(x,t)$ is the density and $u(x,t)$ is the velocity of the gas, which, in general, are both unknown.

2. Conservation of momentum, ρu, using a macroscopic, convective momentum flux and a microscopic flux due to the pressure of the fluid, $p(x,t)$:

$$(\rho u)_t + (\rho u^2 + p)_x = 0. \tag{33b}$$

3. Conservation of energy, $E(x,t)$, using the macroscopic energy flux and a flux in kinetic energy:

$$E_t + ((E+p)u)_x = 0. \tag{33c}$$

   Here, the equation of state for an ideal polytropic gas is used:

$$E = \frac{p}{\gamma - 1} + \frac{1}{2}\rho u^2, \tag{33d}$$

   where the term $\frac{p}{\gamma-1}$ belongs to the internal energy, and $\frac{1}{2}\rho u^2$ is the kinetic energy. Finally, $\gamma = \frac{c_p}{c_v}$ is called the adiabatic exponent, where c_p denotes the specific heat at constant pressure, and c_v is the specific heat at constant volume. For air, γ is approximately equal to 1.4.

Introducing $\mathbf{u} = (\rho, \rho u, E)^\top = (u^{(1)}, u^{(2)}, u^{(3)})^\top$, equations (33) can be written as,

$$\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x = 0, \tag{34a}$$

where,

$$\begin{aligned}\mathbf{f}(\mathbf{u}) &= (\rho u, \rho u^2 + p, u(E+p))^\top \\ &= \left(u^{(2)}, \frac{3-\gamma}{2}\frac{(u^{(2)})^2}{u^{(1)}} + (\gamma-1)u^{(3)}, \frac{u^{(2)}}{u^{(1)}}\left(\gamma u^{(3)} - \frac{\gamma-1}{2}\frac{(u^{(2)})^2}{u^{(1)}}\right)\right)^\top.\end{aligned} \tag{34b}$$

In Appendix A.4, the corresponding DG method is given, together with numerical results.

As an example, we look at a shock tube (Sod, [22]). A diaphragm halfway the tube separates two regions, which have different densities and pressures. These two regions are in a constant state, and both fluids are initially at rest. The following initial conditions are used:

$$\rho(x,0) = \begin{cases} 1, & \text{for } x < 0; \\ 0.125, & \text{for } x \geq 0; \end{cases} \quad p(x,0) = \begin{cases} 1, & \text{for } x < 0; \\ 0.1, & \text{for } x \geq 0; \end{cases} \tag{35a}$$

and

$$u(x,0) \equiv 0, \tag{35b}$$

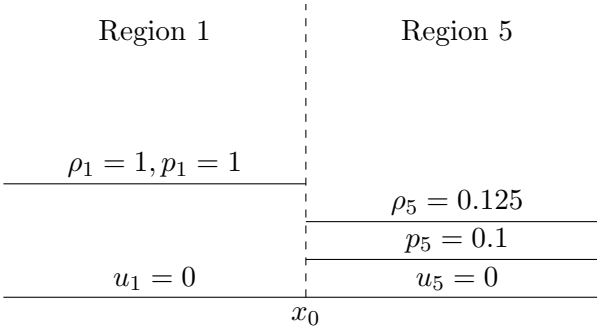
see Figure 6a.
Using that $\gamma = 1.4$, the energy at $T = 0$ is given by,

$$E(x,0) = \frac{p(x,0)}{\gamma - 1} + \frac{1}{2}\rho(x,0)u(x,0)^2 = \begin{cases} 2.5, & \text{for } x < 0; \\ 0.25, & \text{for } x \geq 0. \end{cases} \tag{35c}$$

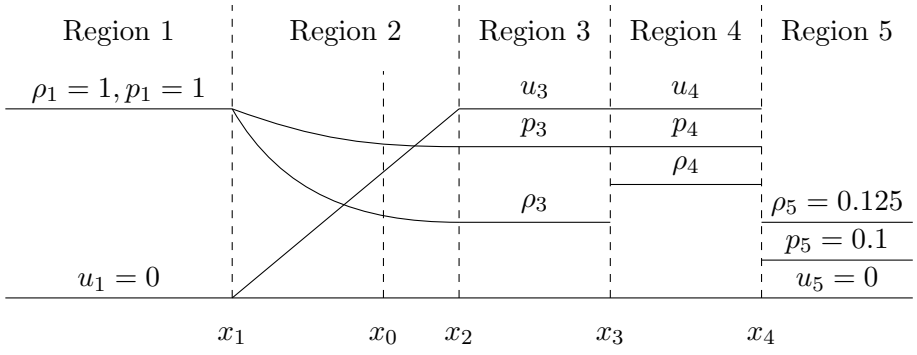
At time $t > 0$ the diaphragm is broken.
The physical domain is assumed to be very long (essentially infinite). The computational domain, however, must be finite, and is set equal to $[-5, 5]$. The boundary conditions must be chosen such that the obtained results using this smaller domain resemble the results, computed on a larger domain. These so-called absorbing or nonreflecting boundary conditions should allow outgoing waves to disappear without generating spurious incoming waves, and are given by (LeVeque, [18]):

$$\rho(x,t) = \begin{cases} 1, & x \leq -5; \\ 0.125, & x \geq 5; \end{cases} \quad p(x,t) = \begin{cases} 1, & x \leq -5; \\ 0.1, & x \geq 5; \end{cases} \quad u(x,t) = 0, \ x \leq -5, x \geq 5, t \geq 0.$$

In Figure 6b the exact results for the shock tube are given for some $t > 0$. This figure is given in the report of Sod, [22], and is analyzed with the help of Smoller, [21].
In Region 2, a rarefaction wave is situated, where both the density and the pressure decrease continuously from (ρ_1, p_1) to (ρ_3, p_3). In this region, some of the derivatives of the fluid quantities may not be continuous. The head x_1 and tail x_2 of the rarefaction wave move to the left with time.
Coordinate x_3 belongs to a contact discontinuity, due to the original discontinuity in the data. This discontinuity is also present in the linear approximations to the equations, and it is the position that an element of fluid initially at x_0 has reached by time t. Across a contact discontinuity the pressure and the normal component of velocity are continuous, whereas the density and the energy are discontinuous.
The point x_4, which moves to the right, is the location of the shock wave: all quantities are discontinuous across x_4. The shock wave is formed due to the nonlinearities in the equations (recall the shock wave in Burgers' equation, Section 1.4.1).



(a) $t = 0$



(b) $t > 0$

Figure 6: Sod's shock tube, divided into different regions:
Region 1: Gas in original state of high pressure;
Region 2: Rarefaction wave, solution is a monotone continuous function of $\frac{x}{t}$;
Region 3: Rarefied gas (gas at lower pressure than in Region 1);
Region 4: Compressed gas (gas at higher density than in Region 5);
Region 5: Gas in original state of low pressure.

Another set of initial conditions for the shock tube is given by Lax, [16], and equals:

$$\rho(x,0) = \begin{cases} 0.445, & \text{for } x < 0; \\ 0.5, & \text{for } x \geq 0; \end{cases} \quad u(x,0) = \begin{cases} 0.698, & \text{for } x < 0; \\ 0, & \text{for } x \geq 0; \end{cases} \tag{36a}$$

and

$$p(x,0) = \begin{cases} 3.528, & \text{for } x < 0; \\ 0.571, & \text{for } x \geq 0, \end{cases} \tag{36b}$$

The shock tube, together with Sod's or Lax's initial conditions, is used in Sections 4.2 and 4.3, to test the limiting techniques that we develop in Section 3.

1.5 Limiters

In practical applications, initial conditions may contain discontinuities (Sod's and Lax's shock tubes, Section 1.4), or nonlinear equations with a smooth initial condition may develop a discontinuity (inviscid Burgers' equation, and the shock tube). In that case, higher order accurate methods fail near these discontinuities, and unphysical oscillations may be generated. Lower order methods have the advantage of keeping the solution monotonically varying in regions where the solution should be monotone, even though the accuracy is not very good, see LeVeque, [18].
Monotonicity has to do with the average $\overline{u}_j^s$, of $u_h(x,t^s)$ in element $I_j, j=0,\ldots,N$, at time $t^s = s\Delta t, s = 0,1,\ldots$, and is given by,

$$\overline{u}_j^s = \frac{1}{\Delta x}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} u_h(x,t^s)dx.$$

Note that, using the orthonormal property of the scaled Legendre polynomials, for $\ell \in \{0,\ldots,k\}$,

$$\int_{-1}^{1}\phi_\ell(x)dx = \int_{-1}^{1}\phi_\ell(x)\cdot 1 dx = \sqrt{2}\int_{-1}^{1}\phi_\ell(x)\phi_0(x)dx = \begin{cases}\sqrt{2} & \text{if } \ell = 0,\\ 0 & \text{otherwise.}\end{cases}$$

Therefore, the average can be computed as follows (using equation (7b)),

$$\begin{aligned}\overline{u}_j^s &= \frac{1}{\Delta x}\int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}}\sum_{\ell=0}^{k} u_j^{(\ell)}(t^s)\phi_\ell\left(\frac{2}{\Delta x}(x-x_j)\right)dx\\ &= \frac{1}{\Delta x}\sum_{\ell=0}^{k} u_j^{(\ell)}(t^s)\cdot\frac{\Delta x}{2}\int_{-1}^{1}\phi_\ell(\xi)d\xi\\ &= \frac{1}{\sqrt{2}}u_j^{(0)}(t^s).\end{aligned} \tag{37}$$

A method is called monotonicity-preserving if,

$$\overline{u}_j^s \geq \overline{u}_{j+1}^s, \forall j \in \{0,\ldots,N\},$$

implies that

$$\overline{u}_j^{s+1} \geq \overline{u}_{j+1}^{s+1}, \forall j \in \{0,\ldots,N\}, s = 0,1,\ldots.$$

This monotone property aids lower order methods in preventing unphysical oscillations that higher order methods produce.

To illustrate this point, we consider the following initial condition for the linear advection equation (given by $u_t + u_x = 0$) on $[-1,1]$, with periodic boundary conditions:

$$u^0(x) = \begin{cases}\frac{1}{6}(G(x,\beta,z-\delta)+G(x,\beta,z+\delta)+4G(x,\beta,z)), & x\in[-0.75,-0.5],\\ 1, & x\in[-0.25,0],\\ 1-|8(x-\frac{3}{8})|, & x\in[0.25,0.5],\\ \frac{1}{6}(F(x,\alpha,a-\delta)+F(x,\alpha,a+\delta)+4F(x,\alpha,z)), & x\in[0.75,1],\\ 0, & \text{else,}\end{cases} \tag{38}$$

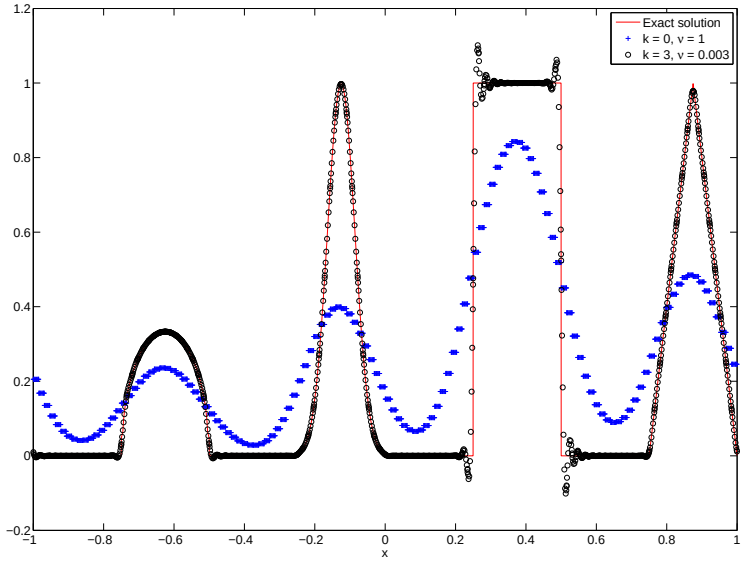


Figure 7: DG solution of linear advection equation on $[-1,1]$, initial condition (38), $T = 0.5$, 128 elements

where $G(x,\beta,z) = e^{-\beta(x-z)^2}, F(x,\alpha,a) = \sqrt{\max(1-\alpha^2(x-a)^2,0)}, a = \frac{7}{8}, z = -\frac{5}{8}$, $\delta = 0.005, \alpha = 8$ and $\beta = \frac{\log(2)}{36\delta^2}$.
In Figure 7 the behaviour of the DG solution using a lower ($k = 0$) and higher ($k \geq 1$) order method is shown. Clearly, the lower order method smears the solution and is not accurate, whereas the higher order method generates unphysical oscillations in the neighbourhood of the discontinuities.

The best features of both lower and higher order methods are combined in so-called high-resolution methods, such that the discontinuous portion of the solution remains nonoscillatory (lower order accurate) while the smooth portion remains accurate (higher order accurate). These high-resolution methods use a limiter that depend on how the solution is behaving. This technique is borrowed from finite volume methodology as described in LeVeque, [18].

Oscillations in the solution are measured by the notion of total variation of a function. In general, the total variation of the discrete function $\mathbf{w} = \{\ldots, w_{-1}, w_0, w_1, \ldots,\}$ is given by,

$$\mathrm{TV}(\mathbf{w}) = \sum_{j=-\infty}^{\infty} |w_j - w_{j-1}|,$$

see LeVeque, [18]. If a numerical method introduces oscillations, then it is expected that the total variation increases with time.

A time stepping method is called total variation diminishing (TVD) if, for any set of data $\mathbf{w}^s$ at time $t^s = s\Delta t, s = 0, 1, \ldots$, the values $\mathbf{w}^{s+1}$ computed by the method satisfy $\mathrm{TV}(\mathbf{w}^{s+1}) \leq \mathrm{TV}(\mathbf{w}^s)$.
The stability of a method can also be considered using total variation boundedness. A numerical method is total variation bounded (TVB) if, for any initial data $\mathbf{w}^0$ with $\mathrm{TV}(\mathbf{w}^0) < \infty$ and for any time T, there is a constant $R > 0$ and a value $\Delta t_0 > 0$ such that $\mathrm{TV}(\mathbf{w}^s) \leq R$ for all $s\Delta t \leq T, s = 0, 1, \ldots,$ whenever $\Delta t < \Delta t_0$ (see [18]).

High-resolution methods that use limiters are distinguished as total variation diminishing in the mean (TVDM), and total variation bounded in the mean (TVBM). Here, the averages of the approximation on each interval I_j, given by $\bar{u}_j^s$ $j = 0, \ldots, N, s = 0, 1, \ldots,$ are used.

Most limiters that are currently used for DG adapt the DG coefficients $u_j^{(\ell)}(t^s)$, computing a limited set $\widetilde{u}_j^{(\ell)}(t^s), j = 0, \ldots, N, \ell = 0, \ldots, k$. Therefore, the following algorithm is used (applying the third order TVD RK time stepping of Section 1.2):

Algorithm 1 DG and TVD RK, combined with limiter

```
for each RK step do
    Enforce boundary conditions;
    Compute limited solution;
    Enforce boundary conditions;
    Apply RK step;
end for
Enforce boundary conditions;
Compute limited solution;
Enforce boundary conditions.
```

Examples of TVDM and TVBM slope limiters are given in Cockburn [8, 9]. Several groups of limiters are discussed in this report, some of them briefly (minmod limiters, projection limiters, WENO limiter, and the multiwavelet limiter that is used today), some of them more intensely (moment limiter, maximum principle limiter).
Using 2^n elements on $[a, b]$, relation (26) allows us to translate each limiter on $u_j^{(\ell)}$ into a limiter for the scaling function coefficients $s_{\ell j}^n, \ell = 0, \ldots, k, j = 0, \ldots, 2^n - 1$.

1.5.1 Minmod limiters

Minmod limiters use the so-called minmod function, given by

$$\mathrm{minmod}(a_1, \ldots, a_q) = \begin{cases} \mathrm{sign}(a_1) \cdot \min_{1 \leq r \leq q} |a_r|, & \text{if sign } (a_1) = \cdots = \mathrm{sign}(a_q), \\ 0, & \text{otherwise.} \end{cases}$$

An example is the monotonized central-difference limiter, proposed by Van Leer and described in LeVeque, [18]: the approximation $u_h(x, t^s)$ of the solution at element I_j,

$j = 0, \ldots, N$, is written as a linear function of the form

$$u_h(x, t^s) = \overline{u}_j^s + \sigma_j^s (x - x_j),$$

where the average $\bar{u}_j^s$ is given by equation (37).
In Van Leer's limiter the slope, σ_j^s, is chosen to be,

$$\sigma_j^s = \text{minmod}\left(\left(\frac{\overline{u}_{j+1}^s - \overline{u}_{j-1}^s}{2\Delta x}\right), 2\left(\frac{\overline{u}_j^s - \overline{u}_{j-1}^s}{\Delta x}\right), 2\left(\frac{\overline{u}_{j+1}^s - \overline{u}_j^s}{\Delta x}\right)\right).$$

Since the solution is written as a linear function on element I_j, the limited DG approximation is given by,

$$\widetilde{u}(x, t^s) = \sum_{\ell=0}^{1} \widetilde{u}_j^{(\ell)}(t^s)\phi_\ell(\xi) = \bar{u}_j^s + \sigma_j^s (x - x_j).$$

Note that the method lowers the order in the neighbourhood of shocks: the approximation is only linear in these elements.

1.5.2 Projection limiters

Several projection limiters are developed by Cockburn and Shu, [9], and one of them is well described in Biswas, [5].
Here, approximation $u_h(x, t^s)$ at time t^s on element $I_j, j = 0, \ldots, N$, is written as,

$$u_h(x, t^s) = u_j^{(0)}(t^s)\phi_0(\xi) + \sum_{\ell=1}^{k} u_j^{(\ell)}(t^s)\phi_\ell(\xi) \equiv \frac{1}{\sqrt{2}} u_j^{(0)} + h_j(\xi, t^s),$$

where $h_j(\xi, t^s)$ contains the higher order terms in the DG approximation.
The limited values $\widetilde{h}_j(\pm 1, t^s)$, on the boundaries $x_{j\pm\frac{1}{2}}$, are given by,

$$\begin{aligned}
\widetilde{h}_j(1, t^s) &= \text{minmod}\left(h_j(1, t^s), \bar{u}_j^s - \bar{u}_{j-1}^s, \bar{u}_{j+1}^s - \bar{u}_j^s\right); \\
-\widetilde{h}_j(-1, t^s) &= \text{minmod}\left(-h_j(-1, t^s), \bar{u}_j^s - \bar{u}_{j-1}^s, \bar{u}_{j+1}^s - \bar{u}_j^s\right).
\end{aligned}$$

If $k = 1$ or $k = 2$, then $\widetilde{u}_j^{(1)}(t^s)$ and $\widetilde{u}_j^{(2)}(t^s)$ can be determined uniquely. If $k \geq 3$ and $\widetilde{h}_j(1, t^s) \neq h_j(1, t^s)$, or $\widetilde{h}_j(-1, t^s) \neq h_j(-1, t^s)$, then the coefficients $\widetilde{u}_j^{(\ell)}(t^s)$ are set to zero for $\ell \geq 3$.
The order of the approximation is lowered in the neighbourhood of shocks and the solution accuracy is reduced: only a first-order accuracy is obtained. Besides that, the limiting flattens solutions near smooth extrema where no limiting is required.

1.5.3 WENO

Another limiter is found by using the weighted essentially nonoscillatory (WENO) finite volume methodology as described in Qiu, [19]. WENO replaces the solution in the troubled cells with reconstructed polynomials which maintain the original cell averages, have the same orders of accuracy as before, but are less oscillatory. These new polynomials are reconstructed from the information of neighbouring cells.
The main difficulty is maintaining the original high order accuracy.

1.5.4 Multiwavelet limiter

The current multiwavelet limiter (Hovhannisyan, [12]) uses a threshold $\epsilon > 0$ such that if $\left|d_{\ell j}^m\right| < \epsilon$, then $d_{\ell j}^m \equiv 0, \ell = 0, \ldots, k, m = 0, \ldots, n-1, j = 0, \ldots, 2^m - 1, n \in \mathbb{N}$. Next, the scaling function coefficients, $s_{\ell j}^n, \ell = 0, \ldots, k, j = 0, \ldots, 2^n - 1$, are computed, and a projection limiter (Section 1.5.2) is used for these coefficients. The limited DG coefficients are found using equation (26).
Note that thresholding the multiwavelet coefficients removes small values only. Therefore, this limiter does not differ that much from the projection limiter itself, and works for low order methods only.
Originally, the goal of my master thesis was to find a new multiwavelet limiter, that works for higher order methods as well. Research on this subject is done in Section 2.

1.5.5 Moment limiter

The moment limiter, described by Krivodonova, [15], uses the same idea as the minmod limiter of Section 1.5.1, but is applied to the DG coefficients $u_j^{(\ell)}$, $j = 0, \ldots, N$, $\ell = k, \ldots, 1$, instead of the entire approximation. This saves on computational cost. The limiter works for higher order methods, but causes problems with multi-dimensions and complex geometries.
Krivodonova uses the unscaled Legendre polynomials instead of the scaled polynomials. Therefore, the definition below differs by a constant from [15].
The DG approximation at time t^s is given by, $u_h(x, t^s) = \sum_{\ell=0}^k u_j^{(\ell)}(t^s)\phi_\ell(\xi)$, for $x \in I_j$, $j = 0, \ldots, N$.

The moment limiter starts at $\ell = k$, and replaces $u_j^{(k)}$ with,

$$\widetilde{u}_j^{(k)} = \text{minmod}\left(u_j^{(k)}, \beta_k\left(u_{j+1}^{(k-1)} - u_j^{(k-1)}\right), \beta_k\left(u_j^{(k-1)} - u_{j-1}^{(k-1)}\right)\right),$$

with $\beta_k = \dfrac{\sqrt{k-\frac{1}{2}}}{\sqrt{k+\frac{1}{2}}}$. If $\widetilde{u}_j^{(k)} = u_j^{(k)}$, the limiting procedure is cut off for this element I_j. Otherwise, $u_j^{(k-1)}$ is limited, continuing until $u_j^{(1)}$ is limited ($u_j^{(0)}$ remains the same, such that the average $\bar{u}_j^s$ is preserved), or stopping the first time $\widetilde{u}_j^{(\ell)} = u_j^{(\ell)}$ for some $\ell = k-1, \ldots, 1$.

Let us use the moment limiter to compute the DG coefficients $\mathbf{u}_j^s, j = 0, \ldots, N$, at time $t^s = s\Delta t$, $s \in \{0, 1, \ldots\}$. Next, we apply the first RK step at time t^{s+1}, to compute $\mathbf{u}_j^*, j = 0, \ldots, N$ (Section 1.2). After forcing the boundary conditions (algorithm 1), the moment limiter is applied, which gives us $\widetilde{\mathbf{u}}_j^*, j = 0, \ldots, N$. The accompanying approximations are denoted as $u_h^*(x, t^{s+1})$ and $\widetilde{u}_h^*(x, t^{s+1})$.
The difference

$$\left|u_h^*(x, t^{s+1}) - \widetilde{u}_h^*(x, t^{s+1})\right| \tag{39}$$

becomes smaller if time increases: the moment limiter smoothes the solution mainly in the first time steps, and the limiter has less impact for later times. For the linear advection equation with initial condition (38), the orders of difference (39) are given in Table 1.

Time step	1	1000	5000
Difference	$\mathcal{O}(10^{-2})$	$\mathcal{O}(10^{-3})$	$\mathcal{O}(10^{-4})$

Table 1: Difference (39) at different times, $\Delta t = 9.3750 \cdot 10^{-5}$, linear advection equation, initial condition (38), $k = 3$, 64 elements

Similarly, we can compute the differences between the multiwavelet decompositions of $u_h^*(x, t^{s+1})$ and $\widetilde{u}_h^*(x, t^{s+1})$, given by,

$$\left|S^0(x) - \widetilde{S}^0(x)\right|, \quad \left|D^m(x) - \widetilde{D}^m(x)\right|, \quad m = 0, \ldots, n-1. \tag{40}$$

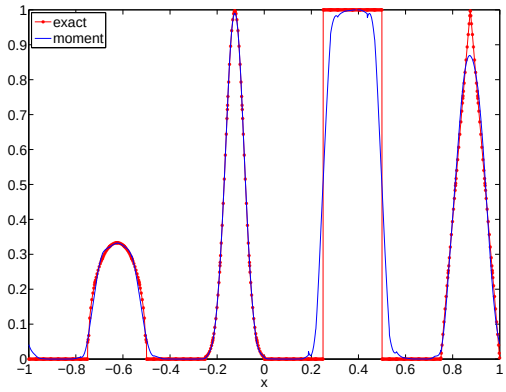
For the linear advection equation with initial condition (38), the orders of differences (40) after 5000 time steps ($T \approx 0.47$)are given in Table 2.

Scaling function	Multiwavelet, level					
	0	1	2	3	4	5
$\mathcal{O}(10^{-9})$	$\mathcal{O}(10^{-9})$	$\mathcal{O}(10^{-8})$	$\mathcal{O}(10^{-7})$	$\mathcal{O}(10^{-6})$	$\mathcal{O}(10^{-5})$	$\mathcal{O}(10^{-4})$

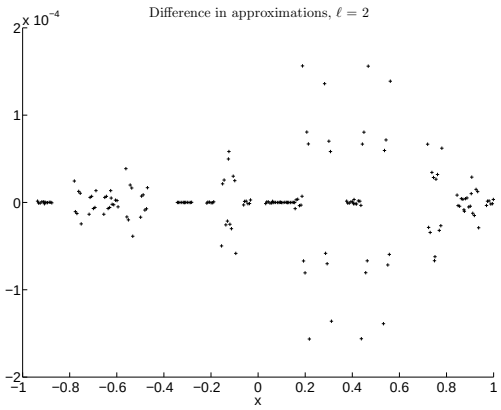
Table 2: Differences (40), linear advection equation with initial condition (38), time step 5000 ($T \approx 0.47$), $\Delta t = 9.3750 \cdot 10^{-5}$, $k = 3, n = 6$ ($2^n = 64$ elements)

It stands out that the moment limiter adapts the highest multiwavelet level at most, and less important are changes in the lower levels and the scaling function contribution. This property holds for all time steps.
If k DG coefficients are not limited in element $I_j, j \in \{0, \ldots, N\}$, then only one DG coefficient is limited, and this must be $u_j^{(k)}$, etc. Therefore the number of DG coefficients that are not limited equals the highest DG coefficient that is limited.
In Figure 8, difference (39) is split up into each limiting level, at $T = 0.5$ (linear advection equation with initial condition (38)). Hereby, only the points are plotted where the limited approximation differs from the original one.
If $\ell = k = 3$, then no coefficient is limited and the approximation after limiting is not different from the approximation before limiting.

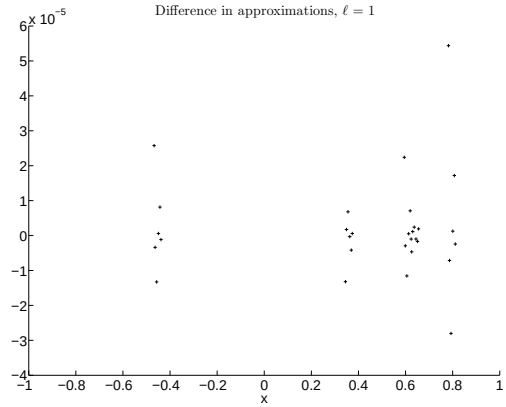
At $T = 0.5$ (after 5334 time steps, $\Delta t = 9.3750 \cdot 10^{-5}$), the moment limited solution is quite smooth in the neighbourhood of the discontinuities. Therefore, in most regions, it is convenient to limit $u_j^{(3)}$ only, as can be seen in Figure 8.



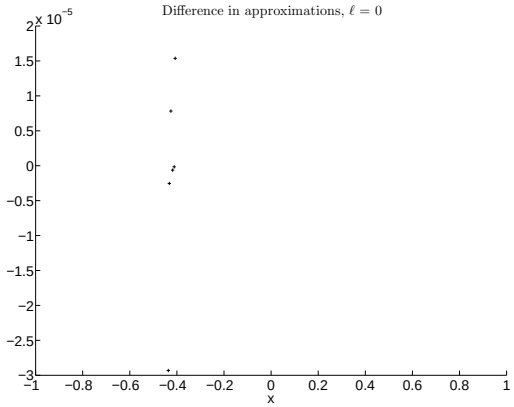
(a) Approximation, moment limiter, $T = 0.5$



(b) $u_j^{(3)}$ adapted only, levels 2,1, and 0 unlimited



(c) $u_j^{(3)}, u_j^{(2)}$ adapted, levels 1 and 0 unlimited



(d) $u_j^{(3)}, u_j^{(2)}$ and $u_j^{(1)}$ adapted, level 0 unlimited

Figure 8: Moment limited approximation at $T = 0.5$, and difference (39), split up into each limiting level, advection equation with initial condition (38), $k = 3$, 64 elements

For the inviscid Burgers' equation, the moment limiter does a really good job. Results at $T = 1$, using initial condition (28), are given in Figure 9.

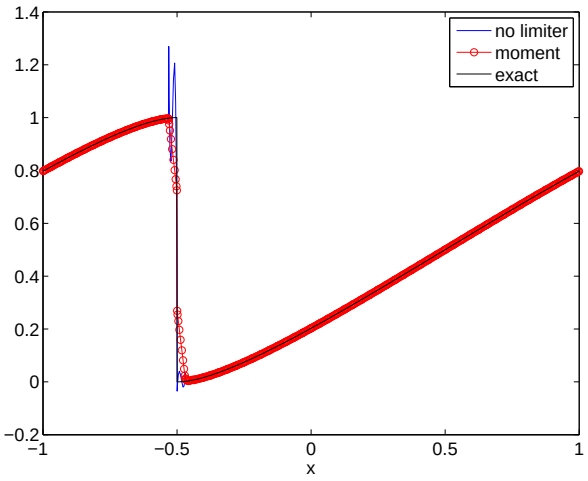


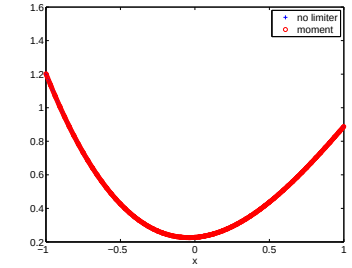
Figure 9: Approximation of the solution at $T = 1$, inviscid Burgers' equation, initial condition (28), $k = 3$, 64 elements

The multiwavelet decomposition of the approximations at $T = 1$ is given in Figure 10. The orders of the multiwavelet contribution for both the unlimited and the moment limited solution are given in Table 3.

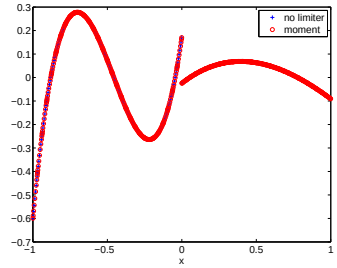
level	0	1	2	3	4	5
no limiter	$\mathcal{O}(10^{-1})$	$\mathcal{O}(10^{-1})$	$\mathcal{O}(10^{-3})$	$\mathcal{O}(10^{-2})$	$\mathcal{O}(10^{-2})$	$\mathcal{O}(10^{-1})$
moment	$\mathcal{O}(10^{-1})$	$\mathcal{O}(10^{-1})$	$\mathcal{O}(10^{-1})$	$\mathcal{O}(10^{-1})$	$\mathcal{O}(10^{-2})$	$\mathcal{O}(10^{-2})$

Table 3: Orders of the multiwavelet contribution for the approximate solution of the inviscid Burgers' equation at $T = 1$, initial condition (28), $n = 6, k = 3$ ($2^n = 64$ elements)

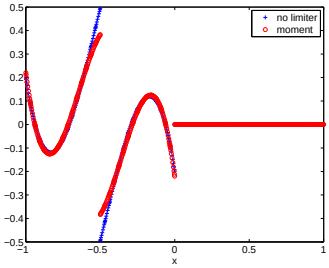
It should be noticed that the moment limited solution is well organized: the contributions decrease if the level m increases, $m = 0, \ldots, n-1$. This means that the approximation is smoothed and seems to be continuous (cf. the results for a continuous function, Figure 61 in Appendix B).
The unlimited solution however, is a mess: level 4 has a contribution of order 10^{-2}, whereas level 5 is $\mathcal{O}(10^{-1})$!
Note that the multiwavelet decomposition can be used to detect the shock region: the multiwavelet contribution of the highest level is much bigger in that region. This gives rise to the idea of using multiwavelets as a shock detector, see Section 3.



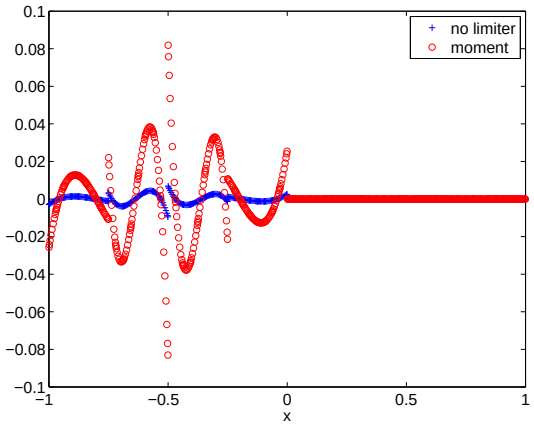
(a) $S^0(x)$



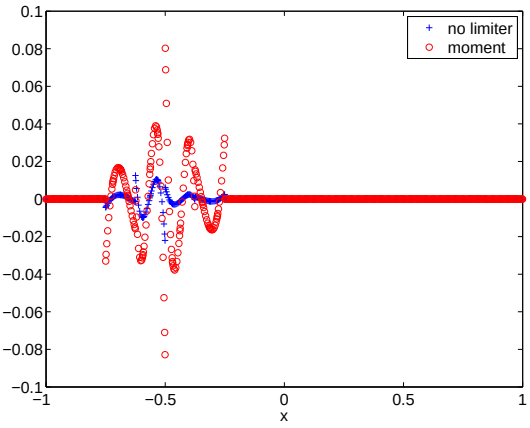
(b) $D^0(x)$



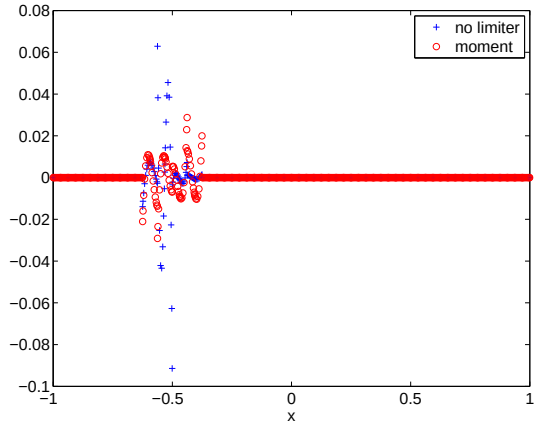
(c) $D^1(x)$



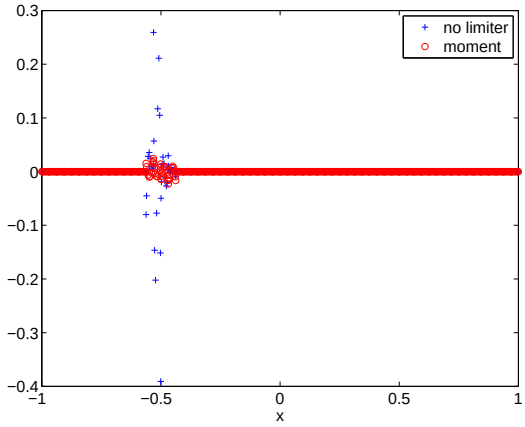
(d) $D^2(x)$



(e) $D^3(x)$



(f) $D^4(x)$



(g) $D^5(x)$

Figure 10: Multiwavelet decomposition of approximations, $T = 1$, inviscid Burgers' equation with initial condition (28), $k = 3, n = 6, 2^n = 64$ elements

Note that the one dimensional Euler equations form a nonlinear system. The DG coefficients for the solution $\mathbf{u}_h(x,t) = \left(u_h^{(1)}, u_h^{(2)}, u_h^{(3)}\right)^\top$, are given by $\mathbf{u}_j^{(\ell)} = \left(u_j^{(1,\ell)}, u_j^{(2,\ell)}, u_j^{(3,\ell)}\right)^\top$, $j = 0, \ldots, N, \ell = 0, \ldots, k$.
The moment limiter is applied to the characteristic variables $\mathbf{w}_j^{(\ell)} = R^{-1}\mathbf{u}_j^{(\ell)}$. The eigenvector matrix R is computed using the Roe averages, as explained in Appendix A.4. Considering element $I_j, j \in \{0, \ldots, N\}$, the moment limiter starts at $\ell = k$ and computes for $b \in \{1, 2, 3\}$:

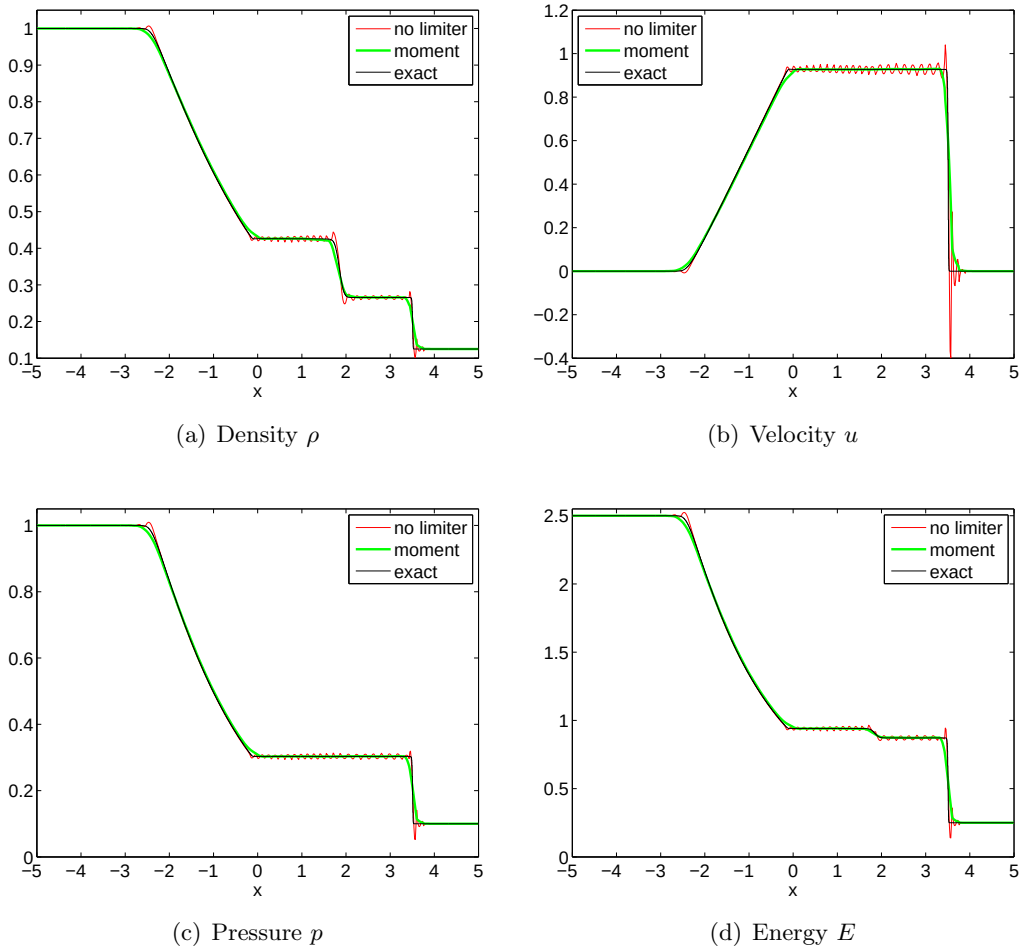
$$\widetilde{w}_j^{(b,\ell)} = \text{minmod}\left(w_j^{(b,\ell)}, \beta_\ell\left(w_{j+1}^{(b,\ell-1)} - w_j^{(b,\ell-1)}\right), \beta_\ell\left(w_j^{(b,\ell-1)} - w_{j-1}^{(b,\ell-1)}\right)\right).$$

If $\widetilde{w}_j^{(b,k)} = w_j^{(b,k)}$, then the limiting procedure is stopped for this element I_j and variable b. Otherwise, the coefficient $w_j^{(b,k-1)}$ is limited, continuing until $w_j^{(b,1)}$ is limited, or stopping the first time that $\widetilde{w}_j^{(b,\ell)} = w_j^{(b,\ell)}$ for some $\ell = k-1, \ldots, 1$. Note that this stopping criterium holds for the b^{th} characteristic variable only.

Finally, the limited values for the conserved variables $\mathbf{u}_j^{(\ell)}, j = 0, \ldots, N, \ell = 0, \ldots, k$, are given by

$$\widetilde{\mathbf{u}}_j^{(\ell)} = R\widetilde{\mathbf{w}}_j^{(\ell)}.$$

In Figure 11, the results are given using the moment limiter for Sod's shock tube, at $T = 2$. The moment limited solution is very good, compared with the unlimited solution. Note that in this case, the exact solution is unknown. Therefore, we use a fine resolution to approximate the exact solution (using 3000 elements).



(a) Density ρ

(b) Velocity u

(c) Pressure p

(d) Energy E

Figure 11: Results of Euler equations with Sod's initial conditions, absorbing boundary conditions, $T = 2, k = 3$, 64 elements. Exact solution is obtained using 3000 elements and $k = 0$.

1.5.6 Maximum principle limiter

The maximum principle limiter is described by Zhang, [25]. This limiter uses an important property of the entropy solution: if for the initial condition $u^0(x), x \in [a,b]$,

$$M = \max_{x\in[a,b]} u^0(x),\ m = \min_{x\in[a,b]} u^0(x),$$

then $u(x,t) \in [m,M]$ for all $x \in [a,b], t \geq 0$ (a strict maximum principle).

Let $p_j(x) = u_h(x,t^s), x \in I_j$ be the polynomial approximation of the solution on element I_j, $j \in \{0,\ldots,N\}$ at time $t^s = s\Delta t, s \in \mathbb{N}$. The maximum principle limiter computes,

$$\widetilde{p}_j(x) = \theta(p_j(x) - \bar{u}_j^s) + \bar{u}_j^s, \tag{41a}$$

where

$$\theta = \min\left\{\left|\frac{M-\bar{u}_j^s}{M_j-\bar{u}_j^s}\right|, \left|\frac{m-\bar{u}_j^s}{m_j-\bar{u}_j^s}\right|, 1\right\}, \tag{41b}$$

$$m = \min_{x\in[a,b]} u^0(x),\ m_j = \min_{x\in I_j} p_j(x), \tag{41c}$$

$$M = \max_{x\in[a,b]} u^0(x),\ M_j = \max_{x\in I_j} p_j(x), \tag{41d}$$

and for the average at time t^s, equation (37) is used:

$$\bar{u}_j^s = \frac{1}{\sqrt{2}} u_j^{(0)}(t^s). \tag{41e}$$

Zhang avoids to evaluate the extrema of a polynomial, and therefore decides to control the values of $\widetilde{p}_j$ at quadrature points only. This means that m_j and M_j are replaced by,

$$m_j = \min_{x\in S_j} p_j(x),\ M_j = \max_{x\in S_j} p_j(x). \tag{42}$$

such that the limiter is sufficient to enforce $\widetilde{p}_j(x) \in [m,M], \forall x \in S_j$. Here, S_j is the set of Legendre Gauss-Lobatto quadrature points for I_j, which are used to compute,

$$\bar{u}_j^s = \frac{1}{\Delta x}\int_{I_j} p_j(x)dx.$$

Note that we use equation (41e) to compute the average of the solution. Therefore, we do not need a quadrature, and we can use any set of points in I_j.
If $p_j(x)$ is replaced by the DG approximation $\sum_{\ell=0}^{k} u_j^{(\ell)}\phi_\ell(\xi)$, where $\xi = \frac{2}{\Delta x}(x - x_j)$, then the maximum principle limiter computes the limited DG coefficients $\widetilde{u}_j^{(\ell)}$. In this

case, equation (41a) can be written as (using equation (41e) and the definition of ϕ_0),

$$\begin{aligned}\sum_{\ell=0}^{k} \widetilde{u}_j^{(\ell)} \phi_\ell(\xi) &= \widetilde{p}_j(x) \\ &= \theta(p_j(x) - \bar{u}_j^s) + \bar{u}_j^s \\ &= \theta\left(\sum_{\ell=0}^{k} u_j^{(\ell)}\phi_\ell(\xi) - \frac{1}{\sqrt{2}}u_j^{(0)}\right) + \frac{1}{\sqrt{2}}u_j^{(0)} \\ &= \theta\left(\sum_{\ell=0}^{k} u_j^{(\ell)}\phi_\ell(\xi) - u_j^{(0)}\phi_0(\xi)\right) + u_j^{(0)}\phi_0(\xi) \\ &= u_j^{(0)}\phi_0(\xi) + \sum_{\ell=1}^{k}(\theta u_j^{(\ell)})\phi_\ell(\xi), \end{aligned} \tag{43}$$

which means that the limited DG coefficients are given by,

$$\widetilde{u}_j^{(0)} = u_j^{(0)}, \text{ and } \widetilde{u}_j^{(\ell)} = \theta u_j^{(\ell)}, \ell = 1, \ldots, k.$$

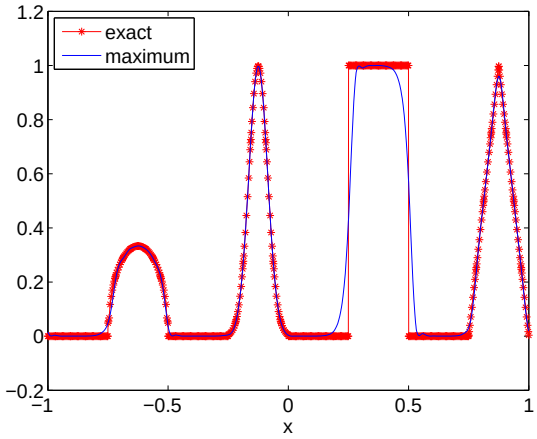
Note that this limiter indeed preserves the average $\bar{u}_j^s$, because the coefficients $u_j^{(0)}$ are unchanged, $j = 0, \ldots, N$.
For the implementation, the following algorithm is used:

Algorithm 2 Maximum principle limiter

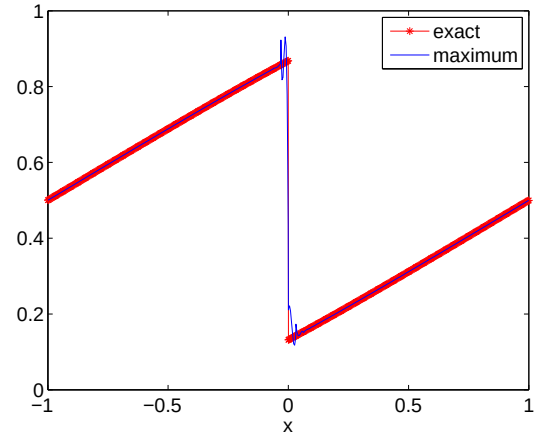
For the initial condition, compute m and M;
for each RK step **do**
  for each element $I_j, j = 0, \ldots, N$ **do**
    Compute $u_h(x,t)$ in local points $\xi = -1, -\frac{3}{4}, -\frac{1}{2}, -\frac{1}{4}, 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$, corresponding to global coordinates $x_j + \frac{\Delta x}{2}\xi$;
    Compute m_j and M_j, using equation (42);
    Compute θ;
    Compute the limited DG coefficients $\widetilde{u}_j^{(0)}, \ldots, \widetilde{u}_j^{(k)}$.
  end for
end for

Results for the linear advection equation with initial condition (38) at $T = 0.5$ can be seen in Figure 12a. Looking back at Figure 8a for the moment limited solution, it should be noted that the behaviour of the maximum principle limiter is much better in case of the triangle, and is similar for the other parts of the initial condition.

For the inviscid Burgers' equation however, the maximum principle limiter does not work. Oscillations arise, which can be seen in Figure 12b for $T = 2$. Following the report of Zhang, [25], the maximum principle limiter and another limiter should be applied simultaneously. The good results using the maximum principle limiter for the linear advection equation where only found by accident, and are due to the special structure of this equation.
The moment limiter for Burgers' equation is very good, and there is no need to apply the maximum principle limiter after having used the moment limiter: the approximation remains equal to the moment limited solution. However, it may be possible that the combination of these limiters give better results for other problems (such as a dam break problem, for which the maximum principle limiter is constructed).



(a) Advection equation, initial condition (38), $T = 0.5$



(b) Burgers' equation, initial condition (28), $T = 2$

Figure 12: Approximations using the maximum principle limiter, $k = 3$, 64 elements

Besides the DG coefficients, it is also possible to look at the impact of the maximum principle limiter on the multiwavelet decomposition, since, for $j = 0, \ldots, N, \ell = 0, \ldots, k$, $2^{-\frac{n}{2}} u_j^{(\ell)} = s_{\ell j}^n$. This is done in Appendix C.

1.5.7 Compare moment limiter to maximum principle limiter

Both the moment limiter (Section 1.5.5) and the maximum principle limiter (Section 1.5.6) give nice results for the linear advection equation. To compare the moment limiter to the maximum principle limiter, the following simple initial condition is used:

$$u^0(x) = \begin{cases} 1, & x \in \left[-\frac{1}{2}, \frac{1}{2}\right], \\ 0, & \text{else.} \end{cases} \tag{44}$$

At $T = 3$, the result using either the moment limiter or the maximum principle limiter is given in Figure 13. Note that the maximum principle limiter gives better results.

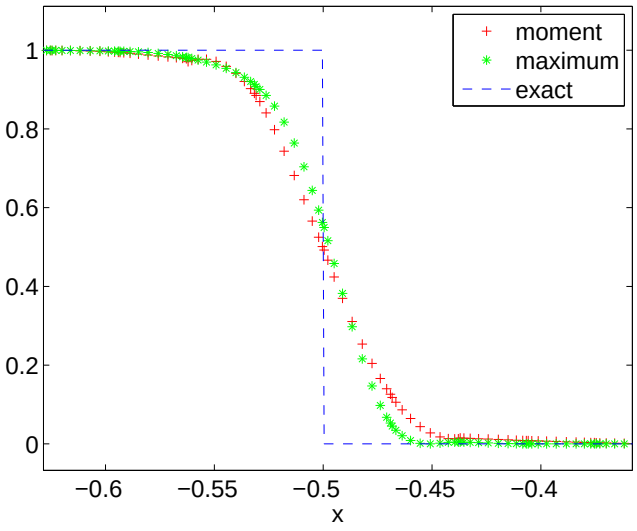


Figure 13: Approximation of the solution, linear advection equation with square wave, $T = 3, k = 3, n = 6$ ($2^n = 64$ elements)

2 Search for a new multiwavelet limiter

If no limiter is used, oscillations arise in the neighbourhood of the discontinuities, and these oscillations grow in time, as we have seen in Section 1.5. Therefore, a limiter is used to avoid obtaining these spurious oscillations. In Section 1.5, various limiters for DG were given. Most of them have problems with high order accuracy ($k \geq 1$), smooth extrema, multi-dimensions and complex geometries.
In this section, I search for a new limiter that uses multiwavelets. Because limiters are applied to discontinuous parts of a function, it is necessary to look at the multiwavelet decomposition of a discontinuous function, which is done in Section 2.1. Then we will attempt to use this information to find a more effective limiter by adapting the multiwavelet contribution (Section 2.2) or adapting the multiwavelet coefficients of the approximation (Section 2.3).

2.1 Multiwavelet decomposition of discontinuous function

As an example, the square wave (44) is decomposed, and given in Figure 14. Here, $n = 6$ is used, such that there are $2^n = 64$ elements. For levels $m = 2, \ldots, n-1$, the multiwavelet contribution is zero. Therefore, in this special case, it holds that

$$u_h(x,0) = P_6^k u_h(x,0) = S^0(x) + \sum_{m=0}^{5} D^m(x) = S^0(x) + \sum_{m=0}^{1} D^m(x).$$

For the linear advection equation ($u_t + u_x = 0$), together with periodic boundary conditions, the exact solution is given by, $u(x,t) = u^0(x-t)$. In general, the discontinuity is not located at the boundary of an element. At $T = 1.5625 \cdot 10^{-2}$, for example, the right shock of the exact solution is located at the center of element $[0.5; 0.53125]$ (using $n = 6, 2^n = 64$ elements, $\Delta x = 0.03125$), which is visualized in Figure 15a.
In this figure, the projection of the exact solution onto Φ^k (the DG approximation space), is visualized as well, which is not accurate in the shock element $[0.5; 0.53125]$.
The oscillations of the exact projection in the shock element are caused by the fact that the functions ϕ_ℓ are continuous (see definition (3) in Section 1.1), whereas the exact solution is discontinuous in this element. It is not possible to approximate a discontinuous function by a polynomial of degree $k < \infty$. The visible projection is the best at approximating a discontinuous function by a continuous polynomial. This appearance is described for Fourier expansions (Gibbs' phenomenon).

Note that the multiwavelet decomposition is deduced from the discontinuous Galerkin approximation (see Section 1.3, equation (26)). Therefore, the multiwavelet decomposition of the exact solution in a shock element is not accurate. Note that, initially, the idea was to use a multiwavelet decomposition of the function within each element (see Appendix D). Unfortunately, this approach did not work, for reasons described in the appendix.

For $T = 0.03125$, the shock of the exact solution is on the boundary of the next element, and the projection of the exact solution onto Φ^k is given in Figure 15b. Note that in this case, this projection can be used to compute the multiwavelet decomposition (no Gibbs' phenomenon). In Figure 16 the multiwavelet decomposition can be seen, together with the decomposition of the numerical approximation with(out) the moment limiter. Here, the boundaries of the intervals I_j^m are visible, for $m = 0, \ldots, 5$, $j = 0, \ldots, 2^m - 1$.

Note that the exact solution and the approximation without a limiter have approximately the same scaling function contribution (Figure 16a), and multiwavelet contributions of levels 0 to 3 (Figures 16b to 16e). Differences are visible in levels 4 and 5 (Figures 16f and 16g). In Table 4, the orders of the errors of the multiwavelet decomposition are stated, given by,

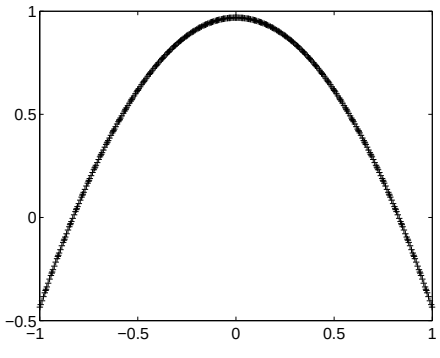
$$\left|D^m(x,T) - D_h^m(x,t)\right|, m = 0, \ldots, n-1,$$

where $D^m(x,T)$ belongs to the exact solution at time T, and $D_h^m(x,T)$ is the multiwavelet contribution on level m of the approximate solution. Note that the moment limited solution is very regular: the order of the error is similar for each multiwavelet level. The unlimited solution however, has a growing contribution to the error for higher levels.
This behaviour (large errors for higher levels of multiwavelet decomposition) is found for all times. Therefore, I expect that the higher levels 4 and 5, are the most important causes for oscillations.

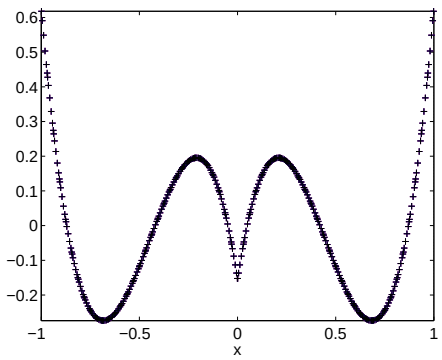
In the next sections, the multiwavelet decomposition of the discontinuous Galerkin approximation is used to limit the solution.

level	0	1	2	3	4	5
no limiter	$\mathcal{O}(10^{-13})$	$\mathcal{O}(10^{-12})$	$\mathcal{O}(10^{-8})$	$\mathcal{O}(10^{-5})$	$\mathcal{O}(10^{-3})$	$\mathcal{O}(10^{-3})$
moment	$\mathcal{O}(10^{-3})$	$\mathcal{O}(10^{-3})$	$\mathcal{O}(10^{-3})$	$\mathcal{O}(10^{-3})$	$\mathcal{O}(10^{-3})$	$\mathcal{O}(10^{-3})$

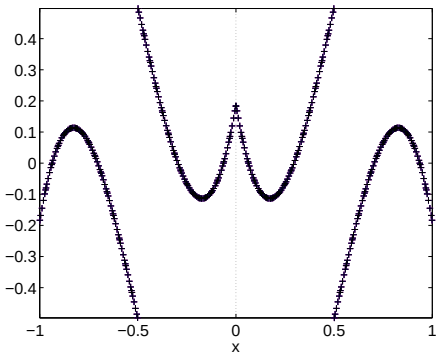
Table 4: Error $|D^m(x,T) - D_h^m(x,T)|$ between exact solution and approximation, linear advection equation using square wave, $T = 0.03125$, $n = 6, k = 3$ ($2^n = 64$ elements)



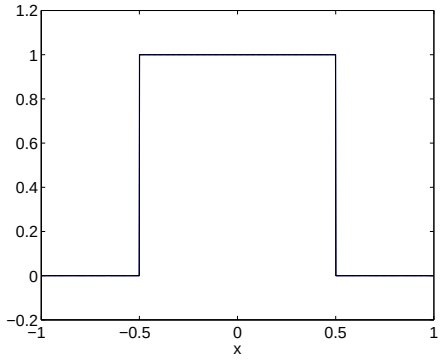
(a) $S^0(x)$



(b) $D^0(x)$

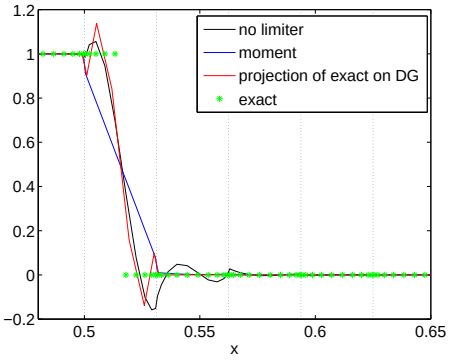


(c) $D^1(x)$

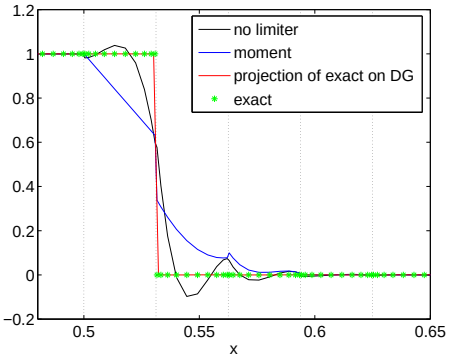


(d) Initial condition

Figure 14: Multiwavelet decomposition of initial condition (44), $k = 3, n = 6$ ($2^n = 64$ elements)

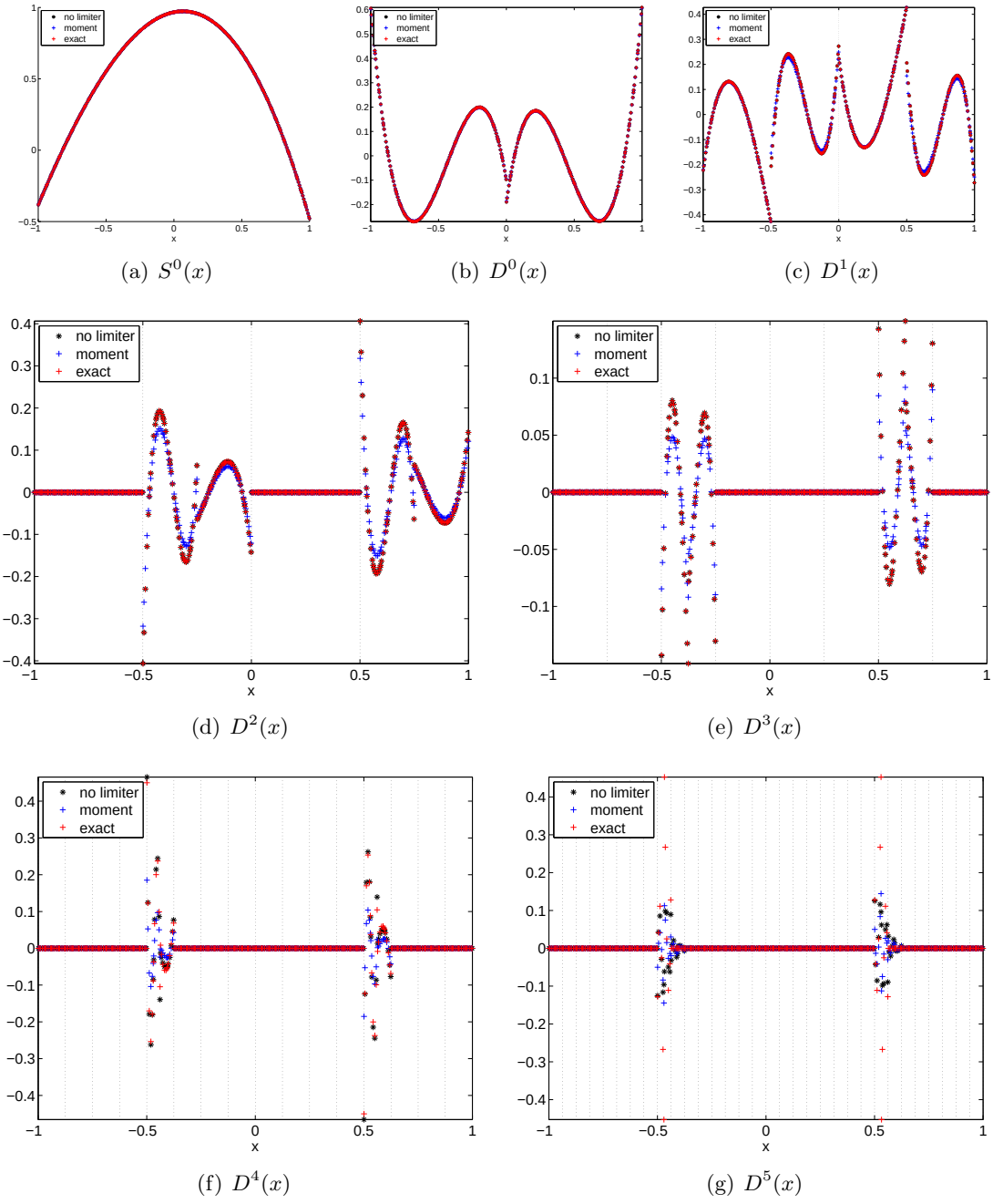


(a) $T = 1.5625 \cdot 10^{-2}$, shock is inside element



(b) $T = 0.03125$, shock is on boundary

Figure 15: Approximations of the solution of the linear advection equation, square wave, $k = 3$, 64 elements



(a) $S^0(x)$ (b) $D^0(x)$ (c) $D^1(x)$

(d) $D^2(x)$ (e) $D^3(x)$

(f) $D^4(x)$ (g) $D^5(x)$

Figure 16: Multiwavelet decomposition of the solution of the linear advection equation, square wave, $T = 0.03125, k = 3, n = 6$ ($2^n = 64$ elements)

2.2 Adapting the multiwavelet contribution

The first way to limit the discontinuous Galerkin approximations has to do with the multiwavelet contribution. In this section, the multiwavelet contribution on level $n-1$, given by $D^{n-1}(x)$, is adapted, such that the approximation,

$$u_h(x,t) = S^0(x) + \sum_{m=0}^{n-1} D^m(x), \tag{45a}$$

is replaced by its limited version,

$$\widetilde{u}_h(x,t) = S^0(x) + \sum_{m=0}^{n-2} D^m(x) + \widetilde{D}^{n-1}(x). \tag{45b}$$

Only level $n-1$ is adapted, because we remember from Section 2.1 that the error belongs mainly to the highest levels of the multiwavelet contribution. Maybe the approximation in the shock elements improves if less details are used: $\widetilde{D}^{n-1}(x) = 0$ in the discontinuous elements, whereas $\widetilde{D}^{n-1}(x) = D^{n-1}(x)$ in continuous regions.
Examples of these limiters are given in Sections 2.2.1 and 2.2.2.

2.2.1 Combine moment limiter and wavelet decomposition

The moment limiter works well and is reliable, as can be seen in Section 1.5.5. Therefore, a logical step is to find a way to improve upon it using multiwavelet decomposition. In this section, a combination of the moment limiter and the multiwavelet decomposition is done. This combined limiter works on elements $I_j, j \in \{0,\ldots,N\}$, where the moment limiter adapts all coefficients $u_j^{(\ell)}, \ell = k,\ldots,1$. In these elements, the multiwavelet contribution $\widetilde{D}^{n-1}(x)$ is set equal to zero. The following algorithm is used:

Algorithm 3 Combination of moment limiter and multiwavelet decomposition

for each RK step **do**
- Compute the moment limited solution (first limiter);
- Compute multiwavelet decomposition of moment limited solution, given by (45a);
- Define $\widetilde{D}^{n-1}(x) = D^{n-1}(x), \forall x \in [-1,1]$;
- Find the elements I_j where $u_j^{(k)}, u_j^{(k-1)}, \ldots, u_j^{(1)}$ are adapted by the moment limiter, $j \in \{0,\ldots,N\}$;
- For $x \in I_j$, set $\widetilde{D}^{n-1}(x) = 0$ (second limiter);
- Compute limited approximation, given by (45b);
- Compute limited coefficients $\widetilde{u}_j^{(\ell)}, j = 0,\ldots,N, \ell = 0,\ldots,k$.

end for

Results at $T = 5$ for the linear advection equation with initial condition (44) are given in Figure 17. Note that the new limiter is worse than the moment limiter on the left side of the shock, and the results are not very accurate in this region. In the next section, I try to use a multiwavelet limiter without the help of the moment limiter.

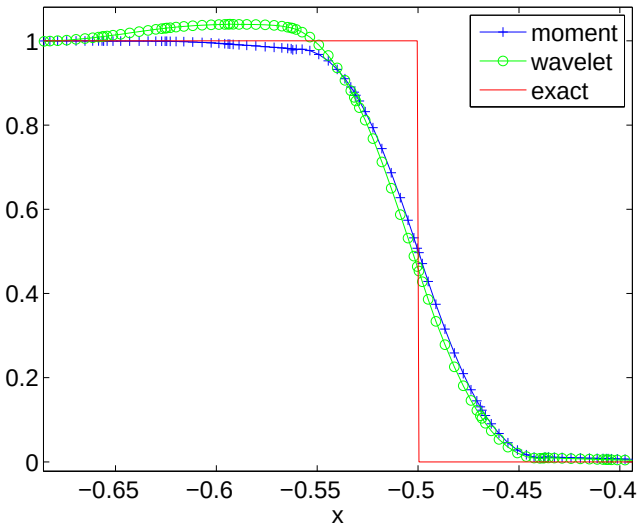


Figure 17: Approximation of the solution of the linear advection equation, square wave, $T = 5$. Wavelet limiter computes moment limited solution and sets multiwavelet contribution equal to zero in elements I_j where $u_j^{(k)}, \ldots, u_j^{(1)}$ are adapted, $j = 0, \ldots, N, k = 3$, 64 elements

2.2.2 Shock detection using exact solution

In the previous section, the moment limiter was combined with the multiwavelet decomposition. Note that the moment limiter acted as a shock detector, because I applied the second limiter only in these elements I_j where $u_j^{(k)}, \ldots, u_j^{(1)}$ were adapted by the moment limiter ($j = 0, \ldots, N$).
In order to test that the multiwavelet limiter is viable, we use the exact solution (which is known for the linear advection equation with periodic boundary conditions) to detect the location of the shock. Similar to the approach in Section 2.2.1, we set $\widetilde{D}^{n-1}(x)$ to zero in the neighbourhood of the discontinuity.
For the linear advection equation with the square wave as an initial condition, six points are used to compute the multiwavelet contribution in element $[0.5; 0.53125]$ (using $\Delta x = 0.03125$), namely,

$$\widetilde{x}_i \in \{0.5011, 0.5053, 0.5119, 0.5194, 0.5260, 0.5302\}, i = 0, \ldots, 5.$$

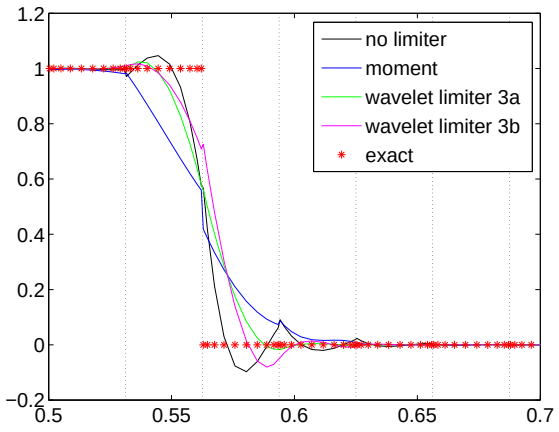
The following algorithm is used to limit the solution:

Algorithm 4 Limiter sets multiwavelet contribution equal to zero around shock

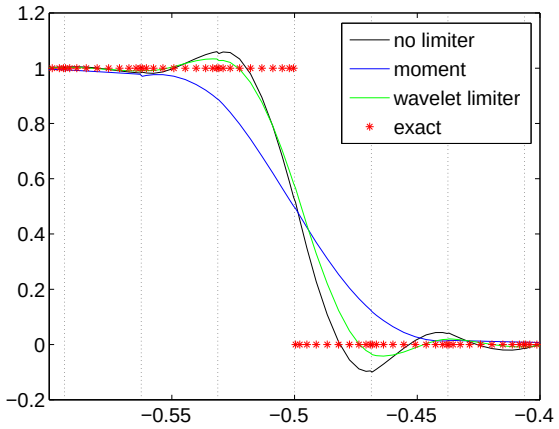
for each RK step **do**
  Compute multiwavelet decomposition of approximation, given by (45a);
  Define $\widetilde{D}^{n-1}(x) = D^{n-1}(x), \forall x \in [-1, 1]$;
  if the shock is in interval $\left[0.5; \frac{\widetilde{x}_0+\widetilde{x}_1}{2}\right]$ **then**
    $\widetilde{D}^{n-1}(x) = 0$ for $x = \widetilde{x}_0, \ldots, \widetilde{x}_5$;
  else if the shock is in interval $\left[\frac{\widetilde{x}_0+\widetilde{x}_1}{2}; \frac{\widetilde{x}_1+\widetilde{x}_2}{2}\right]$ **then**
    $\widetilde{D}^{n-1}(x) = 0$ for $x = \widetilde{x}_1, \ldots, \widetilde{x}_5$;
  else if the shock is in interval $\left[\frac{\widetilde{x}_1+\widetilde{x}_2}{2}; \frac{\widetilde{x}_2+\widetilde{x}_3}{2}\right]$ **then**
    $\widetilde{D}^{n-1}(x) = 0$ for $x = \widetilde{x}_2, \ldots, \widetilde{x}_5$;
  else if the shock is in interval $\left[\frac{\widetilde{x}_2+\widetilde{x}_3}{2}; \frac{\widetilde{x}_3+\widetilde{x}_4}{2}\right]$ **then**
    $\widetilde{D}^{n-1}(x) = 0$ for $x = \widetilde{x}_3, \ldots, \widetilde{x}_5$;
  else if the shock is in interval $\left[\frac{\widetilde{x}_3+\widetilde{x}_4}{2}; \frac{\widetilde{x}_4+\widetilde{x}_5}{2}\right]$ **then**
    $\widetilde{D}^{n-1}(x) = 0$ for $x = \widetilde{x}_4, \widetilde{x}_5$;
  else if the shock is in interval $\left[\frac{\widetilde{x}_4+\widetilde{x}_5}{2}; 0.53125\right]$ **then**
    $\widetilde{D}^{n-1}(x) = 0$ for $x = \widetilde{x}_5$;
  end if
  Compute limited approximation, given by (45b);
  Compute limited DG coefficients $\widetilde{u}_j^{(\ell)}, j = 0, \ldots, N, \ell = 0, \ldots, k$.
end for

In general, a limiter smoothes the solution. This means that we expect that limiting is only necessary for small times $t < T_s$. For $t > T_s$, we assume that the approximate solution is smooth enough, such that no limiter is used anymore.

In Figure 18a two different multiwavelet limiters are used: limiter 3a limits the solution according to algorithm 4 (for $t < 0.03125$), and limiter 3b extends algorithm 4 to element $[0.53125; 0.5625]$ as well (for $t < 0.0625$). Limiter 3a is better: the oscillations are small. It should be noted that these oscillations grow slowly in time, although they remain smaller than the oscillations of the unlimited solution for all times (Figure 18b).



(a) $T = 0.0625$, limiter 3a limits for $t < 0.03125$, limiter 3b for $t < 0.0625$

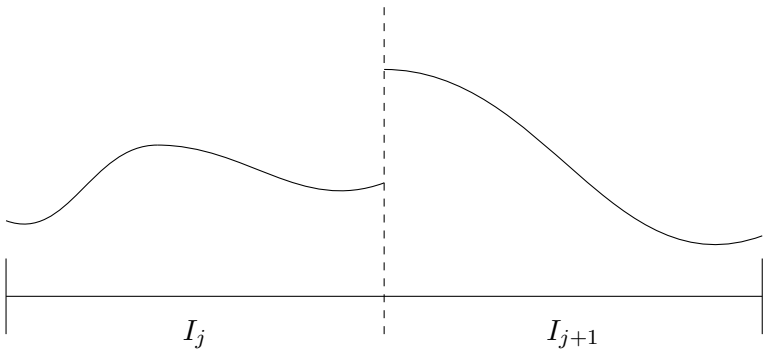


(b) $T = 3$, multiwavelet limiter is used for $t < 0.03125$

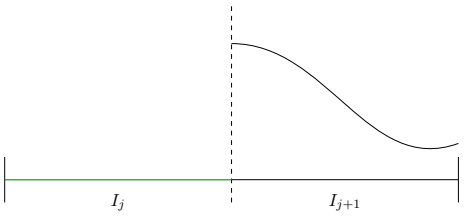
Figure 18: Approximate solution of the linear advection equation, square wave. Wavelet limiter sets multiwavelet contribution in level $n - 1$ equal to zero around shock, $k = 3$, 64 elements

2.2.3 Summary and problems of adapting multiwavelet contribution

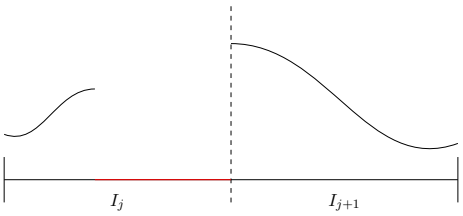
In Sections 2.2.1 and 2.2.2, the contribution of $D^{n-1}(x)$ was adapted to limit the approximation in a shock element, instead of limiting the multiwavelet coefficients $d_{\ell j}^{n-1}$ themselves, $j = 0, \ldots, 2^{n-1} - 1, \ell = 0, \ldots, k$.
As an example, we assume that the shock is found in element $I_j, j \in \{0, \ldots, 2^n - 1\}$. Figure 19 gives an initial multiwavelet contribution $D^{n-1}(x)$ and the limited function $\widetilde{D}^{n-1}(x)$ on element $I_{\frac{j}{2}}^{n-1}$ (the element on level $n-1$ that contains I_j, Section 1.3).



(a) $D^{n-1}(x)$ on element $I_{\frac{j}{2}}^{n-1}$ (level $n-1$)



(b) $\widetilde{D}^{n-1}(x)$ using approach of Section 2.2.1



(c) $\widetilde{D}^{n-1}(x)$ using approach of Section 2.2.2

Figure 19: Limiter adapts $D^{n-1}(x)$ in element I_j into $\widetilde{D}^{n-1}(x)$

If we use the moment limiter to act as a shock detector, we set $\widetilde{D}^{n-1}$ to zero in shock element I_j (Figure 19b). After this we compute $\widetilde{u}_h(x,t)$ using equation (45b). The limited DG coefficients, $\widetilde{u}_j^{(\ell)}, j = 0, \ldots, 2^n - 1, \ell = 0, \ldots, k$, are then computed by projecting $\widetilde{u}_h(x,t)$ onto Φ^k. This projection costs a lot of computation time, which is a disadvantage of this method.
The approach from Section 2.2.2 is visualized in Figure 19c. Here, the discontinuity of $\widetilde{D}^{n-1}(x)$ is not located on one of the boundaries $x_{j\pm\frac{1}{2}}$. This means that the limited approximation $\widetilde{u}_h(x,t)$ has a discontinuity inside element I_j! Instead of solving the problem of a discontinuity in the approximation, new discontinuities are created.
Limiting the multiwavelet contribution causes problems with new discontinuities and computation time. Therefore, it is more convenient to adapt the multiwavelet coefficients $d_{\ell j}^m$ themselves. Examples of this approach are given Section 2.3.

2.3 Adapting the multiwavelet coefficients

In the previous section, we used the multiwavelet contribution $D^{n-1}(x)$ for limiting, and discovered some problems with this approach (Section 2.2.3). In the next sections, the multiwavelet coefficients $d_{\ell j}^m, m = 0, \ldots, n-1, j = 0, \ldots, 2^m - 1, \ell = 0, \ldots, k$, are used for limiting. An advantage of these limiters is that we can easily reconstruct $\widetilde{s}_{\ell j}^n, \ell = 0, \ldots, k, j = 0, \ldots, 2^n - 1$, from $\widetilde{d}_{\ell j}^m, j = 0, \ldots, 2^m - 1$ (reconstruction steps are described in Appendix B). The limited DG coefficients $\widetilde{u}_j^{(\ell)}, j = 0, \ldots, N, \ell = 0, \ldots, k$, are then computed using equation (26) from Section 1.3.

2.3.1 Using the averages of the multiwavelet contribution on each element

In this section, the ideas from Section 2.1 about the multiwavelet decomposition of a discontinuous function are used.
If we look back at Figure 16, we see the following appearance: if the multiwavelet contribution is equal to zero for a certain element I_j^m on level m, so if

$$\sum_{\ell=0}^{k} d_{\ell j}^m \psi_{\ell j}^m(x) = 0, \ x \in I_j^m, m \in \{0, \ldots, n-1\}, j \in \{0, \ldots, 2^m - 1\}, \tag{46}$$

then this contribution is zero for higher levels in that region as well.
Equation (46) should hold for all k, and is therefore satisfied if $d_{\ell j}^m = 0, \ \forall \ell = 0, \ldots, k$.
Using Figure 3, we propose:

$$\begin{aligned} &\text{If for some } m \in \{0, \ldots, n-1\}, j \in \{0, \ldots, 2^m - 1\}, d_{\ell j}^m = 0, \forall \ell = 0, \ldots, k, \\ &\text{then } d_{\ell, 2^{\mu-m} j}^{\mu} = \ldots = d_{\ell, 2^{\mu-m}(j+1)-1}^{\mu} = 0, \forall \ell = 0, \ldots, k, \mu = m+1, \ldots, n-1. \end{aligned} \tag{47}$$

This statement seems to hold for every time and in each level m, but it is not true in general. A counter example is easily found for $k = 0$, and is given in Appendix E.

However, it is possible to use statement (47) to define a new multiwavelet limiter, because it holds for the initial conditions that are used in my master thesis. Therefore, we use a threshold $\epsilon = 10^{-10}$, such that $\widetilde{d}_{\ell j}^m \equiv 0$ if $\left| d_{\ell j}^m \right| < \epsilon, \ell = 0, \ldots, k, m = 0, \ldots, n-1$, $j = 0, \ldots, 2^m - 1$.

Besides statement (47), it is interesting to look at the averages of the multiwavelet contribution in each element. In Figure 20 a zoom in is given of Figure 16g from Section 2.1. Note that the multiwavelet contribution of the exact solution is zero in element $[0.5625; 0.625]$, whereas the unlimited solution is nonzero, but has a small average with respect to the contribution in the neighbouring element $[0.5; 0.5625]$.
This observation introduces the idea to define the averages $\bar{D}_j^m$ of $|D_j^m(x)|$ on element $I_j^m, m = 0, \ldots, n-1, j = 0, \ldots, 2^m - 1$. If $\bar{D}_j^m$ is too small with respect to the average in neighbouring element I_{j-1}^m, given by $\bar{D}_{j-1}^m$, i.e. if

$$\frac{\bar{D}_{j-1}^m}{\bar{D}_j^m} > M,$$

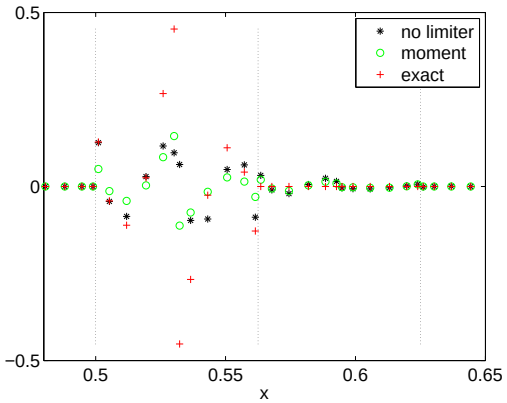


Figure 20: Zoom of Figure 16g, $D^5(x)$ at $T = 0.03125, k = 3, n = 6$ ($2^n = 64$ elements)

the limiter sets $\widetilde{d}^m_{\ell j} \equiv 0, \ell = 0, \ldots, k$. The same happens with respect to the right neighbouring element I^m_{j+1}, for each level m and each element I^m_j (we choose $M = 100$).

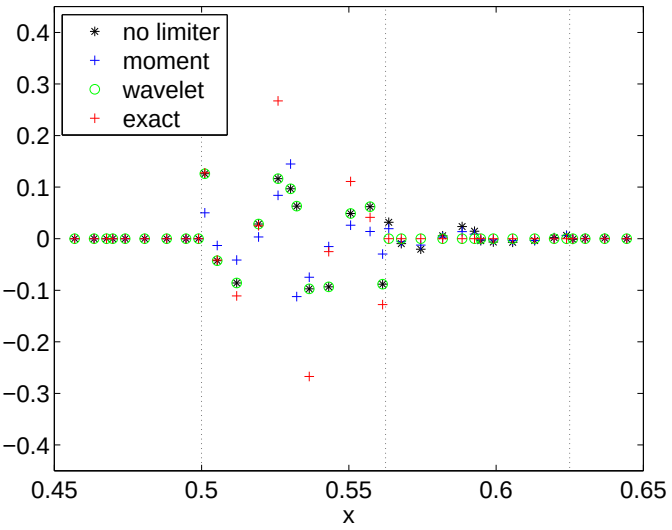
The following limiting algorithm is used:

Algorithm 5 Multiwavelet limiter, based on averages

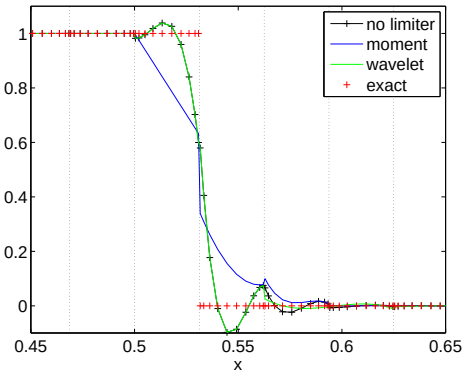
for each RK step **do**
  Compute multiwavelet decomposition;
  Set $\widetilde{d}^m_{\ell j} = d^m_{\ell j}, \ell = 0, \ldots, k, m = 0, \ldots, n-1, j = 0, \ldots, 2^m - 1$;
  if $\left|d^m_{\ell j}\right| < \epsilon = 10^{-10}, \ell = 0, \ldots, k, m = 0, \ldots, n-1, j = 0, \ldots, 2^m - 1$ **then**
    $\widetilde{d}^m_{\ell j} \equiv 0$;
  end if
  if for some element $I^m_j : \widetilde{d}^m_{\ell j} = 0, \forall \ell = 0, \ldots, k$ **then**
    set $\widetilde{d}^\mu_{\ell, 2^{\mu-m} j} = \ldots = \widetilde{d}^\mu_{\ell, 2^{\mu-m}(j+1)-1} = 0, \forall \ell = 0, \ldots, k, \mu = m+1, \ldots, n-1$;
  end if
  Start in level 1: compute $\bar{D}^1_0$ and $\bar{D}^1_1$, using coefficients $\widetilde{d}^m_{\ell j}$;
  if $\bar{D}^1_a$ is small with respect to $\bar{D}^1_b$ (more than 100 times as small), $a, b \in \{0, 1\}$ **then**
    Information in element I^1_a can be neglected: set $\widetilde{d}^1_{\ell a} \equiv 0, \forall \ell = 0, \ldots, k$;
    Set $\widetilde{d}^\mu_{\ell, 2^{\mu-1} a} = \ldots = \widetilde{d}^\mu_{\ell, 2^{\mu-1}(a+1)-1} = 0, \forall \ell = 0, \ldots, k, \mu = 2, \ldots, n-1$;
  end if
  Do the same limiting trick for levels 2 to $n-1$;
  Reconstruct $\widetilde{s}^n_{\ell j}$, and compute $\widetilde{u}^{(\ell)}_j, \ell = 0, \ldots, k, j = 0, \ldots, 2^n - 1$;
end for

Using this algorithm to limit the solution gives results such as those shown in Figure 21 (linear advection equation, square wave as an initial condition), both for $T = 0.0325$ and for $T = 5$. In Figure 21a the multiwavelet contribution in level 5 can be seen for $T = 0.0325$. Indeed, the contribution is now zero in element $[0.5625; 0.625]$, and equals

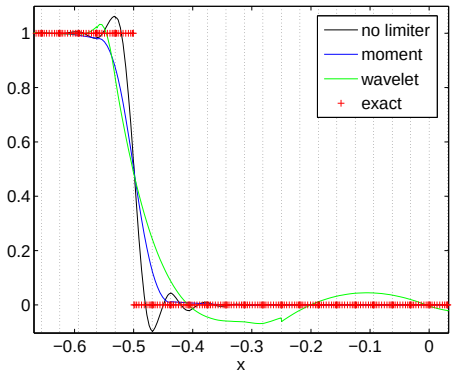
the multiwavelet contribution of the unlimited solution in level $[0.5; 0.5625]$.
Note that for $T = 5$, the oscillations are smeared, see Figure 21c.
Although this limiter removes small information in the multiwavelet decomposition, in the long run the results become bad. Therefore, this limiter is not useful.



(a) $D^5(x)$ at $T = 0.03125$



(b) Approximation at $T = 0.03125$

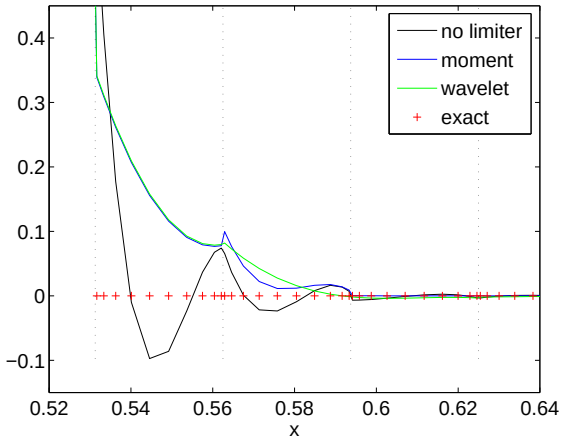


(c) Approximation at $T = 5$

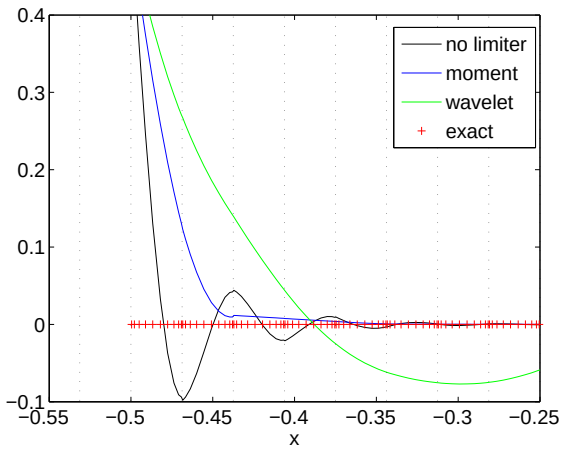
Figure 21: Results using averaging multiwavelet limiter of Section 2.3.1, linear advection equation, square wave, $k = 3, n = 6$ ($2^n = 64$ elements)

2.3.2 Combining moment limiter and averaging idea

In Section 2.3.1, the averages of the multiwavelet contribution where used to limit the solution. When inspecting Figure 20 for the moment limiter results, it should be noticed that the moment limiter has nonzero contribution in element $[0.5625; 0.625]$, whereas this should be zero. Therefore, it may be handy to combine the moment limiter with the idea of Section 2.3.1. At $T = 0.03125$, the result can be seen in Figure 22a. This looks promising, but the limiter is destructive for longer times, see Figure 22b.



(a) $T = 0.03125$



(b) $T = 5$

Figure 22: Combine moment limiter and averaging multiwavelet limiter of Section 2.3.1, linear advection equation, square wave, $k = 3$, 64 elements

2.3.3 Using the maximum principle idea on the multiwavelet coefficients

Another option is to look at the maximum principle limited solution (Zhang, [25], and Section 1.5.6), and to modify this limiter to work on the multiwavelet coefficients. For this, a lot of notation is needed. Because I only use this limiter for the linear advection equation on $[-1,1]$, I do not use a general interval $[a,b]$ in the computations below.
The multiwavelet decomposition of the initial condition, is given by (using equations (21)),

$$u_h(x,0) = S^0(x,0) + \sum_{m=0}^{n-1} D^m(x,0). \tag{48}$$

Next, I define the following global minima and maxima of the multiwavelet decomposition for the initial condition (cf. equations (41c) and (41d)):

$$m^0_{S^0} = \min_{x\in[-1,1]} S^0(x,0), \quad m^0_{D^m} = \min_{x\in[-1,1]} D^m(x,0); \tag{49a}$$

$$M^0_{S^0} = \max_{x\in[-1,1]} S^0(x,0), \quad M^0_{D^m} = \max_{x\in[-1,1]} D^m(x,0), m = 0,\ldots,n-1. \tag{49b}$$

For each RK step, we write $u_h(x,t^s) = S^0(x) + \sum_{m=0}^{n-1}\sum_{j=0}^{2^m-1} D^m_j(x)$ (for readability, I omit t^s from now on).
The goal is to limit each of the polynomials $S^0(x), D^m_j(x)$, for $m = 0,\ldots,n-1$, $j = 0,\ldots,2^m-1$ separatedly (similar to the general maximum principle limiter, which limits $p_j(x)$ for each element I_j, $j = 0,\ldots,N$). Therefore, the local minima and maxima for these functions should be defined, and the average of the function over an interval.
For the minima and maxima, I transform equations (41c) and (41d) into,

$$m_{S^0} = \min_{x\in[-1,1]} S^0(x), \quad m_{D^m_j} = \min_{x\in I^m_j} D^m_j(x); \tag{50a}$$

$$M_{S^0} = \max_{x\in[-1,1]} S^0(x), \quad M_{D^m_j} = \max_{x\in I^m_j} D^m_j(x), m = 0,\ldots,n-1, j = 0,\ldots,2^m-1, \tag{50b}$$

where,

$$I^m_j = (-1+2^{-m+1}j, -1+2^{-m+1}(j+1)], \tag{51}$$

is the interval where $\psi^m_{\ell j}$ is nonzero, $\ell = 0,\ldots,k$ (see Section 1.3).

For the average of the function $S^0(x)$, it should be noted that

$$\bar{S}^0 = \frac{1}{2}\int_{-1}^{1} S^0(x)dx = \frac{1}{2}\sum_{\ell=0}^{k} s^0_{\ell 0}\int_{-1}^{1}\phi_\ell(x)dx = \frac{1}{\sqrt{2}}s^0_{00},$$

because $\phi_0(x) = \frac{1}{\sqrt{2}}, x\in[-1,1]$, and the scaling functions form an orthonormal basis.

For the averages $\bar{D}_j^m, m = 0, \ldots, n-1, j = 0, \ldots, 2^m - 1$, it holds that, using equation (51),

$$\begin{aligned}\bar{D}_j^m &= \frac{1}{2^{-m+1}} \int_{-1+2^{-m+1}j}^{-1+2^{-m+1}(j+1)} D_j^m(x)dx = 2^{m-1} \sum_{\ell=0}^{k} d_{\ell j}^m \int_{-1+2^{-m+1}j}^{-1+2^{-m+1}(j+1)} \psi_{\ell j}^m(x)dx \\ &= 2^{m-1} \sum_{\ell=0}^{k} d_{\ell j}^m \int_{-1+2^{-m+1}j}^{-1+2^{-m+1}(j+1)} 2^{m/2} \psi_\ell(2^m(x+1) - 2j - 1)dx \\ &= 2^{\frac{m}{2}-1} \sum_{\ell=0}^{k} d_{\ell j}^m \int_{-1}^{1} \psi_\ell(y)dy = 2^{\frac{m}{2}-1} \sum_{\ell=0}^{k} d_{\ell j}^m \int_{-1}^{1} \sqrt{2}\psi_\ell(y)\phi_0(y)dy = 0,\end{aligned}$$

because $V_0^k \perp W_0^k$.

Therefore, the following values for the coefficients θ_{S^0}, and $\theta_{D_j^m}, m = 0, \ldots, n-1$, $j = 0, \ldots, 2^m - 1$, are defined, using equations (49) and (50) (cf. equation (41b)):

$$\theta_{S^0} = \min\left\{ \left| \frac{M_{S^0}^0 - \frac{1}{\sqrt{2}} s_{00}^0}{M_{S^0} - \frac{1}{\sqrt{2}} s_{00}^0} \right|, \left| \frac{m_{S^0}^0 - \frac{1}{\sqrt{2}} s_{00}^0}{m_{S^0} - \frac{1}{\sqrt{2}} s_{00}^0} \right|, 1 \right\}, \tag{52a}$$

$$\theta_{D_j^m} = \min\left\{ \left| \frac{M_{D^m}^0}{M_{D_j^m}} \right|, \left| \frac{m_{D^m}^0}{m_{D_j^m}} \right|, 1 \right\}. \tag{52b}$$

The limiter for the scaling function, $S^0(x)$, looks very similar to the limiter in derivation (43):

$$\widetilde{S}^0(x) = \theta_{S^0}\left(S^0(x) - \frac{1}{\sqrt{2}} s_{00}^0 \right) + \frac{1}{\sqrt{2}} s_{00}^0 = s_{00}^0 \phi_0(x) + \sum_{\ell=1}^{k} \left(\theta_{S^0} \cdot s_{\ell 0}^0 \right) \phi_\ell(x),$$

such that

$$\widetilde{s}_{00}^0 = s_{00}^0, \text{ and } \widetilde{s}_{\ell 0}^0 = \theta_{S^0} \cdot s_{\ell 0}^0, \ell = 1, \ldots, k.$$

For the multiwavelets, $D_j^m(x), m = 0, \ldots, n-1, j = 0, \ldots, 2^m - 1$, the fact that $\bar{D}_j^m$ equals zero gives:

$$\widetilde{D}_j^m(x) = \theta_{D_j^m} D_j^m(x),$$

such that

$$\widetilde{d}_{\ell j}^m = \theta_{D_j^m} \cdot d_{\ell j}^m, \ell = 0, \ldots, k.$$

Unfortunately, the maximum principle based limiter does not work! First, I tried to use this limiter for the linear advection equation, with the square wave as an initial condition (equation (44)). The result for $T = 0.125$ can be seen in Figure 24.

Recall that the multiwavelet decomposition at $T = 0$ was given in Figure 14. The problem is that $D^m(x) \equiv 0$ for $m = 2, \ldots, 5$. This means that,

$$m^0_{D^m} = M^0_{D^m} = 0, \text{ for } m = 2, \ldots, 5,$$

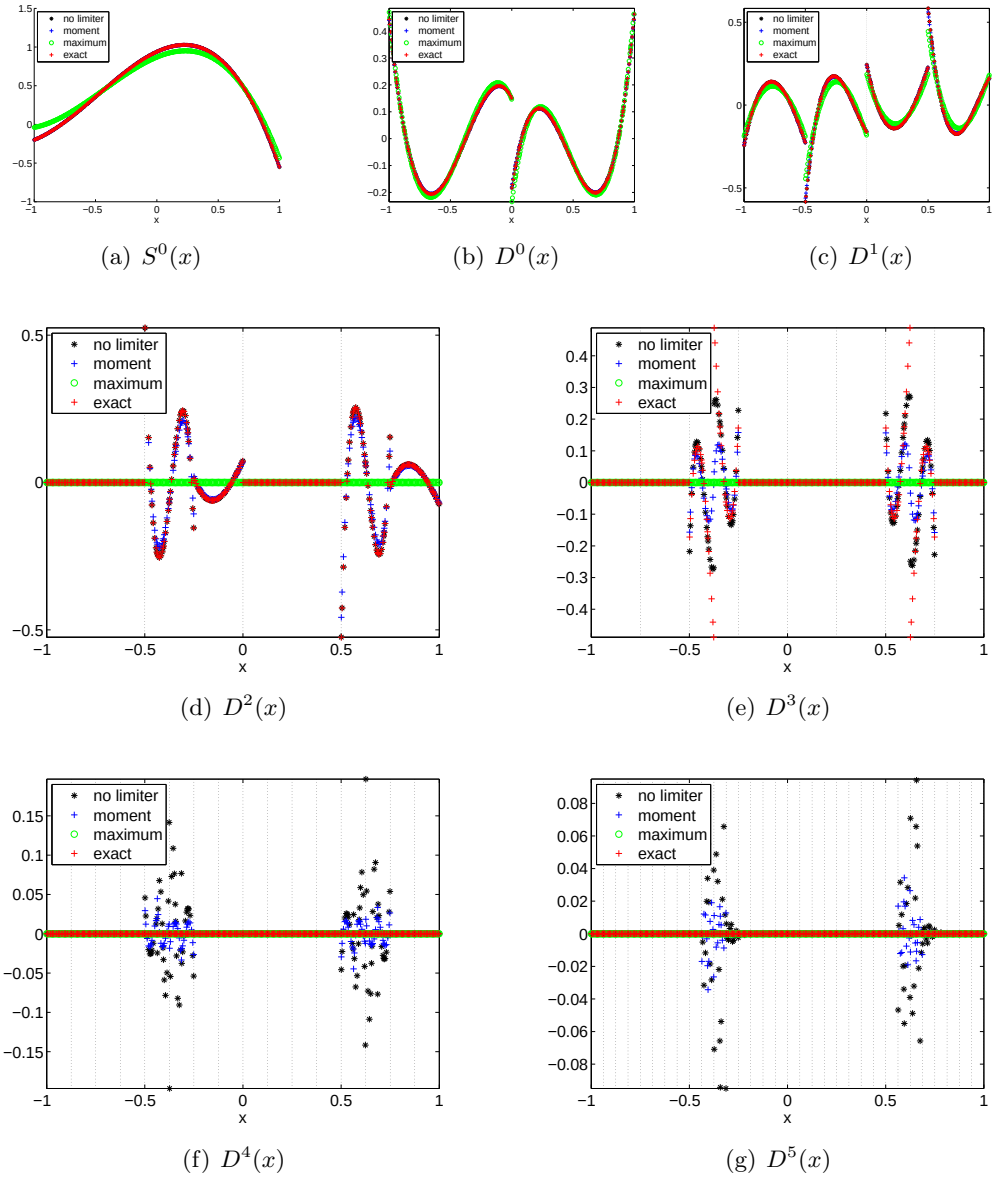
which results in

$$\theta_{D^m_j} = 0, \text{ for } m = 2, \ldots, 5, j = 0, \ldots, 2^m - 1.$$

This means that for each RK step, each level $m \in \{0, \ldots, n-1\}$ and each interval $I^m_j, j = 0, \ldots, 2^m - 1$, we set $d^m_{\ell j}$ to zero, $\ell = 0, \ldots, k$, such that the limiter produces a very inaccurate approximation.
This can be seen very clearly in Figures 23 and 24 for $T = 0.125$. Indeed, the resulting approximation is based on 4 elements only.
Maybe the maximum principle limiter on the multiwavelet decomposition works for functions which do not have zero contribution of the higher levels. The result at $T = 0.125$ using initial condition (38) is given in Figure 25. Looking at the results, it is clear that this limiter fails.

2.3.4 Summary and problems of adapting the multiwavelet coefficients

In Sections 2.3 and 2.2, I tried to find a new limiter for DG based on multiwavelets. We observed that setting the multiwavelet contribution equal to zero (Section 2.2) did not work: either new discontinuities are introduced, or the computation time grows because we need to project the approximation onto the space of scaled Legendre polynomials to find the limited DG coefficients. Setting some multiwavelet coefficients equal to zero is also not useful: the accuracy decreases (Section 2.3), because we effectively use less multiwavelet levels for computation ($m = 0, \ldots, \mu, \mu < n-1$). This is similar to reducing the number of elements that is used for computation (which can be seen very clear in Figures 21c, 23 and 24).
To summarize the search for a new limiter, we can be sure that the limiter should do more than simply setting the multiwavelet contribution/coefficients equal to zero.
However, I observed that the shocks were often very well detected. Therefore, the next section focuses on using the multiwavelets as a shock detector. A standard limiter (moment, maximum principle) is then applied in the detected shock elements only.



(a) $S^0(x)$ (b) $D^0(x)$ (c) $D^1(x)$

(d) $D^2(x)$ (e) $D^3(x)$

(f) $D^4(x)$ (g) $D^5(x)$

Figure 23: Multiwavelet decomposition of the approximation at $T = 0.125$, maximum principle limiter is applied on multiwavelet decomposition, linear advection equation with square wave, $k = 3, n = 6$ ($2^n = 64$ elements)

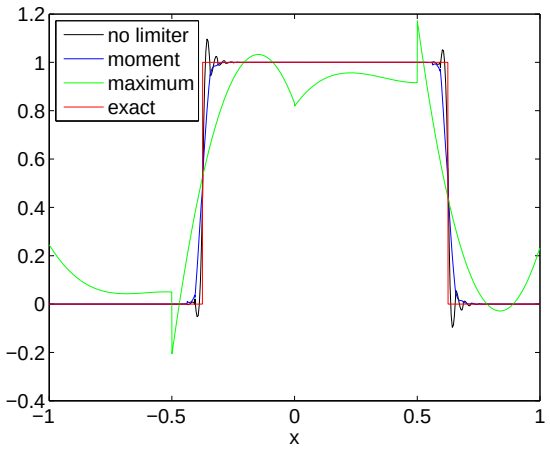


Figure 24: Approximation at $T = 0.125$, maximum principle limiter is applied on multiwavelet decomposition, linear advection equation with initial condition (44), $k = 3$, 64 elements

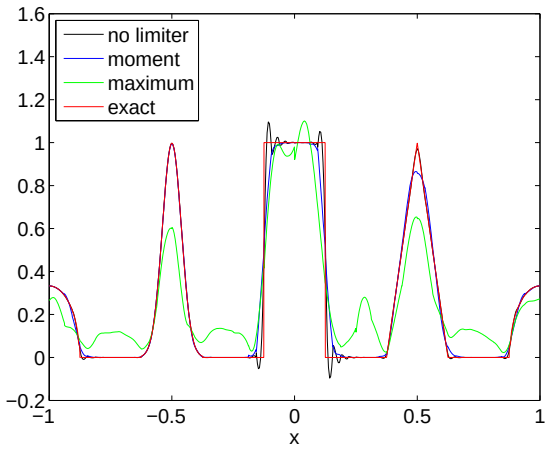


Figure 25: Approximation at $T = 0.125$, maximum principle limiter is applied on multiwavelet decomposition, linear advection equation with initial condition (38), $k = 3$, 64 elements

3 Shock detection methods

In this section, we focus on shock detection methods. Shock detection and limiting often go hand in hand, but it is also possible to separate these two steps. In the following sections, different shock detection methods using multiwavelets are considered (Sections 3.1 and 3.3). The DG coefficients can be used for shock detection as well, as we will discover in Section 3.2.

3.1 Take the ratio of the multiwavelet decomposition at different levels

The first idea for shock detection is to use the multiwavelet decomposition. The multiwavelet decomposition of a function gives an approximation at the coarsest level $m = 0$ $\left(S^0(x)\right)$ and high frequency details at levels $m = 0, \ldots, n-1$ $(D^m(x))$. Because a higher decomposition level corresponds to finer details, it may be possible that $\left|D^{m+1}(x)\right| \leq \left|D^m(x)\right|, x \in [a,b], m = 0, \ldots, n-2$. If this is true, it is expected that

$$\frac{\left|D^{m+1}(x)\right|}{\left|D^m(x)\right|} \leq 1. \tag{53}$$

Higher values for this ratio can arise in the neighbourhood of the boundaries in level $m+1$ (definition (14)), or if $D^m(x)$ is approximately zero. A threshold value ϵ is used to neglect ratio (53) if $|D^m(x)| < \epsilon$ for $x \in [a,b]$.
The question arises if inequality (53) holds. As an example, we look at initial condition (44) on [-1,1]. The multiwavelet contribution of the initial condition is zero for levels 2 and higher (Section 2.1), such that ratio (53) can be computed for $m = 0$ only. Because $\min_{x\in[-1,1]} \left|D^0(x)\right| = 0.0025$, thresholding is not necessary.
In Figure 26, the result can be seen. Unfortunately, statement (53) does not hold and is violated in many points. This idea does not work for the initial condition, so it is obvious that it can not be used for later times.

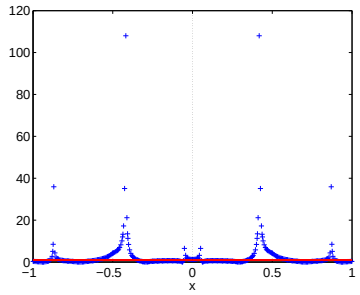


Figure 26: Ratio $\frac{\left|D^1(x)\right|}{\left|D^0(x)\right|}$, of initial condition (44). Red line: $y = 1$.

3.2 Shock detection using DG coefficients

In this section, shock detection is done using DG coefficients. In Section 3.2.1 the derivative inside an element is used for detecting a shock, and in Section 3.2.2 the derivative on the boundary of an element is looked at.

3.2.1 Derivative inside an element

By definition, the DG approximation, $u_h(x,t)$, is continuous on element I_j, $j = 0, \ldots, N$. Therefore, using the notation from Figure 2, it holds that,

$$u_h'(x_j) = \lim_{\Delta x \to 0} \frac{u_h(x^-_{j+\frac{1}{2}}) - u_h(x^+_{j-\frac{1}{2}})}{\Delta x}. \tag{54}$$

Because Δx is assumed to be small, this is replaced by (using equations (10)),

$$\begin{aligned} u_h'(x_j) &\approx \frac{u_h(x^-_{j+\frac{1}{2}}) - u_h(x^+_{j-\frac{1}{2}})}{\Delta x} = \frac{\sum_{\ell=0}^{k} u_j^{(\ell)}(\phi_\ell(1) - \phi_\ell(-1))}{\Delta x} \\ &= \frac{1}{\Delta x} \sum_{\ell=0}^{k} u_j^{(\ell)} \sqrt{\ell + \frac{1}{2}}(1 - (-1)^\ell). \end{aligned}$$

If ℓ is even, then $1 - (-1)^\ell = 0$, such that,

$$u_h'(x_j) \approx \begin{cases} 0, & \text{if } k = 0; \\ \frac{2}{\Delta x}\left(\sqrt{\frac{3}{2}} u_j^{(1)}\right), & \text{if } k = 1, 2; \\ \frac{2}{\Delta x}\left(\sqrt{\frac{3}{2}} u_j^{(1)} + \sqrt{\frac{7}{2}} u_j^{(3)}\right), & \text{if } k = 3, 4. \end{cases} \tag{55}$$

On the other hand, the derivative $u_h'(x_j)$ can be computed using the definition of the DG approximation, given by,

$$u_h(x) = \sum_{\ell=0}^{k} u_j^{(\ell)} \phi_\ell\left(\frac{2}{\Delta x}(x - x_j)\right), x \in I_j, j = 0, \ldots, N.$$

The derivative $u_h'(x)$ on element I_j is given by,

$$u_h'(x) = \sum_{\ell=0}^{k} u_j^{(\ell)} \frac{d}{dx} \phi_\ell\left(\frac{2}{\Delta x}(x - x_j)\right) = \frac{2}{\Delta x} \sum_{\ell=0}^{k} u_j^{(\ell)} \phi_\ell'(\xi),$$

where $\xi = \frac{2}{\Delta x}(x - x_j)$.

Using the definition of $\phi_\ell, \ell = 0, \ldots, 4$, as given in equations (3), the derivatives are given by,

$$\begin{aligned}
\phi_0'(\xi) &= 0;\\
\phi_1'(\xi) &= \sqrt{\frac{3}{2}};\\
\phi_2'(\xi) &= \frac{1}{2}\sqrt{\frac{5}{2}}6\xi;\\
\phi_3'(\xi) &= \frac{1}{2}\sqrt{\frac{7}{2}}(15\xi^2 - 3);\\
\phi_4'(\xi) &= \frac{1}{8}\sqrt{\frac{9}{2}}(140\xi^3 - 60\xi);
\end{aligned}$$

such that,

$$u_h'(x_j) = \begin{cases} 0, & \text{if } k = 0;\\ \frac{2}{\Delta x}\left(\sqrt{\frac{3}{2}}u_j^{(1)}\right), & \text{if } k = 1, 2;\\ \frac{2}{\Delta x}\left(\sqrt{\frac{3}{2}}u_j^{(1)} - \frac{3}{2}\sqrt{\frac{7}{2}}u_j^{(3)}\right), & \text{if } k = 3, 4. \end{cases} \tag{56}$$

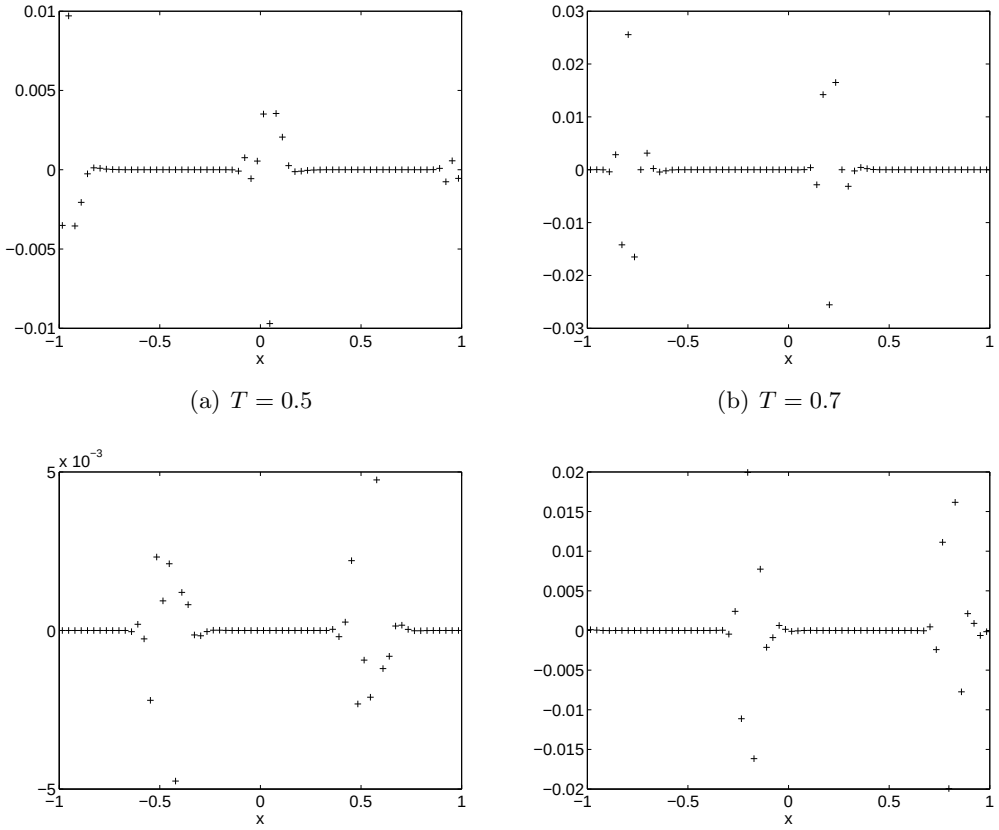
Comparing approximation (55) with equation (56), the definitions for $u_h'(x_j)$ are equal for $k = 0, 1, 2$, by construction. If $k = 3, 4$, then equation (55) equals (56) only if $u_j^{(3)} = 0$. If $|u_j^{(3)}|$ is small in element $I_j, j \in \{0, \ldots, N\}$, then we expect that the exact solution is continuous in this element.

Therefore, we propose to use a shock detector that computes for which element I_j, the DG coefficient $u_j^{(3)}$ is too big, $j \in \{0, \ldots, N\}$. We expect that the exact solution has a discontinuity inside this element.

As an example, the linear advection equation with the square wave as an initial condition is looked at. In Figure 27 the value of $u_j^{(3)}$ can be seen in each element I_j, at different times, $j = 0, \ldots, N$. Indeed, we can recognize the discontinuities in the solution, using this approach.

An important drawback of the shock detector, described in this section, is that this detector uses the DG coefficients of degree 3. Therefore, it can only be applied for higher order DG, using $k = 3$ or $k = 4$.

In Section 4.1.1, we combine this shock detector with a limiter and look at the results.



(a) $T = 0.5$ (b) $T = 0.7$

(c) $T = 1$ (d) $T = 1.3$

Figure 27: The values of $u_j^{(3)}$ of the DG approximation, $j = 0, \ldots, N$, linear advection equation with initial condition (44), $k = 3$, 64 elements

3.2.2 Derivative on the boundary of an element

Besides looking at the derivative inside an element I_j, it is possible to look at the derivative on the boundaries $x_{j\pm\frac{1}{2}}, j = 0, \ldots, N$. By definition, the DG approximation is discontinuous on the element boundary. In general, this discontinuity is not so big, because the exact solution is assumed to be continuous. Therefore, the derivative on the boundary $x_{j+\frac{1}{2}}, j \in \{0, \ldots, N-1\}$, can be approximated by,

$$\begin{aligned} u_h'(x_{j+\frac{1}{2}}) &= \lim_{\Delta x \to 0} \frac{u_h(x_{j+1}) - u_h(x_j)}{\Delta x} \approx \frac{u_h(x_{j+1}) - u_h(x_j)}{\Delta x} \\ &= \frac{\sum_{\ell=0}^{k} u_{j+1}^{(\ell)} \phi_\ell(0) - \sum_{\ell=0}^{k} u_j^{(\ell)} \phi_\ell(0)}{\Delta x} = \frac{1}{\Delta x} \sum_{\ell=0}^{k} \phi_\ell(0) \left(u_{j+1}^{(\ell)} - u_j^{(\ell)} \right). \end{aligned} \tag{57}$$

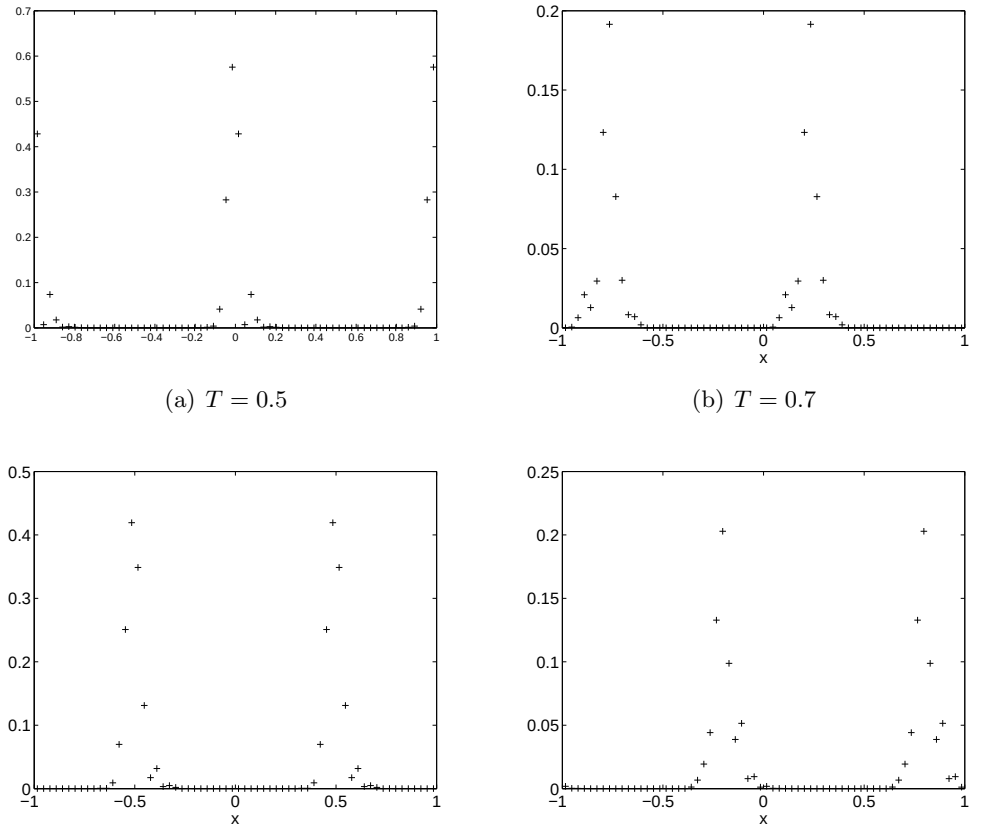
On the other hand, if the function is continuous on the boundary, the derivative can be approximated by,

$$u_h'(x_{j+\frac{1}{2}}) \approx \frac{1}{2} \left(u_h'(x_{j+\frac{1}{2}}^-) + u_h'(x_{j+\frac{1}{2}}^+) \right) = \frac{1}{2} \frac{2}{\Delta x} \left(\sum_{\ell=0}^{k} u_j^{(\ell)} \phi_\ell'(1) + \sum_{\ell=0}^{k} u_{j+1}^{(\ell)} \phi_\ell'(-1) \right). \tag{58}$$

If the distance $\left| u_h(x_{j+\frac{1}{2}}^-) - u_h(x_{j+\frac{1}{2}}^+) \right|$ is small, it is expected that the difference between (57) and (58), given by,

$$\Delta u'_{j+\frac{1}{2}} \equiv \left| \sum_{\ell=0}^{k} \phi_\ell(0) \left(u_{j+1}^{(\ell)} - u_j^{(\ell)} \right) - \sum_{\ell=0}^{k} \left(u_j^{(\ell)} \phi_\ell'(1) + u_{j+1}^{(\ell)} \phi_\ell'(-1) \right) \right|, \tag{59}$$

is small. Difference (59) will be big in the neighbourhood of a serious discontinuity.
In Figure 28, the value of equation (59) can be seen for the unlimited solution of the linear advection equation, using the square wave as an initial condition. It is very easy to recognize the discontinuities, so the value of equation (59) seems to be a useful shock detector.
Note that both equation (57) and (58) are approximations of the derivative on the element boundary: the derivative itself is undefined, because the DG approximation is discontinuous on element boundaries. This is one disadvantage of this shock detector.
Section 4.1.1 contains examples, where the shock detector is applied in combination with a limiter.



(a) $T = 0.5$ (b) $T = 0.7$

(c) $T = 1$ (d) $T = 1.3$

Figure 28: The value of $\Delta u'_{j+\frac{1}{2}}, j = 0, \ldots, N-1$, for the DG approximation, linear advection equation with initial condition (44), $k = 3$, 64 elements

3.3 Shock detection using multiwavelets

In this section, we look back at Figures 10, 16 and 23, in which the multiwavelet decomposition of a discontinuous function is given. These figures have a special structure: in the neighbourhood of a discontinuity, the multiwavelet contribution of the higher levels suddenly becomes big with respect to this contribution in the continuous regions. This means that multiwavelets can act as a shock detector! Therefore, I propose to use a limiter only in these elements where the average of the multiwavelet contribution of the highest level $n-1$ is maximal.
For each RK step, I compute the averages $\bar{D}_j^{n-1}$ of the multiwavelet contribution in element $I_j^{n-1}, j = 0, \ldots, 2^{n-1}-1$. If the maximum value of the average is found in I_J^{n-1}, for some $J \in \{0, \ldots, 2^{n-1}-1\}$, i.e. $\bar{D}_J^{n-1} = \max_{j=0,\ldots,2^{n-1}-1} \bar{D}_j^{n-1}$, then we expect that there is a discontinuity somewhere in the elements I_{2J} and I_{2J+1}, because $I_J^{n-1} = I_{2J}^n \cup I_{2J+1}^n$ (definition (14) and Figure 3 in Section 1.3). Therefore, a limiter can be used in these two elements I_{2J} and I_{2J+1} only.

Besides this, we can use a combination of the information from levels $n-1$ and $n-2$, as we saw in the decompositions that level $n-2$ contains information about the discontinuities as well. In that case we compute the averages $\bar{D}_j^{n-2}, j = 0, \ldots, 2^{n-2}-1$, and mark the elements $I_{4J}, \ldots, I_{4J+3}$ to be discontinuous if $\bar{D}_J^{n-2} = \max_{j\in\{0,\ldots,2^{n-2}-1\}} \bar{D}_j^{n-2}$.

A disadvantage of this limiter, is that the multiwavelet decomposition costs a lot of computation time.
In Section 4.1.2, the performance of this shock detector, in combination with a limiter, is examined.

4 Combine shock detectors with existing limiters

In this section, we look at the performance of the new shock detectors of Sections 3.2 and 3.3, in combination with an existing limiter. In Section 4.1 two test problems are used: the linear advection equation and the inviscid Burgers' equation. In Sections 4.2 and 4.3 we look at the one dimensional Euler equations, together with Sod's or Lax's initial conditions.

4.1 Linear advection equation and inviscid Burgers' equation

First, we consider two test problems. The first one is the linear advection equation on $[-1,1]$, with the square wave as an initial condition. The square wave contains two discontinuities, but they are of equal length. This means that the shock detectors (which use absolute values) always detect both discontinuities simultaneously.
The second test problem is the inviscid Burgers' equation on $[-1,1]$, using initial condition (28). For this equation, one shock is found for $T > \frac{2}{\pi}$, as we argued in Section 1.4.1.

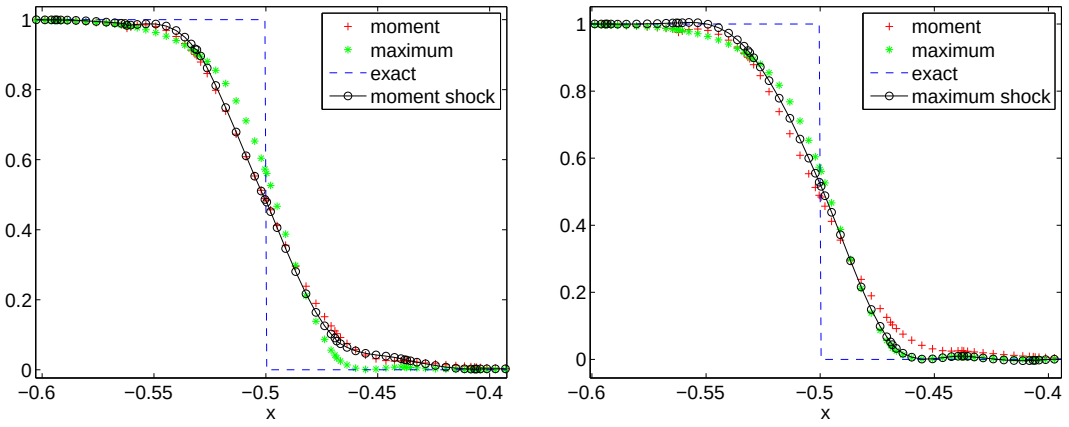
4.1.1 Shock detectors based on DG coefficients

In Section 3.2 we constructed two shock detectors based on the DG coefficients. In this section, we combine these shock detectors with existing limiters: the limiter is applied only in discontinuous regions.
First of all, we want to use the shock detector, based on $|u_j^{(3)}|, j = 0, \ldots, N$. We expect discontinuous regions in these elements where $|u_j^{(3)}|$ is big. Note that we need to make a choice about the threshold $U > 0$, such that I_j is marked as a discontinuous element if $|u_j^{(3)}| \geq U$. I found the best results using $U = \max_{j\in\{0,\ldots,N\}} |u_j^{(3)}|$, which means that the limiter is applied only on element I_j where $|u_j^{(3)}|$ is maximal, $j = 0, \ldots, N$.
For the linear advection equation, there are allways two different elements I_{j_1} and $I_{j_2}, j_1, j_2 \in \{0, \ldots, N\}$ for which $|u_{j_1}^{(3)}| = |u_{j_2}^{(3)}| = \max_{j=0,\ldots,N} |u_j^{(3)}|$, because the square wave is symmetric. In these two elements, I used either the moment limiter (Figure 29a) or the maximum principle limiter (Figure 29b). Inspecting the results, we see that the combination with the maximum principle limiter works best: there is a great improvement for $x \in [-0.6, -0.53]$ in Figure 29b.
Recall that for the inviscid Burgers' equation, the maximum principle limiter fails. This is the reason why I use the moment limiter in the shock elements ($|u_j^{(3)}|$ is maximal) for this problem. The results for $T = 1$ and $T = 3$ are given in Figure 30. Also in this case, the new limiter (moment limiter applied in shock elements where $|u_j^{(3)}|$ is maximal) gives better results in some regions. Note that for $T = 3$, both the standard moment limiter and the combination with the shock detector give bad results in the region $[0.52, 0.54]$.

The second shock detector using DG coefficients was defined in Section 3.2.2. We expect a discontinuity in the elements $I_j, I_{j+1}, j \in \{0, \ldots, N-1\}$, for which $\Delta u'_{j+\frac{1}{2}}$ (equation (59)) is maximal. The limiter is then applied in both elements I_j and I_{j+1}.
In Figure 31 the results are computed for the linear advection equation with the square wave. The moment limiter (Figure 31a) or the maximum principle limiter (Figure 31b)

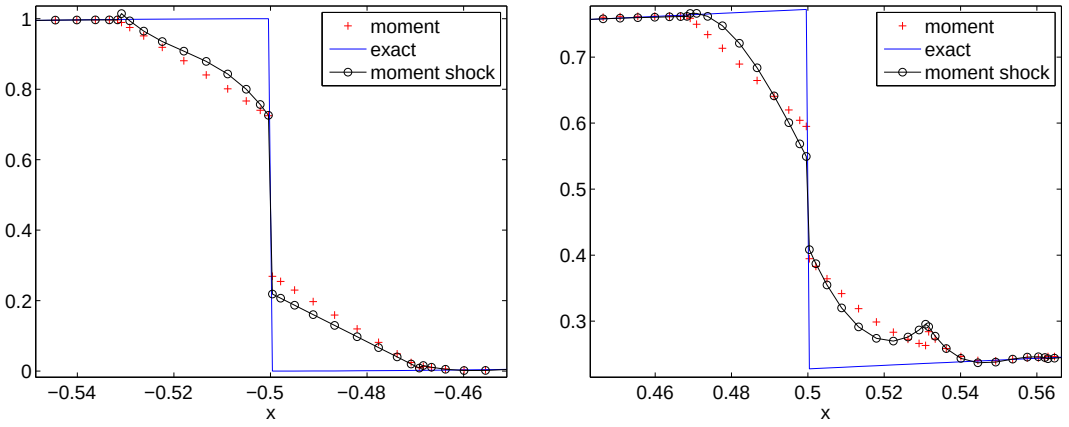
is used in the shock elements. For the inviscid Burgers' equation with initial condition (28), the results using this shock detector in combination with the moment limiter can be seen in Figure 32.

The results look very similar to the results in Figures 29 and 30, where we used the value of $|u_j^{(3)}|$ for shock detection. To compare the results for the shock detectors, based on the DG coefficients, we inspect Figures 33 (linear advection equation, shock detectors combined with maximum principle limiter) and 34 (inviscid Burgers' equation, shock detectors combined with moment limiter). In both figures, the first shock detector, using $|u_j^{(3)}|, j = 0, \ldots, N$, gives the best results.



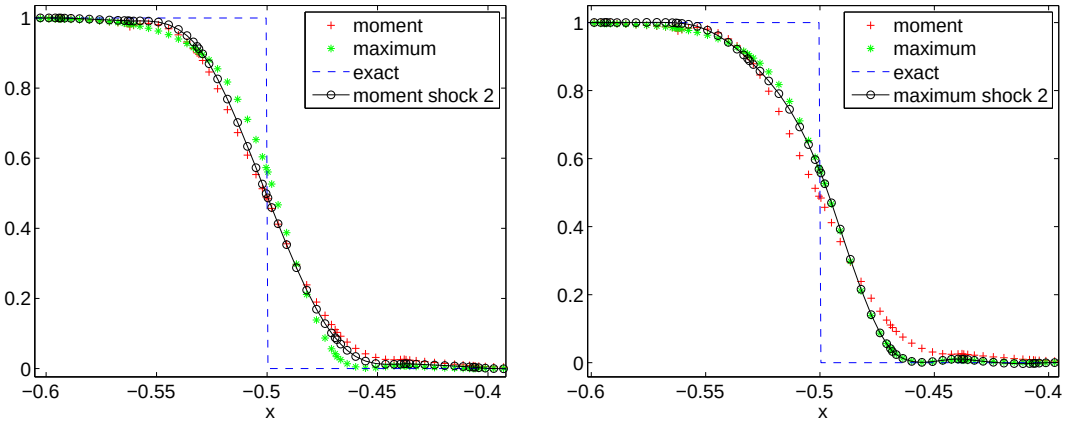
(a) Moment limiter in shock element (b) Maximum principle limiter in shock element

Figure 29: Results for the linear advection equation, square wave, at $T = 1$. Limiter is applied only in element I_j, where $|u_j^{(3)}|$ is maximal, $j = 0, \ldots, N, k = 3$, 64 elements



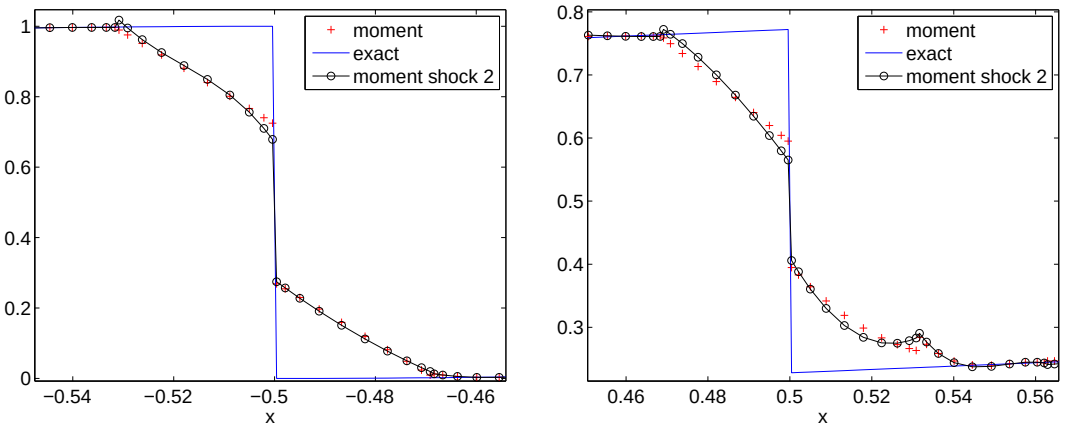
(a) $T = 1$ (b) $T = 3$

Figure 30: Results for inviscid Burgers' equation, initial condition (28). Moment limiter is applied only in element I_j, where $|u_j^{(3)}|$ is maximal, $j = 0, \ldots, N, k = 3$, 64 elements



(a) Moment limiter in shock elements (b) Maximum principle limiter in shock elements

Figure 31: Results for the linear advection equation, square wave, at $T = 1$. Limiter is applied only in elements I_j and I_{j+1} for which $\Delta u'_{j+\frac{1}{2}}$ is maximal, $j = 0, \ldots, N-1, k = 3$, 64 elements



(a) $T = 1$ (b) $T = 3$

Figure 32: Results for the inviscid Burgers' equation, initial condition (28). Moment limiter is applied only in elements I_j and I_{j+1} for which $\Delta u'_{j+\frac{1}{2}}$ is maximal, $j = 0, \ldots, N-1, k = 3$, 64 elements

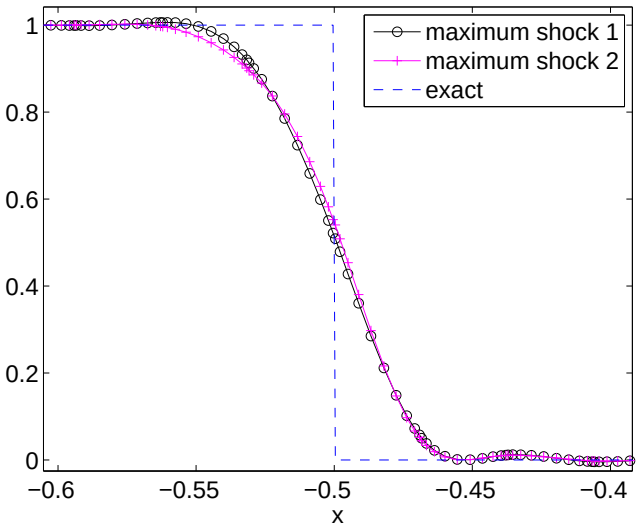


Figure 33: Results for the linear advection equation using square wave, at $T = 3$. Shock detector 1 is based on $|u_j^{(3)}|$, shock detector 2 is based on $\Delta u'_{j+\frac{1}{2}}$, combined with maximum principle limiter, $k = 3$, 64 elements

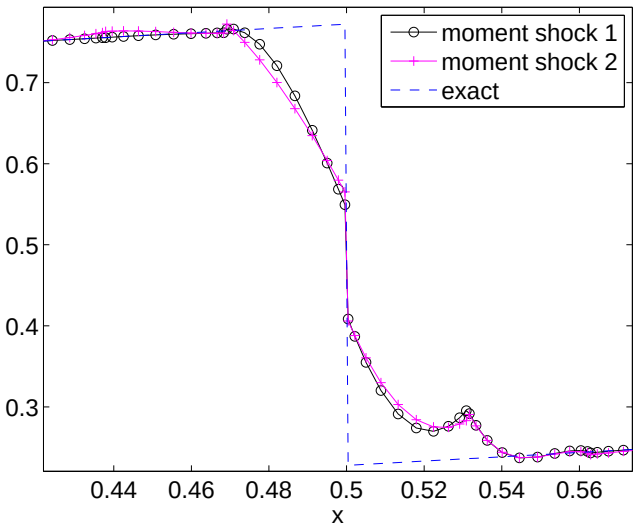


Figure 34: Results for the inviscid Burgers' equation using initial condition (28), $T = 3$. Shock detector 1 is based on $|u_j^{(3)}|$, shock detector 2 is based on $\Delta u'_{j+\frac{1}{2}}$, combined with moment limiter, $k = 3$, 64 elements

4.1.2 Multiwavelet shock detector

In this section, the results using multiwavelets as a shock detector (Section 3.3), in combination with a limiter, are examined. The results are computed using the maximum principle limiter for the linear advection equation (Figure 35), and the moment limiter for the inviscid Burgers' equation (Figure 36) in the shock elements I_{2J} and I_{2J+1}, for which $\bar{D}_J^{n-1} = \max_{j\in\{0,\ldots,2^{n-1}-1\}} \bar{D}_j^{n-1}$ (the average of multiwavelet contribution in level $n-1$ is maximal).

If we include level $n-2$ for the shock detection as well (using the averages $\bar{D}_j^{n-2}$), the results are nearly equal to the original approximations, where the limiter is applied on every element. This means that we limit the same elements as the limiter itself does, and we do not add anything using this shock detector.

A comparison between the DG coefficient shock detector based on $|u_j^{(3)}|, j = 0, \ldots, N$, and the multiwavelet shock detector is done in Figures 37 and 38. Note that for the linear advection equation, both approximations produce a small wiggle around $x = -0.56$ (Figure 37b). The results using the shock detector based on $|u_j^{(3)}|, j = 0, \ldots, N$, are worse in this region. The same holds for the results of the inviscid Burgers' equation (Figure 38) around $x = 0.47$.
For the region $[0.5, 0.53]$ in Figure 38, the shock detector based on $|u_j^{(3)}|$ gives the best results, but the differences are small.

In Sections 4.2 and 4.3, we look at the behaviour of the various shock detectors for Sod's and Lax's shock tubes.

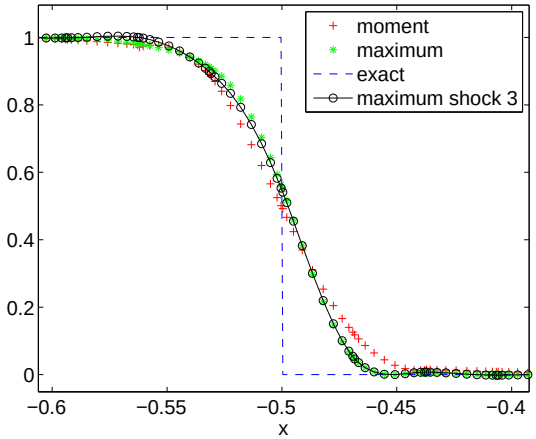


Figure 35: Results for the linear advection equation, square wave, at $T = 3$. Maximum principle limiter is applied in elements where average of multiwavelet contribution in level $n - 1$ is maximal, $k = 3$, 64 elements

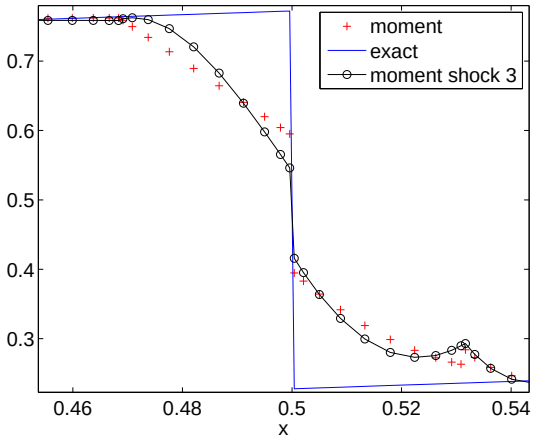
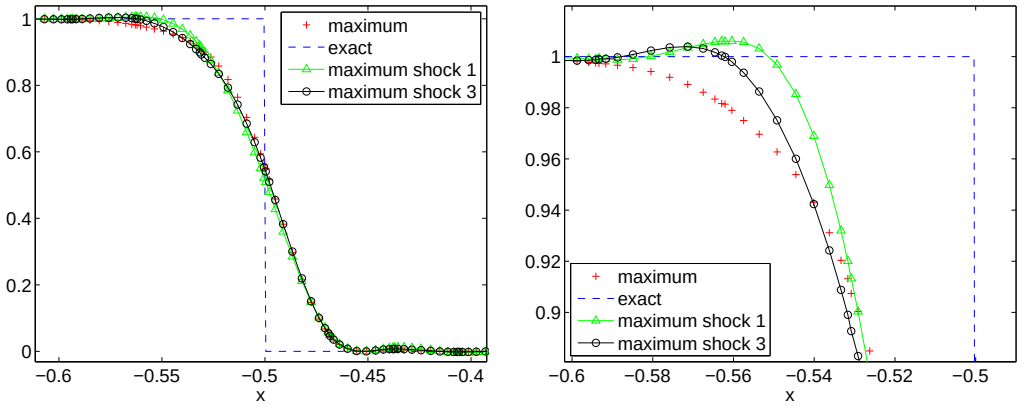


Figure 36: Results for the inviscid Burgers' equation, initial condition (28), $T = 3$. Moment limiter is applied in elements where average of multiwavelet contribution in level $n - 1$ is maximal, $k = 3$, 64 elements



(a) $T = 3$ (b) $T = 3$, extra zoom

Figure 37: Results for the linear advection equation, square wave. Shock detector 1 is based on $|u_j^{(3)}|$, and shock detector 3 is based on $\bar{D}_j^{n-1}$, combined with maximum principle limiter, $k = 3$, 64 elements

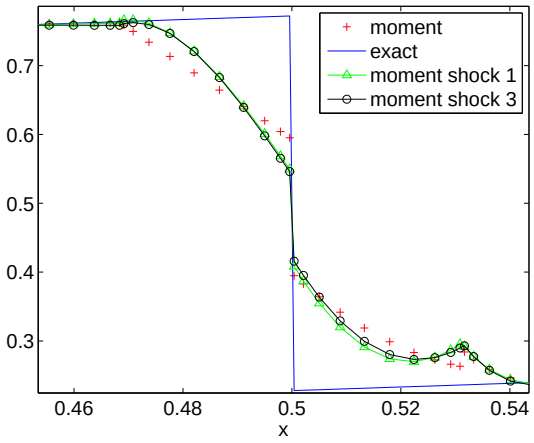


Figure 38: Results for the inviscid Burgers' equation, initial condition (28), $T = 3$. Shock detector 1 is based on $|u_j^{(3)}|$, and shock detector 3 is based on $\bar{D}_j^{n-1}$, combined with the moment limiter, $k = 3$, 64 elements

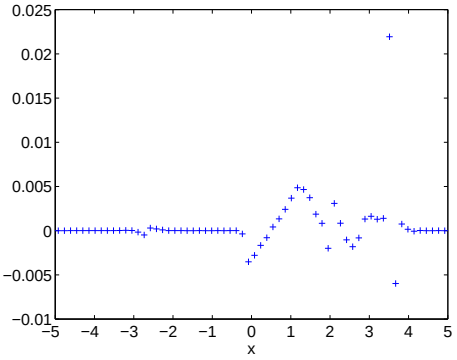
4.2 Sod

In the previous sections, we looked at the results using a shock detector in combination with an existing limiter. There, we used simple test problems, namely the linear advection equation and the inviscid Burgers' equation. Recall that for these problems, there was only one discontinuity (Burgers' equation), or two discontinuities of equal length (square wave, same maximal value of $|u_j^{(3)}|$ and $\Delta u'_{j+\frac{1}{2}}$). In that case, the discontinuities where detected by computing for which element I_j, the value of $|u_j^{(3)}|$ was maximal, $j = 0, \ldots, N$ (or for which elements I_j and I_{j+1}, $\Delta u'_{j+\frac{1}{2}}$ was maximal, $j = 0, \ldots, N-1$).

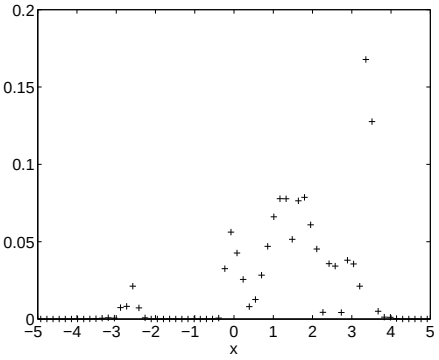
In this section, we look at a more difficult problem: the one dimensional Euler equations, in combination with Sod's initial conditions (shock tube), as defined in Section 1.4.2, equations (35). Without using a limiter, the approximation is very oscillatory, as we saw already in Figure 11. The moment limiter performs well, but we try to improve this limiter, by adding a shock detector.
Note that for Sod's shock tube, there are several discontinuities (and there are continuous regions where oscillations arise as well). This has consequences for the shock detectors that use DG coefficients (Section 3.2). As an example, we look at Figure 39, where the unlimited approximation is used to explore the working of the shock detectors. The left figures show the values of the DG coefficients $u_j^{(b,3)}, j = 0, \ldots, N$, $b = 1, 2, 3$, (where b denotes which variable $u^{(b)}$ is considered, as defined in Section 1.4.2 and Appendix A.4). The right figures visualize the values of $\Delta u'_{j+\frac{1}{2}}, j = 0, \ldots, N-1$. In each plot, we can easily recognize the shock at $x \approx 3.5$, because a global maximum is reached in that region. The local maximum at $x \approx -2.6$ belongs to the rarefaction wave (recall Figure 6b). The other discontinuities and oscillatory regions are difficult to locate exactly.
To summarize these observations: the DG coefficients can play a role in shock detection, but it is not easy to determine where the discontinuities and oscillatory regions are found.

The multiwavelet decomposition however, works excellent in detecting the various discontinuities! Recall that we compute the DG approximation for the density, momentum, and energy. For the multiwavelet shock detector, I only use the information from ρ, because we know (see Figure 6 in Section 1.4.2) that the location of the rarefaction wave, contact discontinuity, and shock, is the same for each variable.
The multiwavelet decomposition is given in Figure 40 for $T = 0.5$, Figure 41 for $T = 1$, and Figure 42 for $T = 2$, for $\rho_h = u_h^{(1)}$. I only plotted levels $m = 3, 4, 5$, and marked the elements where $\bar{D}_j^m$ is maximal.

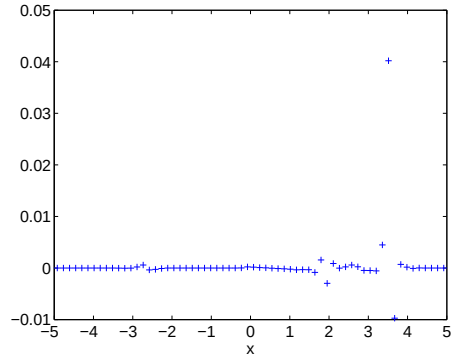
It is interesting to see that each level perfectly recognizes one of the discontinuities (and for small times, the rarefaction wave is selected as well)! It seems like level $n-1$ always selects the shock, but there is no pattern in the detected elements on levels $n-2$ and $n-3$. More research should be done on this subject.



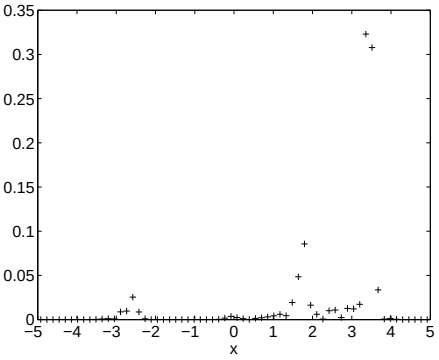
(a) $u_j^{(1,3)}$, belong to density ρ



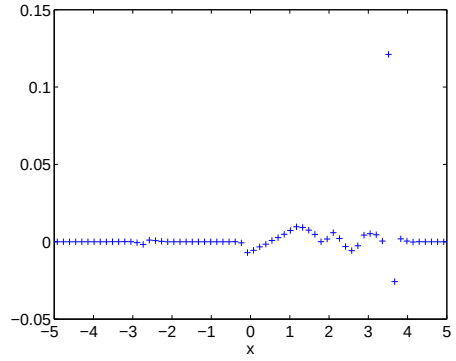
(b) $\Delta u'_{j+\frac{1}{2}}$ for ρ



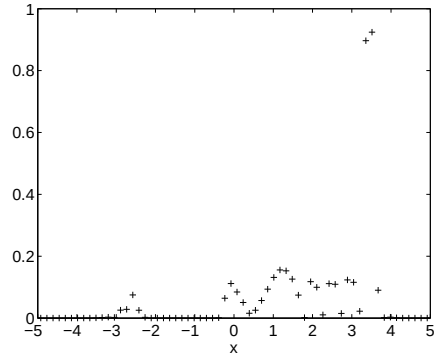
(c) $u_j^{(2,3)}$, belong to momentum ρu



(d) $\Delta u'_{j+\frac{1}{2}}$ for ρu

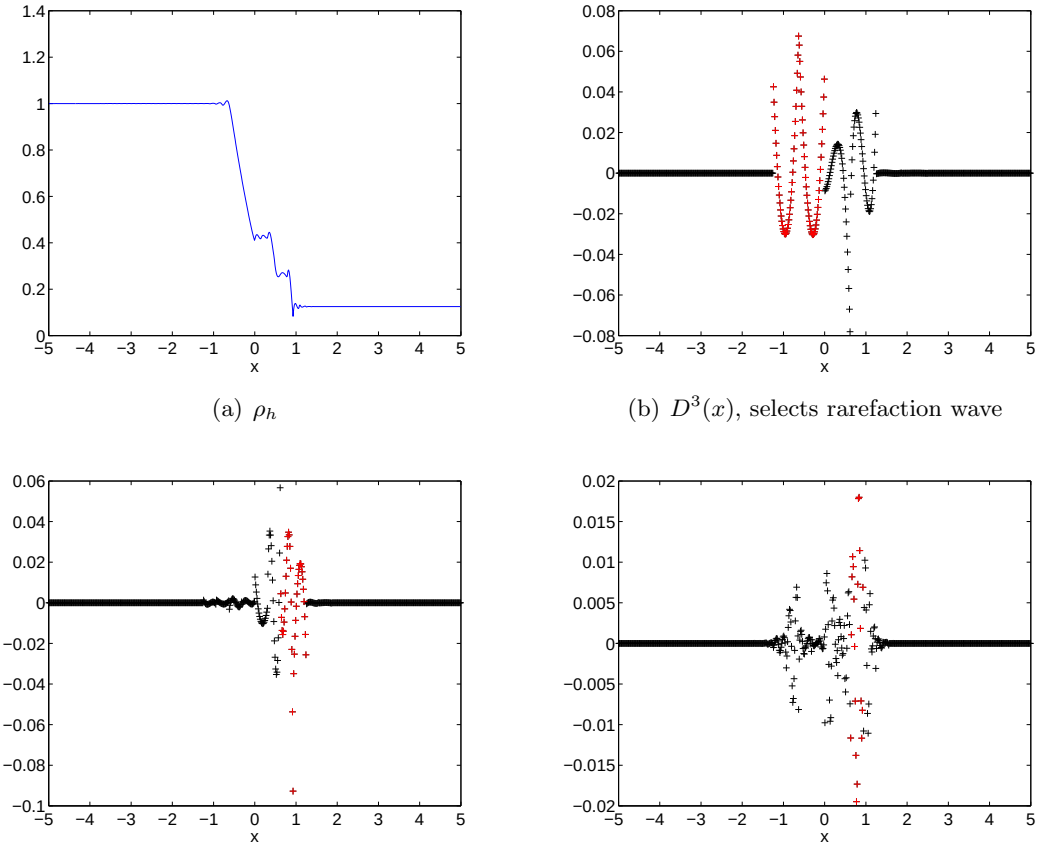


(e) $u_j^{(3,3)}$, belong to energy E



(f) $\Delta u'_{j+\frac{1}{2}}$ for E

Figure 39: Values of $u_j^{(b,3)}, j = 0, \ldots, N, b = 1, 2, 3$, and $\Delta u'_{j+\frac{1}{2}}, j = 0, \ldots, N-1$, DG approximation of Euler equations with Sod's initial conditions, $T = 2, k = 3$, 64 elements.



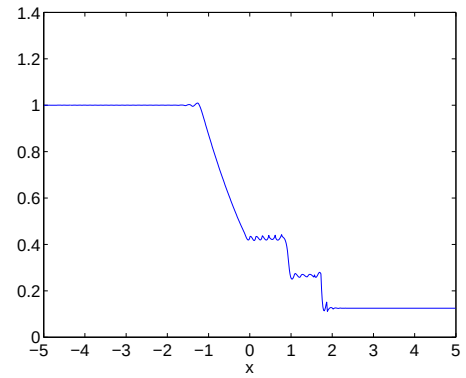
(a) ρ_h

(b) $D^3(x)$, selects rarefaction wave

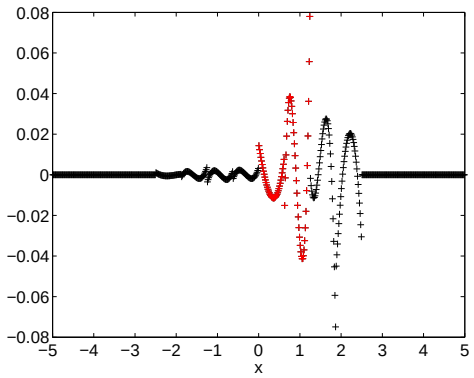
(c) $D^4(x)$, selects shock

(d) $D^5(x)$, selects shock

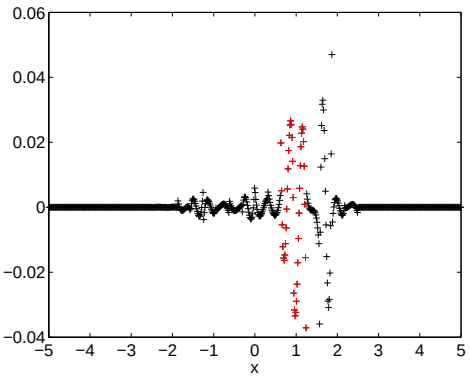
Figure 40: Multiwavelet decomposition of DG approximation, Sod's shock tube, $T = 0.5$, $k = 3, n = 6, 2^n = 64$ elements for ρ_h



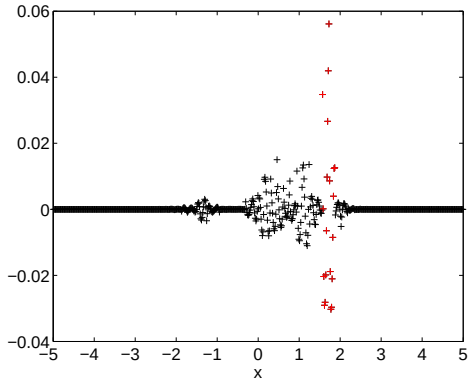
(a) ρ_h



(b) $D^3(x)$, selects contact discontinuity

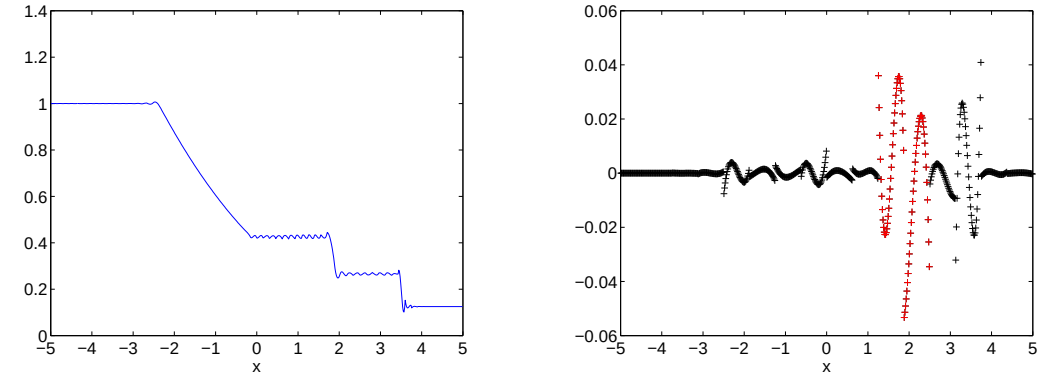


(c) $D^4(x)$, selects contact discontinuity



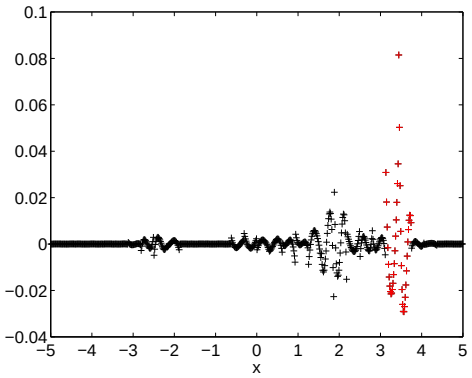
(d) $D^5(x)$, selects shock

Figure 41: Multiwavelet decomposition of DG approximation, Sod's shock tube, $T = 1$, $k = 3, n = 6, 2^n = 64$ elements for ρ_h

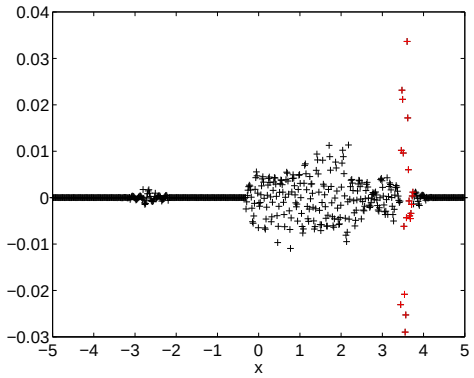


(a) ρ_h

(b) $D^3(x)$, selects contact discontinuity



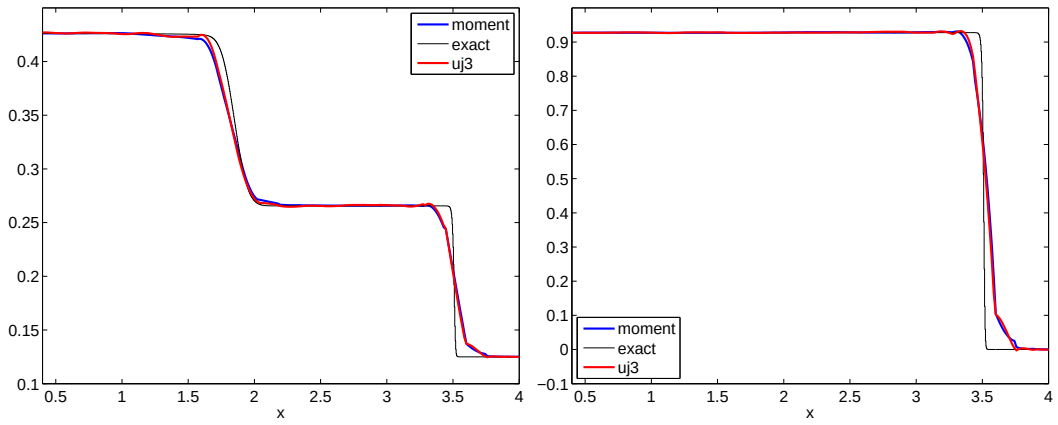
(c) $D^4(x)$, selects shock



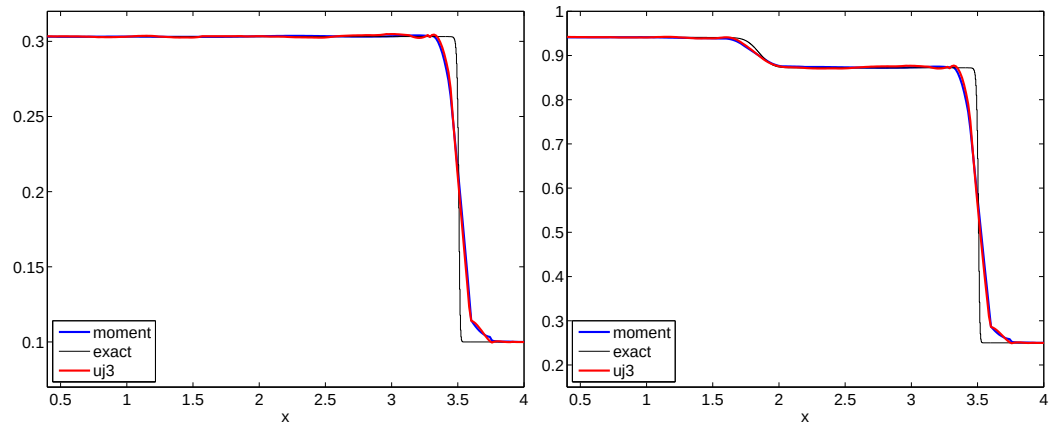
(d) $D^5(x)$, selects shock

Figure 42: Multiwavelet decomposition of of DG approximation, Sod's shock tube, $T = 2$, $k = 3, n = 6, 2^n = 64$ elements for ρ_h

Next, we apply the different shock detectors in combination with the moment limiter. In Figures 43-47, the results are given using each shock detector. We focus on the contact discontinuities and shocks, because the oscillations in the neighbourhood of the rarefaction wave are very small, and the impact of the shock detector as well.
When we compare the results, we see that the multiwavelet shock detector on level $n-1$ (Figure 45) and the one on levels $n-2$ and $n-1$ (Figure 46) introduce new oscillations, which we do not want to get. If we also include multiwavelet level $n-3$ for shock detection (Figure 47), we get approximately the same results as the moment limiter on all elements does (we limit many elements). It seems like the shock detector using $\Delta u'_{j+\frac{1}{2}}, j = 0, \ldots, N-1$, gives the best results (Figure 44).

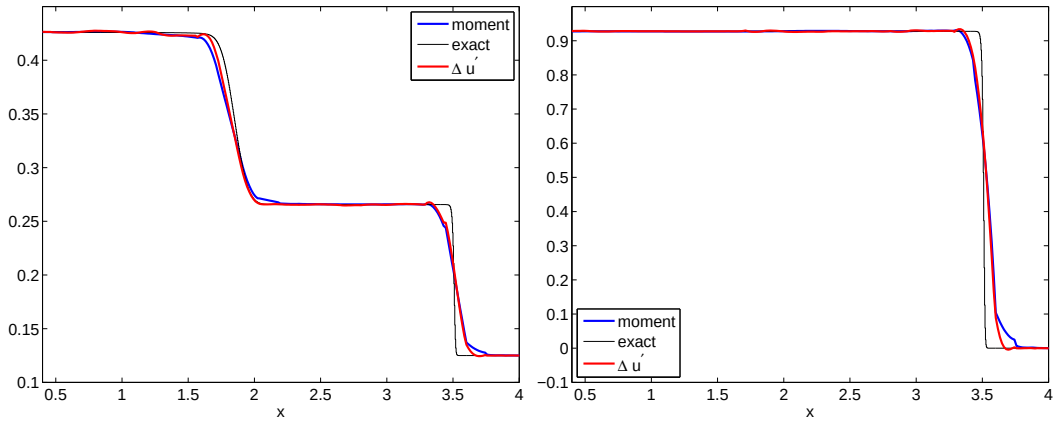


(a) Density ρ (b) Velocity u

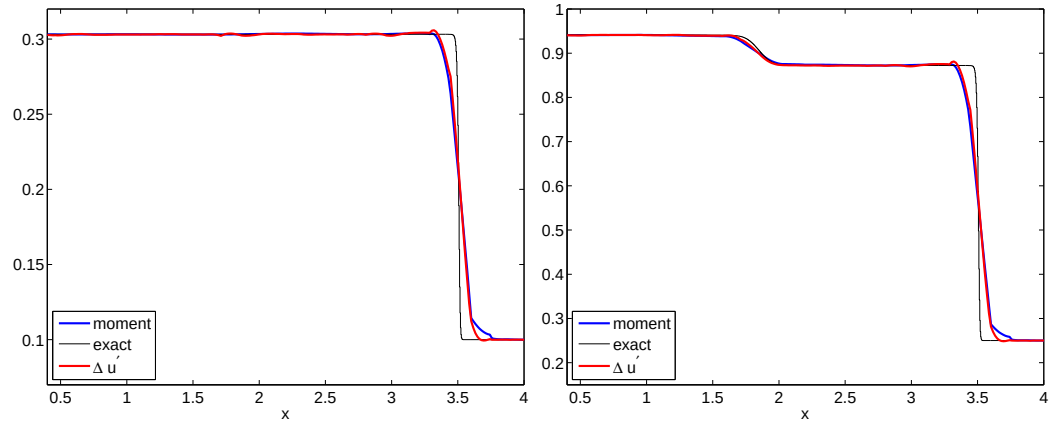


(c) Pressure p (d) Energy E

Figure 43: Results of Euler equations with Sod's initial conditions, absorbing boundary conditions, $T = 2, k = 3$, 64 elements. Moment limiter is applied in elements where $|u_j^{(3)}|$ is maximal, $j = 0, \ldots, N$. Exact solution is obtained using 3000 elements and $k = 0$.

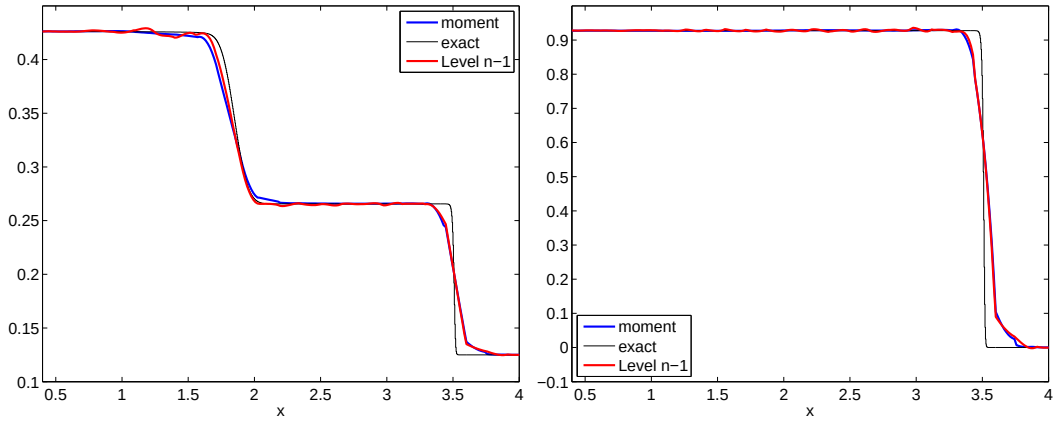


(a) Density ρ (b) Velocity u

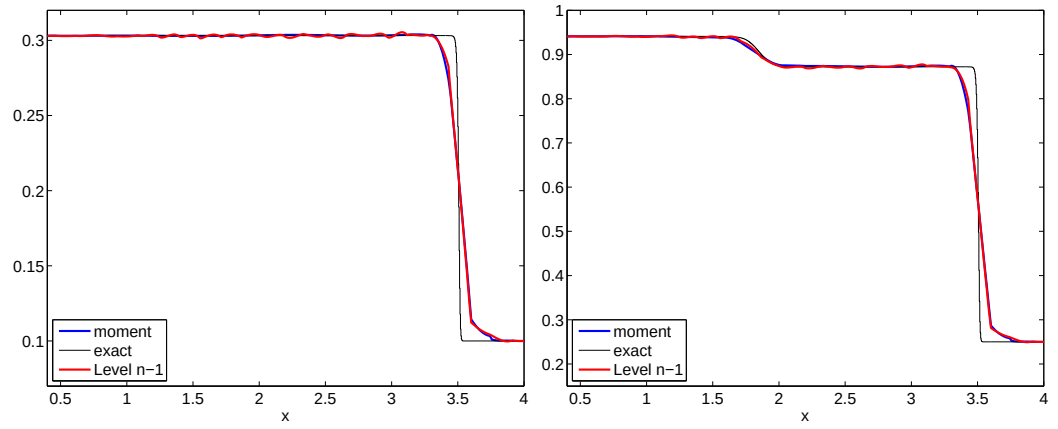


(c) Pressure p (d) Energy E

Figure 44: Results of Euler equations with Sod's initial conditions, absorbing boundary conditions, $T = 2, k = 3$, 64 elements. Moment limiter is applied in elements where $\Delta u'_{j+\frac{1}{2}}$ is maximal, $j = 0, \ldots, N-1$. Exact solution is obtained using 3000 elements and $k = 0$.

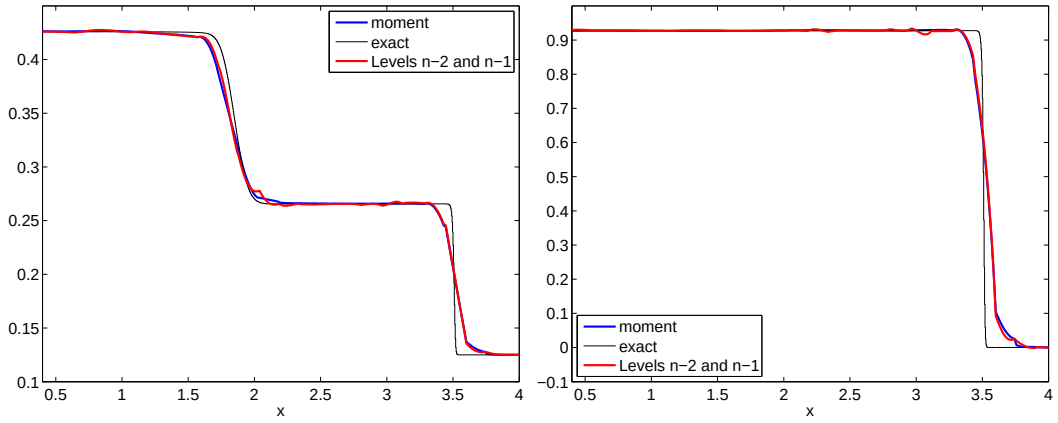


(a) Density ρ (b) Velocity u

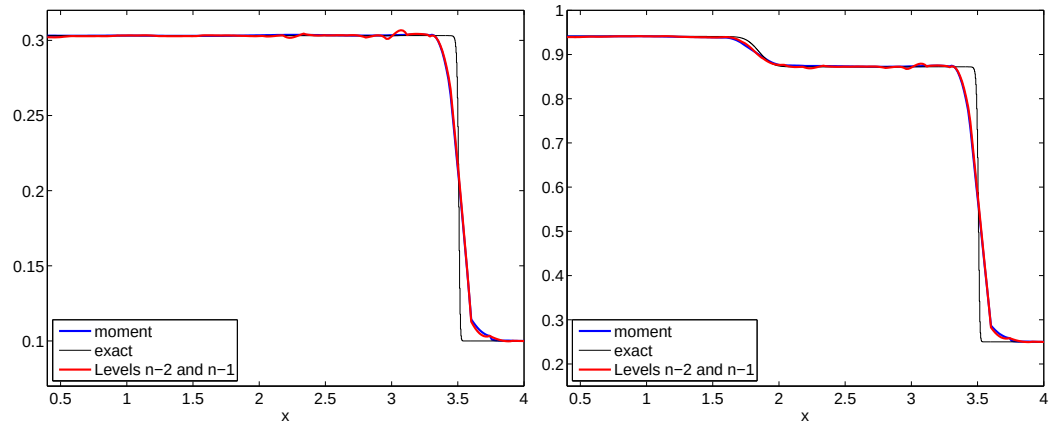


(c) Pressure p (d) Energy E

Figure 45: Results of Euler equations with Sod's initial conditions, absorbing boundary conditions, $T = 2, k = 3$, 64 elements. Moment limiter is applied in elements where $\bar{D}^{n-1}$ is maximal. Exact solution is obtained using 3000 elements and $k = 0$.

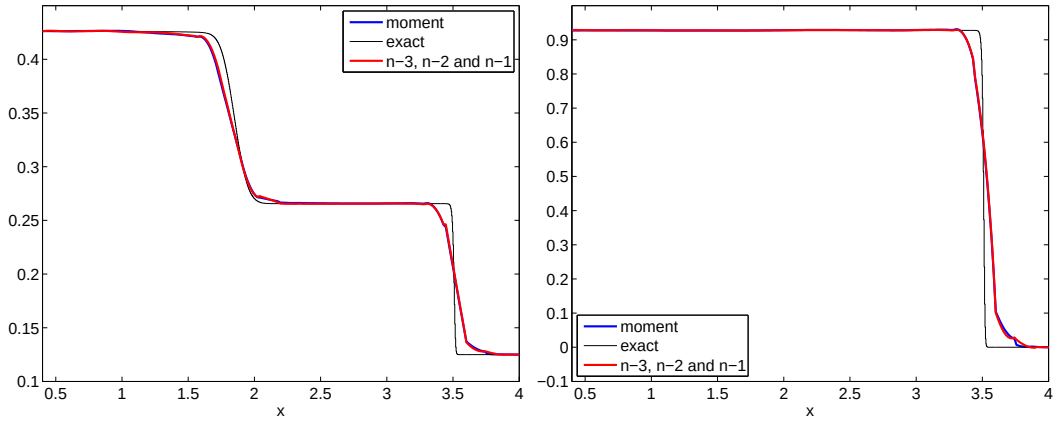


(a) Density ρ (b) Velocity u



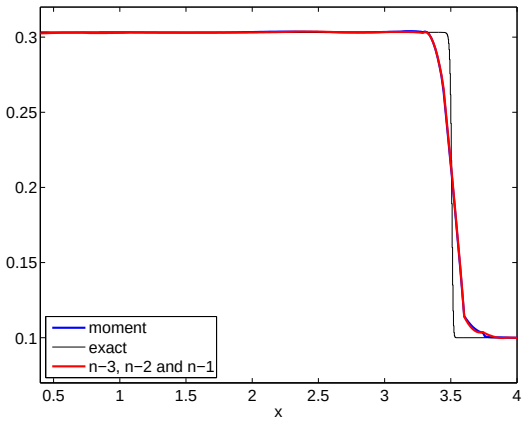
(c) Pressure p (d) Energy E

Figure 46: Results of Euler equations with Sod's initial conditions, absorbing boundary conditions, $T = 2, k = 3$, 64 elements. Moment limiter is applied in elements where $\bar{D}^{n-2}$ or $\bar{D}^{n-1}$ is maximal. Exact solution is obtained using 3000 elements and $k = 0$.

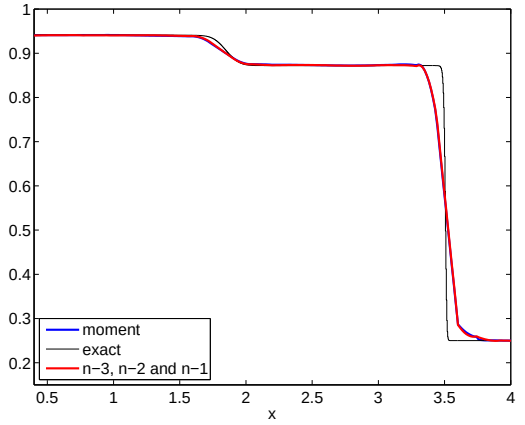


(a) Density ρ

(b) Velocity u



(c) Pressure p

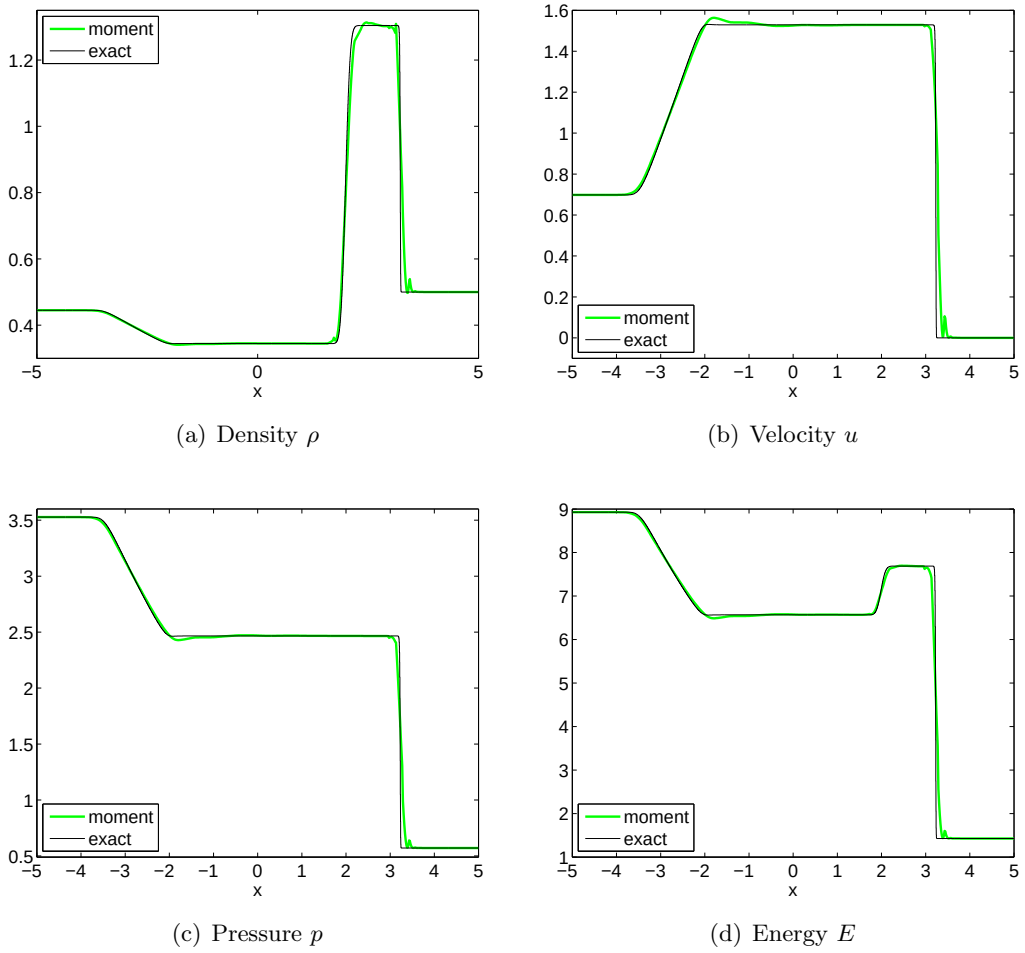


(d) Energy E

Figure 47: Results of Euler equations with Sod's initial conditions, absorbing boundary conditions, $T = 2, k = 3$, 64 elements. Moment limiter is applied in elements where $\bar{D}^{n-3}$, $\bar{D}^{n-2}$ or $\bar{D}^{n-1}$ is maximal. Exact solution is obtained using 3000 elements and $k = 0$.

4.3 Lax

For Lax's shock tube, it is not possible to compute the approximation without the use of a limiter: the oscillations grow very fast with time. The moment limited solution however, is not bad, as can be seen in Figure 48. Note that the moment limiter still introduces some oscillations.



(a) Density ρ

(b) Velocity u

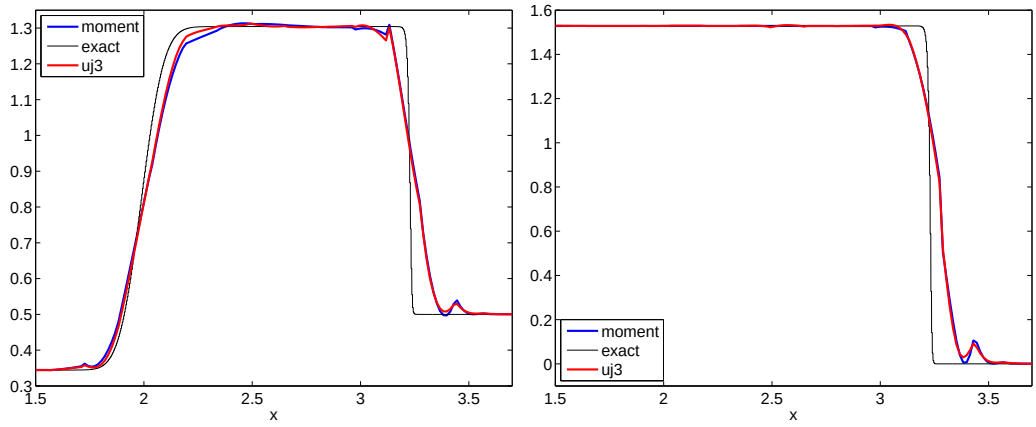
(c) Pressure p

(d) Energy E

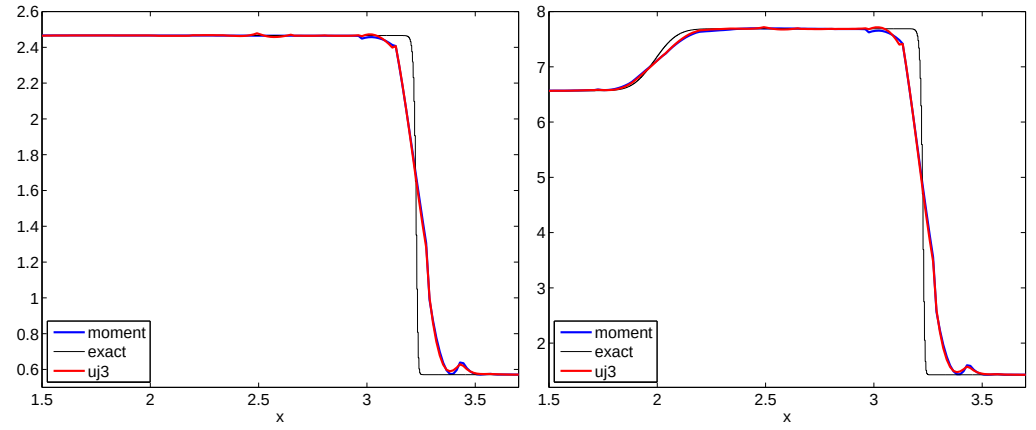
Figure 48: Results of Euler equations with Lax's initial conditions, absorbing boundary conditions, $T = 1.3, k = 3$, 64 elements. Exact solution is obtained using 3000 elements and $k = 0$.

Next, we apply a shock detector, in combination with the moment limiter, to Lax's shock tube. The results for the contact discontinuity and shock are given in Figures 49-53. The shock detector using $\Delta u'_{j+\frac{1}{2}}, j = 0, \ldots, N-1$ (Figure 50), and the multiwavelet shock detector using level $n-1$ (Figure 51), produce oscillations, which we do not want to see.
The contact discontinuity (region around $x = 2$) is best captured by the multiwavelet shock detector on levels $n-2$ and $n-1$ (Figure 52). Also for the shock, this detector gives good results.
The shock detectors based on $|u_j^{(3)}|, j = 0, \ldots, N$ (Figure 49), and levels $n-3, n-2, n-1$ (Figure 53), do not differ so much from the moment limiter itself.

Because the moment limited solution oscillates around the right end of the rarefaction wave, we look at the results in that region as well (Figures 54-58). None of the shock detectors is able to remove these oscillations. The shock detector, based on $\Delta u'_{j+\frac{1}{2}}, j = 0, \ldots, N-1$, is the only one that tries to smooth these oscillations.

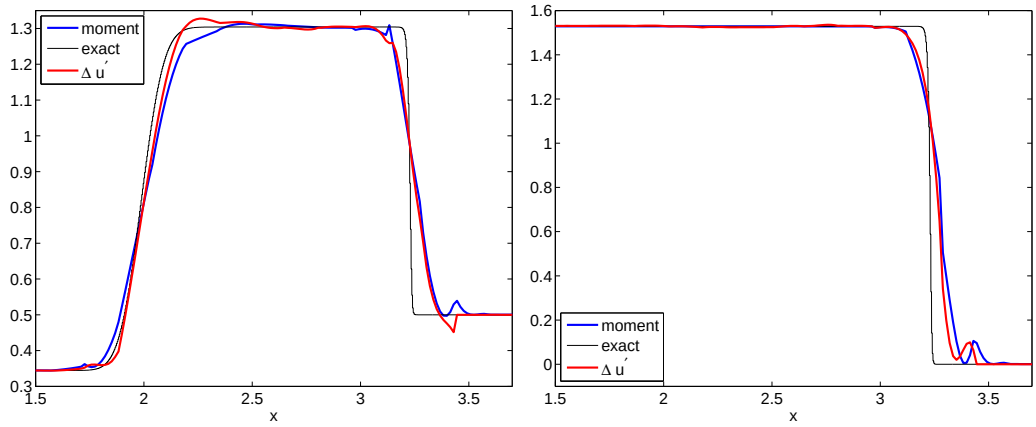


(a) Density ρ (b) Velocity u

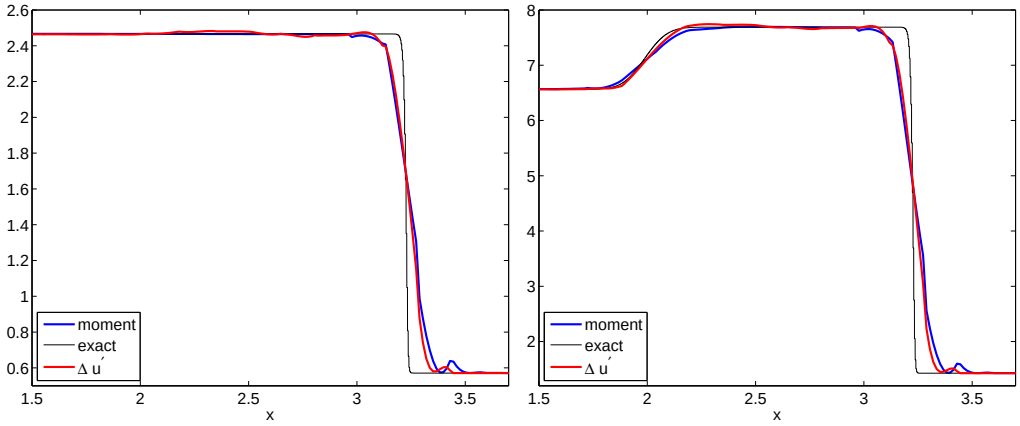


(c) Pressure p (d) Energy E

Figure 49: Results of Euler equations with Lax's initial conditions, absorbing boundary conditions, $T = 2, k = 3$, 64 elements. Moment limiter is applied in elements where $|u_j^{(3)}|$ is maximal, $j = 0, \ldots, N$. Exact solution is obtained using 3000 elements and $k = 0$.

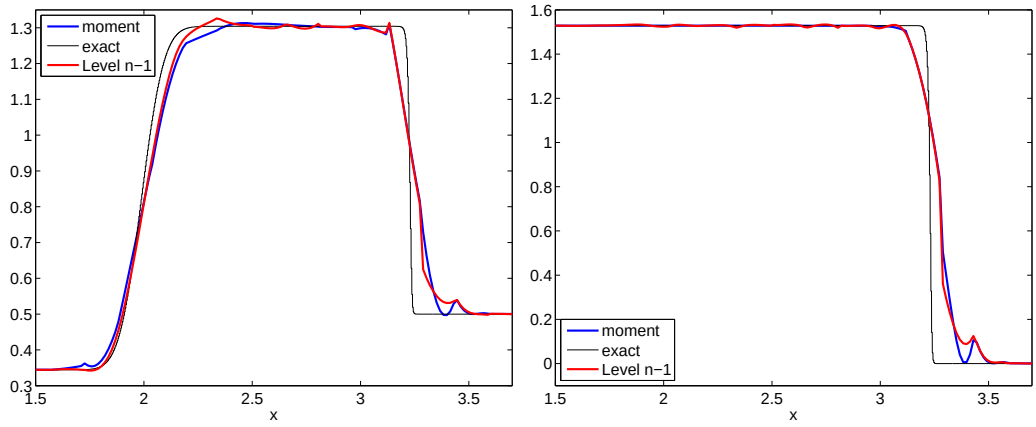


(a) Density ρ (b) Velocity u

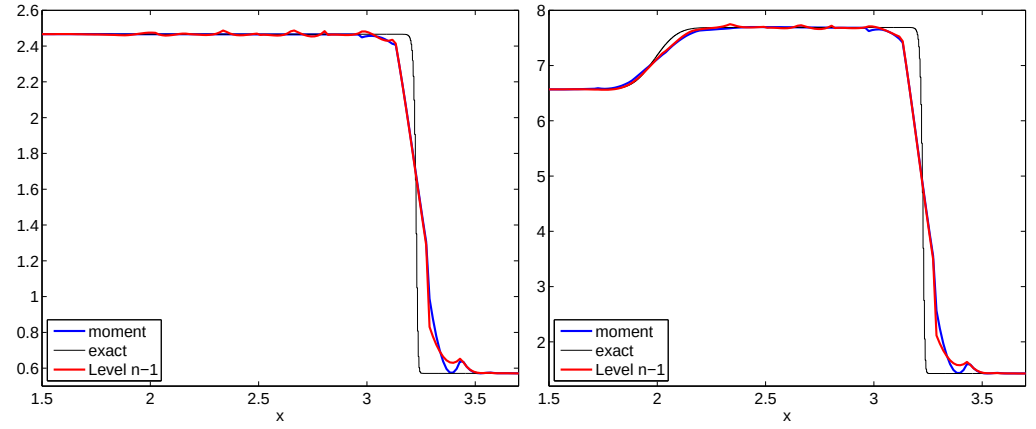


(c) Pressure p (d) Energy E

Figure 50: Results of Euler equations with Lax's initial conditions, absorbing boundary conditions, $T = 2, k = 3$, 64 elements. Moment limiter is applied in elements where $\Delta u'_{j+\frac{1}{2}}$ is maximal, $j = 0, \ldots, N-1$. Exact solution is obtained using 3000 elements and $k = 0$.

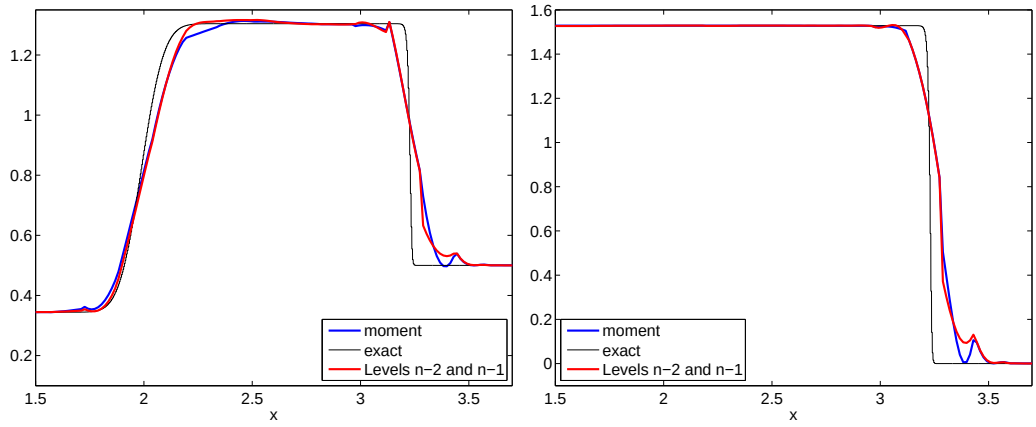


(a) Density ρ (b) Velocity u

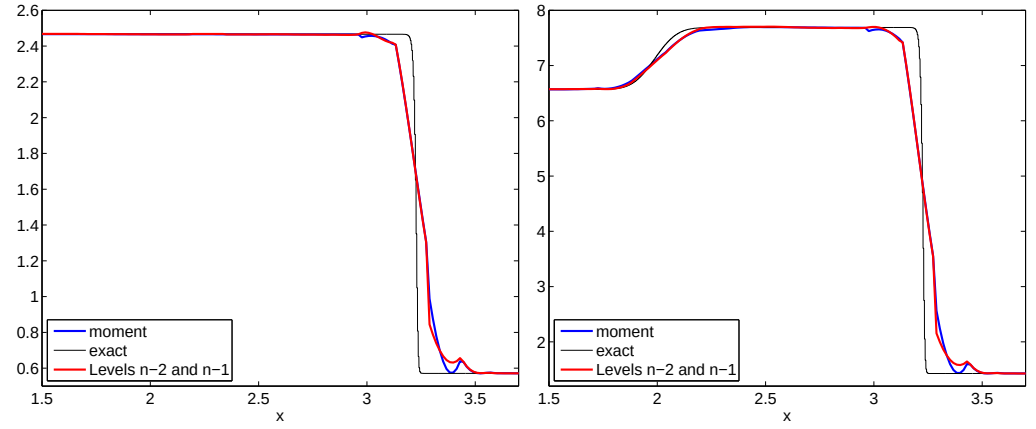


(c) Pressure p (d) Energy E

Figure 51: Results of Euler equations with Lax's initial conditions, absorbing boundary conditions, $T = 2, k = 3$, 64 elements. Moment limiter is applied in elements where $\bar{D}^{n-1}$ is maximal. Exact solution is obtained using 3000 elements and $k = 0$.

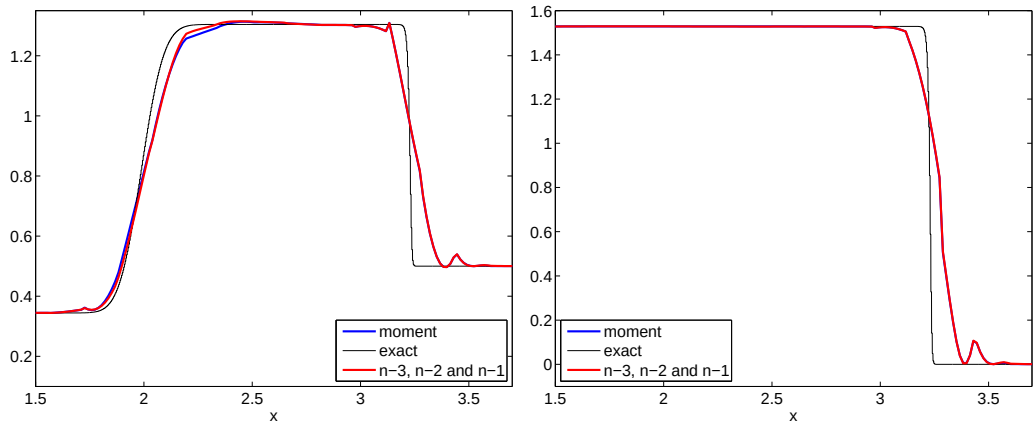


(a) Density ρ (b) Velocity u

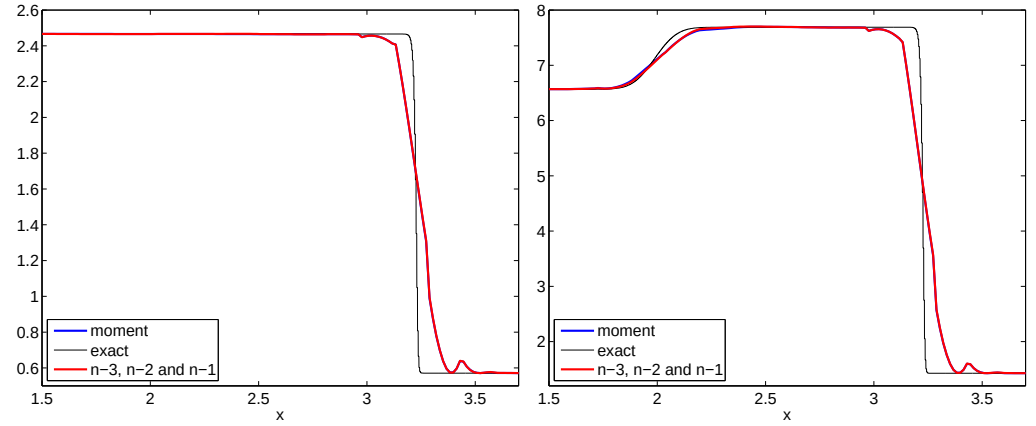


(c) Pressure p (d) Energy E

Figure 52: Results of Euler equations with Lax's initial conditions, absorbing boundary conditions, $T = 2, k = 3$, 64 elements. Moment limiter is applied in elements where $\bar{D}^{n-2}$ or $\bar{D}^{n-1}$ is maximal. Exact solution is obtained using 3000 elements and $k = 0$.

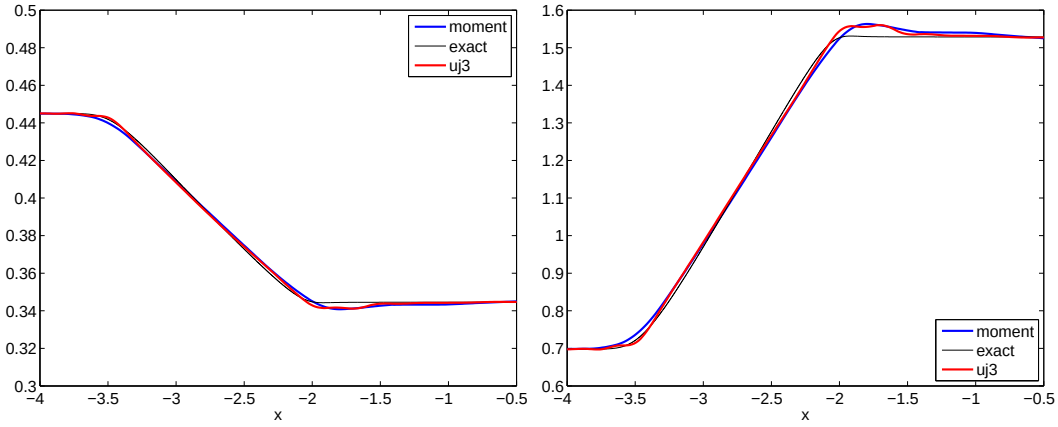


(a) Density ρ (b) Velocity u

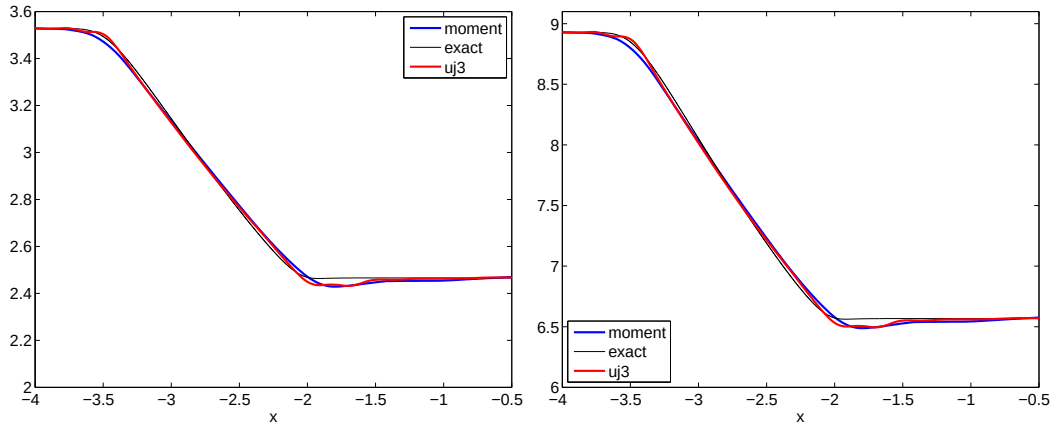


(c) Pressure p (d) Energy E

Figure 53: Results of Euler equations with Lax's initial conditions, absorbing boundary conditions, $T = 2, k = 3$, 64 elements. Moment limiter is applied in elements where $\bar{D}^{n-3}$, $\bar{D}^{n-2}$ or $\bar{D}^{n-1}$ is maximal. Exact solution is obtained using 3000 elements and $k = 0$.

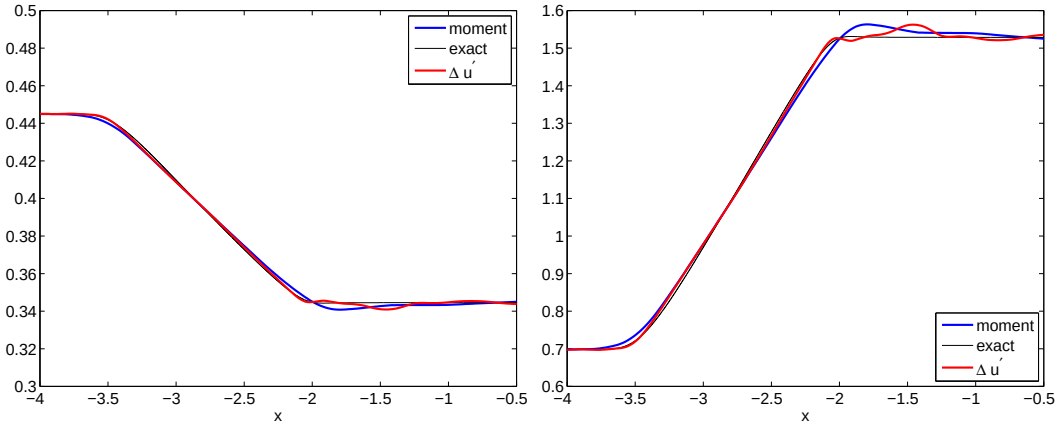


(a) Density ρ (b) Velocity u

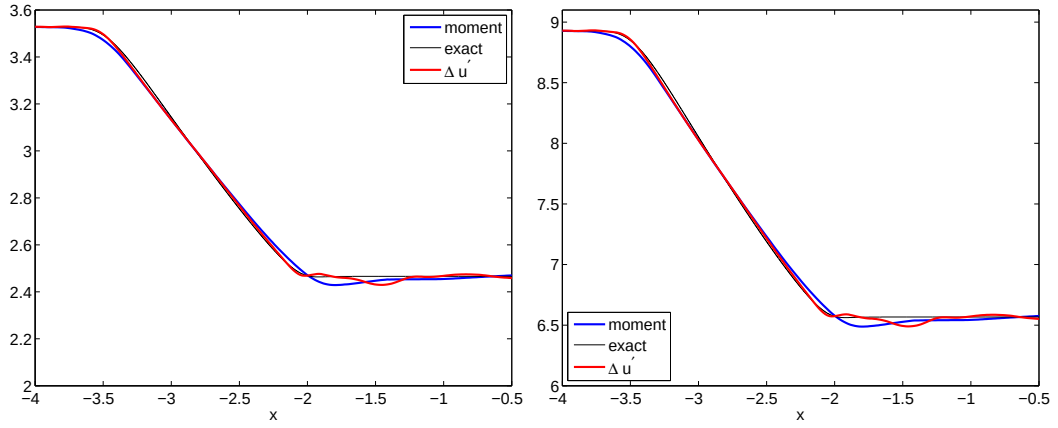


(c) Pressure p (d) Energy E

Figure 54: Results of Euler equations with Lax's initial conditions, absorbing boundary conditions, $T = 2, k = 3$, 64 elements. Moment limiter is applied in elements where $|u_j^{(3)}|$ is maximal, $j = 0, \ldots, N$. Exact solution is obtained using 3000 elements and $k = 0$.

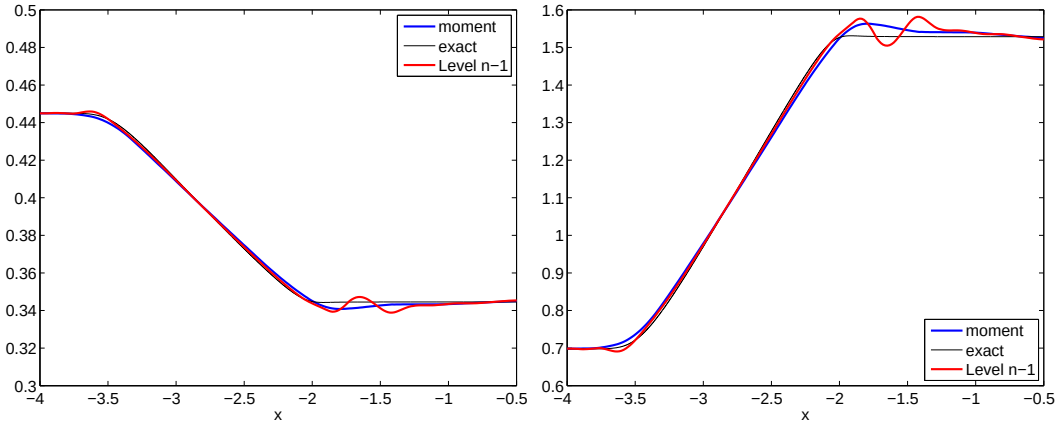


(a) Density ρ (b) Velocity u

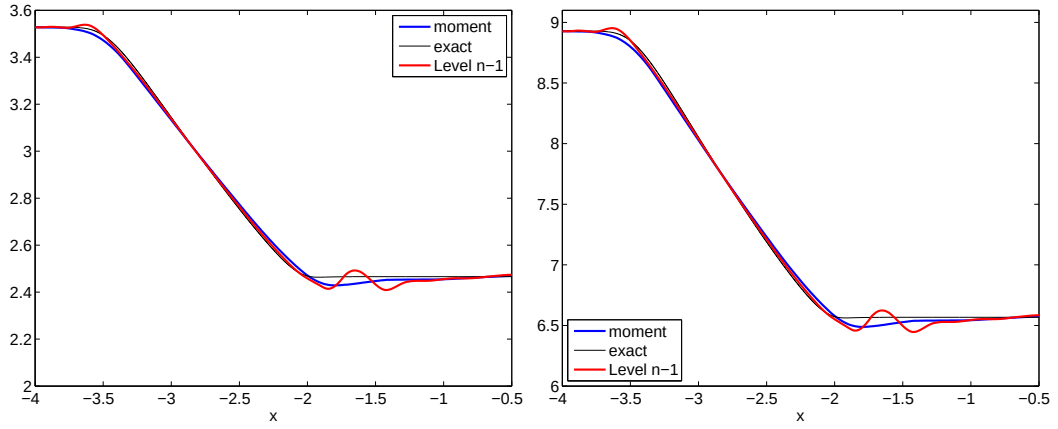


(c) Pressure p (d) Energy E

Figure 55: Results of Euler equations with Lax's initial conditions, absorbing boundary conditions, $T = 2, k = 3$, 64 elements. Moment limiter is applied in elements where $\Delta u'_{j+\frac{1}{2}}$ is maximal, $j = 0, \ldots, N-1$. Exact solution is obtained using 3000 elements and $k = 0$.

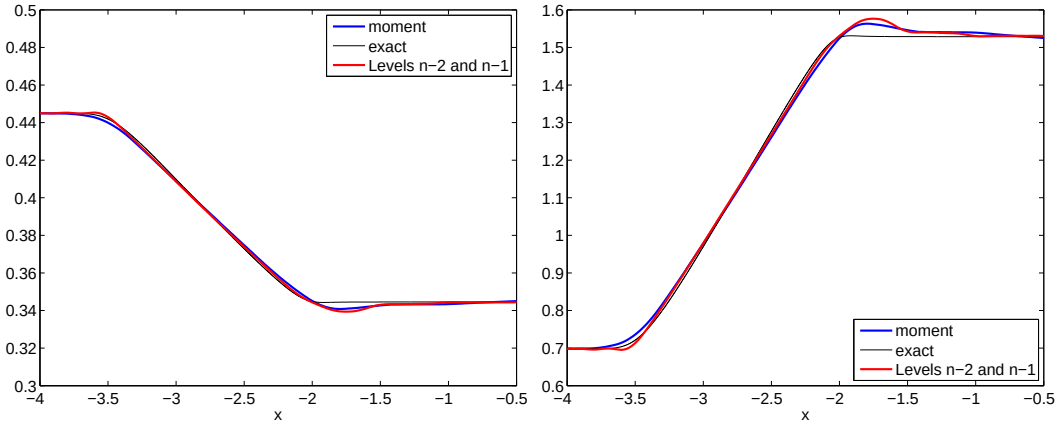


(a) Density ρ (b) Velocity u

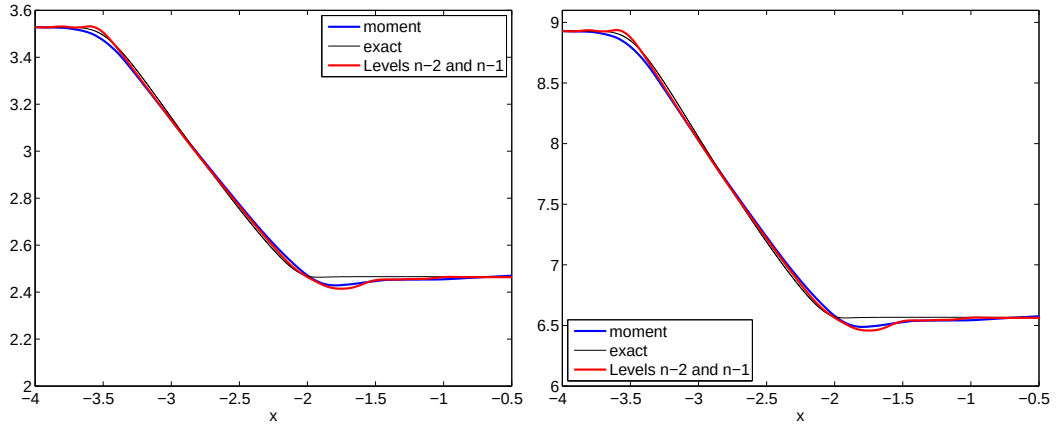


(c) Pressure p (d) Energy E

Figure 56: Results of Euler equations with Lax's initial conditions, absorbing boundary conditions, $T = 2, k = 3$, 64 elements. Moment limiter is applied in elements where $\bar{D}^{n-1}$ is maximal. Exact solution is obtained using 3000 elements and $k = 0$.

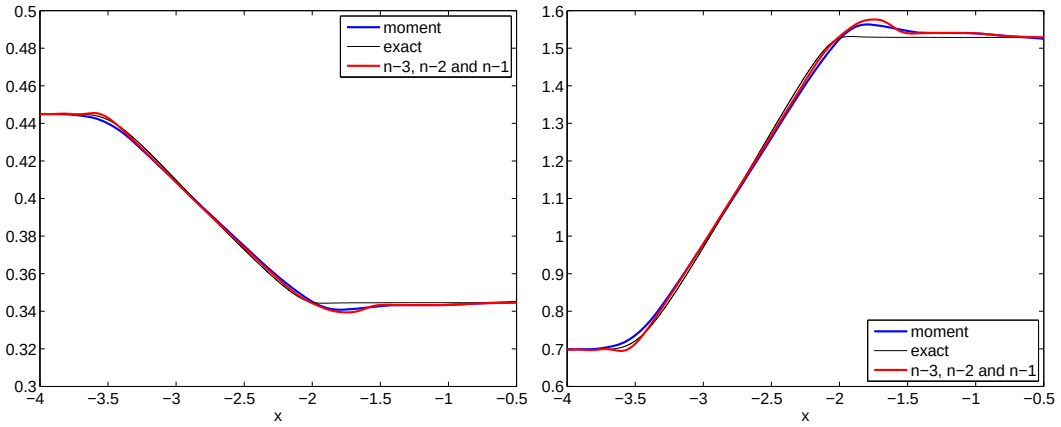


(a) Density ρ (b) Velocity u

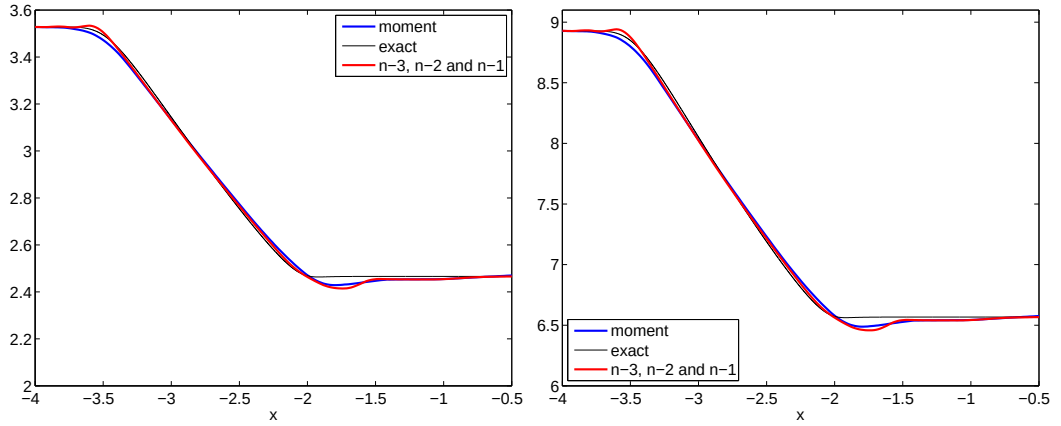


(c) Pressure p (d) Energy E

Figure 57: Results of Euler equations with Lax's initial conditions, absorbing boundary conditions, $T = 2, k = 3$, 64 elements. Moment limiter is applied in elements where $\bar{D}^{n-2}$ or $\bar{D}^{n-1}$ is maximal. Exact solution is obtained using 3000 elements and $k = 0$.



(a) Density ρ (b) Velocity u



(c) Pressure p (d) Energy E

Figure 58: Results of Euler equations with Lax's initial conditions, absorbing boundary conditions, $T = 2, k = 3$, 64 elements. Moment limiter is applied in elements where $\bar{D}^{n-3}$, $\bar{D}^{n-2}$ or $\bar{D}^{n-1}$ is maximal. Exact solution is obtained using 3000 elements and $k = 0$.

Conclusion

This master thesis was devoted to the search for a new limiter for DG, using multiwavelets. During our project, we discovered that multiwavelets can play an important role for shock detection. In this section, we summarize the established work, and draw some conclusions.

The multiwavelet decomposition (Section 1.3) is perfectly able to localize different kinds of discontinuities (shocks, contact discontinuities, rarefaction waves). Unfortunately, it is very hard to use this feature to construct a new multiwavelet limiter. In Section 2.2, we tried to limit the solution by adapting the multiwavelet contribution of the higher levels. This introduced problems with computation time (for each RK step, the limited approximation should be projected onto DG space), and new discontinuities are constructed within the elements.
Secondly, we tried to limit the DG approximation by limiting its multiwavelet coefficients on the higher levels (Section 2.3). This approach effectively reduces the number of elements that is used, such that a very coarse grid is found, where the DG approximation is defined.

Because the multiwavelets work for detecting discontinuities, we switched to shock detection (Section 3). We have seen that both the DG coefficients themselves, and the multiwavelet decomposition, can be used for detecting discontinuities in the solution.
The DG coefficient shock detectors are based on the derivative of the DG approximation inside an element (works only for higher order methods), or the derivative on the element boundaries (Section 3.2).
The multiwavelet shock detectors that we constructed, use the maximum average of the multiwavelet contribution for the highest multiwavelet levels. A drawback of this detector is that the multiwavelet decomposition on each RK step causes an increase of the computation time.

In Section 4, we used the various shock detectors in combination with an existing limiter. We applied the limiter only in the detected shock elements.
For the test problems (linear advection equation in 1D, inviscid Burgers' equation, Section 4.1), this idea works very well, and the DG approximation of the solution improves in the neighbourhood of the discontinuities.
More involved problems were given by Sod's and Lax's shock tubes (Sections 4.2 and 4.3). Also for these problems, it was possible to improve the moment limited solution, and most improvements were found in the direct neighbourhood of the discontinuities.

Summarizing the work established in this master thesis, it turns out that we are able to improve the currently used moment limiter for DG locally, in the discontinuous regions of the solution. We discovered several shock detectors that work excellent in detecting discontinuities.
However, more research must be done on this subject. Suggestions for further research are given in the next section.

Further research

In this section, we make some notes for further research.

First of all, we want to find the reason why multiwavelets can be used for shock detection. Therefore, we need to examine why the multiwavelet contribution of the highest levels grows in the neighbourhood of the discontinuities in the solution. This is a theoretical question, and the answer may be given in the literature.
Further, we ask ourselves if it is possible to recognize which type of discontinuity is found by the multiwavelet decomposition (shock, contact discontinuity, rarefaction wave).

Secondly, we should think about the total variation, and check if the moment limiter, applied in combination with our shock detectors, gives total variation bounded results.

Furthermore, we can focus on the multiwavelet decomposition of the characteristic variables, $\mathbf{w} = R^{-1}\mathbf{u}$, for the Euler equations, instead of looking at the conserved variables, $\mathbf{u}$. Maybe the extrema of these variables contain extra information about the type of discontinuity that is found.

Another interesting idea is that the derivative of the solution is discontinuous in the end points of the rarefaction wave. It is possible that this region is selected if we use a shock detector for $u_h'(x,t)$.

More improvements to the multiwavelet shock detector can be made by modifying the ideas in Section 3.3. In Section 3.3, we used the averages $\bar{D}_j^m$ of $|D_j^m(x)|$ on element $I_j^m, j = 0, \ldots, 2^m - 1$. However, notice that, per definition, multiwavelets are discontinuous in the center of element I_j^m (see Figure 60b in Appendix B). Therefore, it is interesting to distinguish these two parts in each element, and to look at the averages $\bar{D}^m(i)$ of $|D^m(x)|$ on elements $I_i^{m+1}, i = 0, \ldots, 2^{m+1} - 1$.
The current approach of Section 3.3 results in 2^m averages on level m, whereas the new approach defines 2^{m+1} averages on level m. This should work better for shock detection, because we find 2^n averages on level $n-1$ (which is the number of elements of the DG discretization).

In order to answer these questions, we need to look at more sets of initial conditions for the one dimensional Euler equations. For example the blast wave, or the sine-entropy wave, can be used as an initial condition. For these problems, the moment limiter looses accuracy in the neighbourhood of smooth extrema. We expect that these extrema are not detected by our shock detectors, and the approximation may improve if we combine the moment limiter with one of our shock detectors.

Lastly, to truly check the viability, the working of the shock detectors in 2D should be examined.

References

[1] M. Abramowitz and I.A. Stegun, editors. *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*, volume 55 of *Applied Mathematics*. National Bureau of Standards, Washington, D.C., 1964.

[2] B.K. Alpert. A Class of Bases in L^2 for the Sparse Representation of Integral Operators. *SIAM J.on Mathematical Analysis*, 24:246–262, 1993.

[3] B.K. Alpert, G. Beylkin, D. Gines, and L. Vozovoi. Adaptive Solution of Partial Differential Equations in Multiwavelet Bases. *Journal of Computational Physics*, 182:149–190, 2002.

[4] R.K. Archibald, G.I. Fann, and W.A. Shelton. Adaptive Discontinuous Galerkin Methods with Multiwavelets Bases. *Applied Numerical Mathematics*, 61(7):879–890, 2011.

[5] R. Biswas, K.D. Devine, and J.E. Flaherty. Parallel, adaptive finite element methods for conservation laws. *Applied Numerical Mathematics*, 14:255–283, 1994.

[6] R.L. Burden and J.D. Faires. *Numerical Analysis*. Brooks/Cole, Boston, ninth edition, 1985.

[7] C.S. Burrus, R.A. Gopinath, and H. Guo. *Introduction to Wavelets and Wavelet Transforms*. Prentice-Hall, New Jersey, 1998.

[8] B. Cockburn. *Advanced Numerical Approximation of Nonlinear Hyperbolic Equations*, volume 1697 of *Lecture Notes in Mathematics*, chapter An introduction to the Discontinuous Galerkin method for convection-dominated problems. Springer Berlin/Heidelberg, 1998.

[9] B. Cockburn and C.-W. Shu. TVB Runge-Kutta Local Projection Discontinuous Galerkin Finite Element Method for Conservation Laws II: General Framework. *Mathematics of Computation*, 52(186):411–435, 1989.

[10] B. Cockburn and C.-W. Shu. Runge-Kutta Discontinuous Galerkin Methods for Convection-Dominated Problems. *Journal of Scientific Computing*, 16(3):173–261, 2001.

[11] S. Gottlieb and C.-W. Shu. Total Variation Diminishing Runge-Kutta Schemes. *Mathematics of Computation*, 67(221):73–85, 1998.

[12] N. Hovhannisyan, S. Müller, and R. Schäfer. Adaptive Multiresolution Discontinuous Galerkin Schemes for Conservation Laws. Report 311, Institut für Geometrie und Praktische Mathematik, Aachen, sep 2010. http://www.igpm.rwth-aachen.de/en/reports2010.

[13] F. Iacono, G. May, S. Müller, and R. Schäfer. A High-Order Discontinuous Galerkin Discretization with Multiwavelet-Based Grid Adaptation for Compressible Flows. Technical Report AICES-2011/08-02, Aachen Institute for Advanced Study in Computational Engineering Science, Aachen, aug 2011. http://www.aices.rwth-aachen.de/preprints.

[14] F. Keinert. *Wavelets and Multiwavelets*. Studies in Advanced Mathematics. Chapman and Hall/CRC, Florida, 2004.

[15] L. Krivodonova. Limiters for high-order discontinuous Galerkin methods. *Journal of Computational Physics*, 226:879–896, 2007.

[16] P.D. Lax. Weak Solutions of Nonlinear Hyperbolic Equations and Their Numerical Computation. *Communications on Pure and Applied Mathematics*, 7:159–193, 1954.

[17] R.J. LeVeque. *Numerical Methods for Conservation Laws*. Lectures in Mathematics. Birkhäuser Verlag, Basel, second edition, 1992.

[18] R.J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, New York, sixth edition, 2002.

[19] J.X. Qiu and C.-W. Shu. Runge-Kutta Discontinuous Galerkin Method using WENO limiters. *Society for Industrial and Applied Mathematics*, 26(3):907–929, 2005.

[20] D.K. Ruch and P.J. van Fleet. *Wavelet theory: An Elementary Approach With Applications*. John Wiley & Sons, Inc., New Jersey, 2009.

[21] J. Smoller. *Shock Waves and Reaction-Diffusion Equations*, volume 258 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, New York, second edition, 1983.

[22] G.A. Sod. A Survey of Several Finite Difference Methods for Systems of Nonlinear Hyperbolic Conservation Laws. *Journal of Computational Physics*, 27:1–31, 1978.

[23] W.G. Strang. *Linear algebra and its applications*. Harcourt Brace Jovanovich, San Diego, third edition, 1976.

[24] J.W. Thomas. *Numerical Partial Differential Equations: Conservation Laws and Elliptic Equations*, volume 33 of *Texts in Applied Mathematics*. Springer-Verlag, New York, first edition, 1999.

[25] X. Zhang and C.-W. Shu. On maximum-principle-satisfying high order schemes for scalar conservation laws. *Journal of Computational Physics*, 229(9):3091–3120, 2010.

A Numerical details for DG with TVD RK

In this appendix, several test problems for DG are considered. We compute the weak formulation, and determine the CFL number that is used for computation. Furthermore, we look at the order of the error after the application of the TVD Runge Kutta time stepping.

A.1 The linear advection equation in 1D

In this appendix, the linear advection equation on $[-1,1]$ is considered, given by

$$u_t + u_x = 0, \quad x \in [-1,1], t \geq 0, \tag{60a}$$

$$u(x,0) = u^0(x), \quad x \in [-1,1], \tag{60b}$$

with periodic boundary conditions.
Note that the exact solution of equations (60) is known: the initial condition given in (60b) is advected with constant speed along the characteristics $x - t = c$, with $c \in \mathbb{R}$, see LeVeque, [18].
For the numerical solution, it should be noticed that in this case, equation (9) equals,

$$\frac{\Delta x}{2}\frac{du_j^{(m)}}{dt} = \int_{-1}^{1} \phi_\ell(\xi)\phi_m'(\xi)d\xi + \hat{F}_{j-\frac{1}{2}} v^+_{h,j-\frac{1}{2}} - \hat{F}_{j+\frac{1}{2}} v^-_{h,j+\frac{1}{2}}. \tag{61}$$

Hereby, the integral in equation (61) takes the values,

$$\int_{-1}^{1} \phi_\ell(\xi)\frac{d}{d\xi}\phi_m(\xi)d\xi = \begin{cases} 2\sqrt{(\ell+\frac{1}{2})(m+\frac{1}{2})}, & \text{if } m > \ell \text{ and } m+\ell \text{ is odd;} \\ 0, & \text{else,} \end{cases} \tag{62}$$

and the local Lax Friedrichs flux through the boundaries $x_{j\pm\frac{1}{2}}$ (equations (11)) is given by,

$$\begin{aligned} \hat{F}_{j-\frac{1}{2}} &= \frac{1}{2}\left(u^-_{j-\frac{1}{2}} + u^+_{j-\frac{1}{2}} - (u^+_{j-\frac{1}{2}} - u^-_{j-\frac{1}{2}})\right) \\ &= u^-_{j-\frac{1}{2}} = \sum_{\ell=0}^{k} u^{(\ell)}_{j-1}\phi_\ell(1), \end{aligned} \tag{63a}$$

$$\hat{F}_{j+\frac{1}{2}} = u^-_{j+\frac{1}{2}} = \sum_{\ell=0}^{k} u^{(\ell)}_{j}\phi_\ell(1). \tag{63b}$$

Using equations (10) and (63), equation (61) transforms to the following differential equation:

$$\begin{aligned} \tfrac{\Delta x}{2}\tfrac{du_j^{(m)}}{dt} &= \sum_{\ell=0}^{k} u_j^{(\ell)} \textstyle\int_{-1}^{1} \phi_\ell(\xi)\phi_m'(\xi)d\xi + \\ &- \sqrt{m+\tfrac{1}{2}}\left(\sum_{\ell=0}^{k}\sqrt{\ell+\tfrac{1}{2}}u_j^{(\ell)}\right) + (-1)^m\sqrt{m+\tfrac{1}{2}}\left(\sum_{\ell=0}^{k}\sqrt{\ell+\tfrac{1}{2}}u_{j-1}^{(\ell)}\right); \end{aligned}$$

where $u_j^{(m)}(t), j = 0, \ldots, N, m = 0, \ldots, k$ are the unknown DG coefficients. Written in matrix-vector form, this looks like,

$$M\frac{d}{dt}\mathbf{u}_j = S_1\mathbf{u}_j + S_2\mathbf{u}_{j-1}, j = 1, \ldots, N, \tag{64}$$

where,

$$\begin{aligned}
\mathbf{u}_j &= \left(u_j^{(0)}(t) \;\; u_j^{(1)}(t) \;\; \ldots \;\; u_j^{(k)}(t)\right)^\top, \\
M &= \operatorname{diag}\left(\tfrac{\Delta x}{2}\right), \\
S_1(m+1,\ell+1) &= -\sqrt{(m+\tfrac{1}{2})(\ell+\tfrac{1}{2})} + \textstyle\int_{-1}^{1} \phi_\ell(\xi)\frac{d}{d\xi}\phi_m(\xi)d\xi, \\
S_2(m+1,\ell+1) &= (-1)^m\sqrt{(m+\tfrac{1}{2})(\ell+\tfrac{1}{2})}.
\end{aligned}$$

The three matrices all have the same size of $(k+1)\times(k+1)$.
For element I_0, the periodic boundary condition is used, which results in the following equation,

$$M\frac{d}{dt}\mathbf{u}_0 = S_1\mathbf{u}_0 + S_2\mathbf{u}_N. \tag{65}$$

The initial DG coefficients $u_j^{(\ell)}(0), j = 0, \ldots, N, \ell = 0, \ldots, k$, are found using a projection onto $V_h(I_j)$. Let this projected initial condition $u_h(x,0) \in V_h(I_j)$ be given by,

$$u_h(x,0) = \sum_{\ell=0}^{k} u_j^{(\ell)}(0)\phi_\ell\left(\frac{2}{\Delta x}(x-x_j)\right), x \in I_j, j = 0, \ldots, N, \tag{66}$$

see equation (7b). Multiplying this equation by $\phi_m\left(\frac{2}{\Delta x}(x-x_j)\right)$, and integrating over I_j, gives,

$$\int_{I_j} u_h(x,0)\phi_m\left(\frac{2}{\Delta x}(x-x_j)\right)dx = \int_{I_j}\sum_{\ell=0}^{k} u_j^{(\ell)}(0)\phi_\ell\left(\frac{2}{\Delta x}(x-x_j)\right)\phi_m\left(\frac{2}{\Delta x}(x-x_j)\right)dx.$$

Using the change of coordinates $\xi = \frac{2}{\Delta x}(x-x_j)$ and the orthonormal property of the scaled Legendre polynomials, as defined in (4), gives,

$$\frac{\Delta x}{2}\int_{-1}^{1} u_h\left(x_j+\frac{\Delta x}{2}\xi, 0\right)\phi_m(\xi)d\xi = \frac{\Delta x}{2}\sum_{\ell=0}^{k} u_j^{(\ell)}(0)\int_{-1}^{1}\phi_\ell(\xi)\phi_m(\xi)d\xi = \frac{\Delta x}{2}u_j^{(m)}(0).$$

Therefore, the initial DG coefficients are given by,

$$\begin{aligned}
u_j^{(m)}(0) &= \int_{-1}^{1} u_h\left(x_j+\frac{\Delta x}{2}\xi, 0\right)\phi_m(\xi)d\xi \\
&\approx \int_{-1}^{1} u^0\left(x_j+\frac{\Delta x}{2}\xi\right)\phi_m(\xi)d\xi.
\end{aligned} \tag{67}$$

To compute the above stated integral numerically, a Gauss-Legendre quadrature using six points[2] is used, see Abramowitz and Stegun, [1]. This rule is stated as,

$$\int_{-1}^{1} f(x)dx \approx \sum_{i=0}^{5} w_i f(\hat{x}_i),$$

where $\hat{x}_i, w_i, i = 0, \ldots, 5$ are given in Table 5.

i	0	1	2	3	4	5
$\hat{x}_i$	-0.93247	-0.66121	-0.23862	0.23862	0.66121	0.93247
w_i	0.17132	0.36076	0.46791	0.46791	0.36076	0.17132

Table 5: Points and weights using the six point Gauss-Legendre quadrature

For the time stepping using the third order TVD RK, the value for Δt should be chosen such that the CFL number $\nu = \frac{\Delta t}{\Delta x}$ is small enough (LeVeque, [18]). In that case, the time stepping errors can be neglected with respect to the spatial errors. For the linear advection equation with approximation space Φ^k, the following stability condition holds:

$$\nu \leq \frac{1}{2k+1}, \ k = 0, 1, 2,$$

Cockburn, [8].
In practice, this bound is somewhat too restrictive. For $k \leq 2$, the used stability limit is given in Table 6 (Cockburn [10]). Note that in this case, the time stepping error is $\mathcal{O}(\Delta t^3)$ whereas the dominating spatial errors are of order $k + 1 \leq 3$.

Approximation space	Φ^0	Φ^1	Φ^2
$\nu_{\max}$	1.256	0.409	0.209

Table 6: Maximum CFL number for stability using DG and TVD RK to solve $u_t + u_x = 0$

If $k \geq 3$, then it is possible that the time stepping errors dominate. In general, for $k = 3$ it should hold that $\nu \leq 0.13$, but if another number of elements is used, then the new CFL number $\widetilde{\nu}$ should be adapted such that the spatial errors still dominate. This is achieved by using,

$$\widetilde{\nu} = \nu \left(\frac{N+1}{\widetilde{N}+1} \right)^{\frac{k+1}{\mu}}, \tag{68}$$

where ν is the old stability bound, $N + 1$ is the number of elements in the old mesh, $\widetilde{N} + 1$ is the number of elements on the current mesh, $k + 1$ is the order of the DG scheme, and μ is the order of the time stepping scheme (in this case $\mu = 3$), see Table 7. The most restrictive value for ν is chosen, such that $\frac{\Delta t}{\Delta x} \leq 0.003 \leq 0.13 \cdot \left(\frac{1}{2}\right)^{\frac{16}{3}} \approx 0.0032$. Note that the chosen value for ν depends on the initial mesh. Relation (68) is an empirical formula which can not be proven thoroughly.

[2]The associated formula is given at `http://pathfinder.scar.utoronto.ca/~dyer/csca57/book_P/node44.html`

$N+1$	10	20	40	80	160
ν	0.13	$0.13 \cdot \left(\frac{10}{20}\right)^{\frac{4}{3}}$	$0.13 \cdot \left(\frac{1}{2}\right)^{\frac{8}{3}}$	$0.13 \cdot \left(\frac{1}{2}\right)^{\frac{12}{3}}$	$0.13 \cdot \left(\frac{1}{2}\right)^{\frac{16}{3}}$

Table 7: CFL number such that spatial errors dominate, using approximation space Φ^3 to solve $u_t + u_x = 0$

It should be noticed that the exact solution of the linear advection equation is known, which is given by $u(x,t) = u^0(x-t)$. Therefore, the accuracy of the numerical results can be computed using the pointwise error norm, which is defined as,

$$||e(T)||_\infty = \max\left\{|u(x,T) - u_h(x,T)|, x \in [-1,1]\right\}.$$

Following Theorem 2.2 in Cockburn, [8], a regular initial condition (e.g. $u^0(x) = \sin(2\pi x)$) using approximation space Φ^k, should give an approximation error of order $(\Delta x)^{k+1}$, if the correct CFL number is chosen. To verify this relation, two different values of N are used. Note that if,

$$\begin{aligned} ||e(T)|| &= C(\Delta x)^{k+1} = C\left(\tfrac{2}{N+1}\right)^{k+1}, \\ ||\widetilde{e}(T)|| &= C(\widetilde{\Delta x})^{k+1} = C\left(\tfrac{2}{\widetilde{N}+1}\right)^{k+1}, \end{aligned}$$

with $N, \widetilde{N} \in \mathbb{N}, C \in \mathbb{R}$, then it should hold that,

$$\frac{||e(T)||}{||\widetilde{e}(T)||} = \left(\frac{\widetilde{N}+1}{N+1}\right)^{k+1}.$$

Taking the logarithm of both sides gives,

$$\log\left(\frac{||e(T)||}{||\widetilde{e}(T)||}\right) = (k+1)\log\left(\frac{\widetilde{N}+1}{N+1}\right),$$

such that the approximation is good enough if,

$$\text{order} = \frac{\log\left(\frac{||e(T)||}{||\widetilde{e}(T)||}\right)}{\log\left(\frac{\widetilde{N}+1}{N+1}\right)} \geq k+1. \tag{69}$$

In Table 8, the errors $||e(0.5)||_\infty$ are given, using initial condition $u^0(x) = \sin(2\pi x)$. The order of accuracy can also be seen: column j gives the order of the approximation using the mesh of column $j-1$ and j. Indeed, the order equals $k+1$.
Due to the limitations of the processing chip, using polynomial basis Φ^4 or higher causes a dominating roundoff error. The order of the error is not equal to $k+1$ in that case.

In the case that the CFL number is not adapted for $k = 3$ for the mesh, the order of the error is not equal to $k+1$, but equals approximately 3.5: the error is dominated by time stepping, and not by the spatial error as wanted.

	$\Phi^0, \nu = 1$				
$N+1$	10	20	40	80	160
$\|\|e(0.5)\|\|_\infty$	1.0107	0.7040	0.4221	0.2326	0.1223
order	-	0.5217	0.7380	0.8596	0.9271
	$\Phi^1, \nu = 0.4$				
$N+1$	10	20	40	80	160
$\|\|e(0.5)\|\|_\infty$	0.1659	0.0435	0.0130	0.0035	8.9903e-4
order	-	1.9329	1.7420	1.8997	1.9532
	$\Phi^2, \nu = 0.2$				
$N+1$	10	20	40	80	160
$\|\|e(0.5)\|\|_\infty$	0.0244	0.0030	3.9854e-4	5.0421e-5	6.3179e-6
order	-	3.0146	2.9236	2.9826	2.9965
	$\Phi^3, \nu = 0.003$				
$N+1$	10	20	40	80	160
$\|\|e(0.5)\|\|_\infty$	0.0017	1.1501e-4	7.1912e-6	4.5587e-7	2.8542e-8
order	-	3.9017	3.9993	3.9795	3.9975

Table 8: Errors and orders of the DG approximation at $T = 0.5$, linear advection equation with initial condition $u^0(x) = \sin(2\pi x)$

A.2 The linear advection equation in 2D

In two dimensions, the linear advection equation for $u = u(x, y, t)$ is given by,

$$\begin{array}{ll} u_t + u_x + u_y = 0, & x, y \in [-1, 1], t \geq 0, \\ u(x, y, 0) = u^0(x, y), & x, y \in [-1, 1], \end{array}$$

using periodic boundary conditions.
The space discretization is done by defining N_x and N_y such that $N_x + 1$ elements are found in the $x-$direction and N_y+1 elements in the $y-$direction (N_x does not necessarily equal N_y). The elements are given by, $I_{ij} = (x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}) \times (y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}), i = 0, \ldots, N_x$, $j = 0, \ldots, N_y$, following the definition in one dimension, see Section 1.1.
The approximation space is defined as,

$$V_h(I_{ij}) = \{v : v \in \Phi^k(I_{ij})\}, i = 0, \ldots, N_x, j = 0, \ldots, N_y\},$$

where the basis for Φ^k consists of a multiplication of the scaled Legendre polynomials $\phi_\ell, \ell = 0, \ldots, k$ (equations (3)).
Using local coordinates $\xi = \frac{2}{\Delta x}(x - x_i)$ and $\eta = \frac{2}{\Delta y}(y - y_j)$, the DG approximation is given by,

$$u_h(x, y, t) = \sum_{\ell_x=0}^{k} \sum_{\ell_y=0}^{k} u_{ij}^{(\ell_x,\ell_y)}(t)\phi_{\ell_x}(\xi)\phi_{\ell_y}(\eta), \tag{70}$$

which is a direct extension of Section 1.1. The weak form of the differential equation can be derived by multiplying by an arbitrary function, $v \in C^1(I_{ij})$.
Integrating over element I_{ij} gives, using partial integration and the same notation as in Section 1.1:

$$\begin{array}{l} \iint_{I_{ij}} (u_t + u_x + u_y) v dx dy = \\ \iint_{I_{ij}} u_t v dx dy + \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \left(vu\big|_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} - \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} uv_x dx \right) dy \\ \quad + \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \left(vu\big|_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} - \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} uv_y dy \right) dx \\ \quad\quad \approx \iint_{I_{ij}} u_t v dx dy - \iint_{I_{ij}} u(v_x + v_y) dx dy \\ \quad\quad\quad + \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} (v_{i,j+\frac{1}{2}}^- u_{i,j+\frac{1}{2}}^- - v_{i,j-\frac{1}{2}}^+ u_{i,j-\frac{1}{2}}^-) dx \\ \quad\quad\quad + \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} (v_{i+\frac{1}{2},j}^- u_{i+\frac{1}{2},j}^- - v_{i-\frac{1}{2},j}^+ u_{i-\frac{1}{2},j}^-) dy = 0. \end{array} \tag{71}$$

Hereby, the local Lax Friedrichs flux is used to compute the fluxes through the boundaries.
Using the DG approximation, u_h, a test function, $v_h(x, y) = \phi_{m_x}(\xi)\phi_{m_y}(\eta)$ (where $m_x, m_y \in \{0, \ldots, k\}$), and the fact that (from equations (10)),

$$\begin{aligned} v_{i+\frac{1}{2},j}^- u_{i+\frac{1}{2},j}^- &= \sqrt{m_x + \tfrac{1}{2}} \sum_{\ell_x=0}^{k} \sum_{\ell_y=0}^{k} \sqrt{\ell_x + \tfrac{1}{2}} u_{ij}^{(\ell_x,\ell_y)}(t)\phi_{m_y}(\eta)\phi_{\ell_y}(\eta), \\ v_{i-\frac{1}{2},j}^+ u_{i-\frac{1}{2},j}^- &= (-1)^{m_x} \sqrt{m_x + \tfrac{1}{2}} \sum_{\ell_x=0}^{k} \sum_{\ell_y=0}^{k} \sqrt{\ell_x + \tfrac{1}{2}} u_{i-1,j}^{(\ell_x,\ell_y)}(t)\phi_{m_y}(\eta)\phi_{\ell_y}(\eta), \end{aligned}$$

for $y \in [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$ (the same structure is used for $v^-_{i,j+\frac{1}{2}} u^-_{i,j+\frac{1}{2}}$ and $v^+_{i,j-\frac{1}{2}} u^-_{i,j-\frac{1}{2}}$), differential equation (71) is rewritten as,

$$\begin{aligned}
&\sum_{\ell_x=0}^{k}\sum_{\ell_y=0}^{k} \frac{d}{dt} u_{ij}^{(\ell_x,\ell_y)}(t) \iint\limits_{I_{ij}} \left(\phi_{\ell_x}(\xi)\phi_{\ell_y}(\eta)\right)\left(\phi_{m_x}(\xi)\phi_{m_y}(\eta)\right) dxdy = \\
&\sum_{\ell_x=0}^{k}\sum_{\ell_y=0}^{k} u_{ij}^{(\ell_x,\ell_y)}(t) \iint\limits_{I_{ij}} \phi_{\ell_x}(\xi)\phi_{\ell_y}(\eta)\left(\phi_{m_y}(\eta)\frac{d}{dx}\phi_{m_x}(\xi) + \phi_{m_x}(\xi)\frac{d}{dy}\phi_{m_y}(\eta)\right) dxdy \\
&\quad - \sqrt{m_x+\frac{1}{2}} \sum_{\ell_x=0}^{k}\sum_{\ell_y=0}^{k} \sqrt{\ell_x+\frac{1}{2}} u_{ij}^{(\ell_x,\ell_y)}(t) \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \phi_{\ell_y}(\eta)\phi_{m_y}(\eta)dy \\
&\quad + (-1)^{m_x}\sqrt{m_x+\frac{1}{2}} \sum_{\ell_x=0}^{k}\sum_{\ell_y=0}^{k} \sqrt{\ell_x+\frac{1}{2}} u_{i-1,j}^{(\ell_x,\ell_y)}(t) \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \phi_{\ell_y}(\eta)\phi_{m_y}(\eta)dy \\
&\quad - \sqrt{m_y+\frac{1}{2}} \sum_{\ell_x=0}^{k}\sum_{\ell_y=0}^{k} \sqrt{\ell_y+\frac{1}{2}} u_{ij}^{(\ell_x,\ell_y)}(t) \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \phi_{\ell_x}(\xi)\phi_{m_x}(\xi)dx \\
&\quad + (-1)^{m_y}\sqrt{m_y+\frac{1}{2}} \sum_{\ell_x=0}^{k}\sum_{\ell_y=0}^{k} \sqrt{\ell_y+\frac{1}{2}} u_{i,j-1}^{(\ell_x,\ell_y)}(t) \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \phi_{\ell_x}(\xi)\phi_{m_x}(\xi)dx. \qquad (72)
\end{aligned}$$

Using the coordinate transformation to ξ and η and equation (4), equation (72) becomes,

$$\begin{aligned}
\frac{\Delta x}{2}\cdot\frac{\Delta y}{2}\cdot\frac{d}{dt}u_{ij}^{(m_x,m_y)}(t) &= \frac{\Delta y}{2}\sum_{\ell_x=0}^{k} u_{ij}^{(\ell_x,m_y)}(t)\int_{-1}^{1}\phi_{\ell_x}(\xi)\phi'_{m_x}(\xi)d\xi \\
&+ \frac{\Delta x}{2}\sum_{\ell_y=0}^{k} u_{ij}^{(m_x,\ell_y)}(t)\int_{-1}^{1}\phi_{\ell_y}(\eta)\phi'_{m_y}(\eta)d\eta \\
&+ \sqrt{m_x+\frac{1}{2}}\frac{\Delta y}{2}\sum_{\ell_x=0}^{k}\sqrt{\ell_x+\frac{1}{2}}\left(-u_{ij}^{(\ell_x,m_y)}(t) + (-1)^{m_x}u_{i-1,j}^{(\ell_x,m_y)}(t)\right) \\
&+ \sqrt{m_y+\frac{1}{2}}\frac{\Delta x}{2}\sum_{\ell_y=0}^{k}\sqrt{\ell_y+\frac{1}{2}}\left(-u_{ij}^{(m_x,\ell_y)}(t) + (-1)^{m_y}u_{i,j-1}^{(m_x,\ell_y)}(t)\right). \qquad (73)
\end{aligned}$$

The time stepping is again done using the TVD RK method of order three. Therefore, for every element I_{ij}, the following $(k+1)\times(k+1)$ matrices are defined:

$$\mathbf{u}_{ij}(\ell_x{+}1,\ell_y{+}1) = u_{ij}^{(\ell_x,\ell_y)},\ \mathbf{u}_{i-1,j}(\ell_x{+}1,\ell_y{+}1) = u_{i-1,j}^{(\ell_x,\ell_y)},\ \mathbf{u}_{i,j-1}(\ell_x{+}1,\ell_y{+}1) = u_{i,j-1}^{(\ell_x,\ell_y)},$$

such that

$$\frac{d}{dt}\mathbf{u}_{ij} = L(\mathbf{u}_{ij},\mathbf{u}_{i-1,j},\mathbf{u}_{i,j-1}), i = 0,\ldots,N_x; j = 0,\ldots,N_y.$$

The CFL condition in the two dimensional case is given by

$$\nu_{2D} = \frac{\Delta t}{\Delta x} + \frac{\Delta t}{\Delta y} \leq 1.$$

If $\Delta x = \Delta y$, this means that $\nu_{2D} = 2 \cdot \frac{\Delta t}{\Delta x} = 2\nu_{1D}$. This means that for every choice of N_x the value of Δt_{2D} should be one half times the value of Δt_{1D}.
The initial DG coefficients are found using the same procedure as in Appendix A.1: for each element I_{ij}, the initial condition,

$$u_h(x,y,0) = \sum_{\ell_x=0}^{k} \sum_{\ell_y=0}^{k} u_{ij}^{(\ell_x,\ell_y)}(0) \phi_{\ell_x}\left(\frac{2}{\Delta x}(x-x_i)\right) \phi_{\ell_y}\left(\frac{2}{\Delta y}(y-y_j)\right),$$

is multiplied by $\phi_{m_x}\left(\frac{2}{\Delta x}(x-x_i)\right) \phi_{m_y}\left(\frac{2}{\Delta y}(y-y_j)\right), m_x, m_y \in \{0,\ldots,k\}$, and integrated over I_{ij}. Using coordinate transformations to ξ and η as well as the orthonormal property of the scaled Legendre polynomials, the coefficients are then given by,

$$u_{ij}^{(m_x,m_y)}(0) = \int_{-1}^{1}\int_{-1}^{1} u_h(x_i + \frac{\Delta x}{2}\xi, y_j + \frac{\Delta y}{2}\eta, 0)\phi_{m_x}(\xi)\phi_{m_y}(\eta)d\xi d\eta.$$

The stated integral can be approximated using the Gauss Legendre quadrature. If two sets of quadrature points and weights are given by,

$$\{\hat{x}_0,\ldots,\hat{x}_{q_x-1}, v_0,\ldots,v_{q_x-1}\} \text{ and } \{\hat{y}_0,\ldots,\hat{y}_{q_y-1}, w_0,\ldots,w_{q_y-1}\},$$

then for an arbitrary function $g \in L^2[(-1,1)\times(-1,1)]$,

$$\begin{aligned}
\int_{-1}^{1}\int_{-1}^{1} g(\xi,\eta)d\xi d\eta &= \int_{-1}^{1}\left(\sum_{r_x=0}^{q_x-1} v_{r_x} g(\hat{x}_{r_x},\eta)\right)d\eta \\
&= \sum_{r_y=0}^{q_y-1} w_{r_y}\left(\sum_{r_x=0}^{q_x-1} v_{r_x} g(\hat{x}_{r_x},\hat{y}_{r_y})\right) \\
&= \sum_{r_y=0}^{q_y-1}\sum_{r_x=0}^{q_x-1} w_{r_y} v_{r_x} g(\hat{x}_{r_x},\hat{y}_{r_y}).
\end{aligned}$$

The accuracy of the initial condition can be measured using the same procedure as in Section A.1 in two dimensions. Using the theory of Taylor series, the error of the initial condition is expected to be,

$$\begin{aligned}
||u(x,y,T) - u_h(x,y,T)|| &= ||u(x,y,T) - \sum_{\ell_x=0}^{k}\sum_{\ell_y=0}^{k} u_{ij}^{(\ell_x,\ell_y)}(T)\phi_{\ell_x}(\xi)\phi_{\ell_y}(\eta)|| \\
&= \mathcal{O}((\Delta x)^{k+1} + (\Delta y)^{k+1}).
\end{aligned}$$

If $\Delta x = \Delta y$, this reduces to $\mathcal{O}(\Delta x)^{k+1}$. In Table 9 the found values for the order are stated for $T = 0.5$, using equation (69) and initial condition $u^0(x,y) = \sin(2\pi x)\cdot\sin(2\pi y)$. As can be seen in this table, the method is of order $k+1$ for $k \geq 1$. For $k = 0$, the method is not so accurate: this is also found by Cockburn [8], where it is stated that the order is proven to be $k+1$ only for $k = 1,2,3$. For $k = 1$, the order of the error increases if the mesh is refined.

	$\Phi^0, \nu = 0.5$				
$N_x + 1$	10	20	40	80	160
$\|\|e(0.5)\|\|_\infty$	0.9991	0.8672	0.6448	0.4016	0.2246
order	-	0.2042	0.4275	0.6830	0.8384
	$\Phi^1, \nu = 0.2$				
$N_x + 1$	10	20	40	80	160
$\|\|e(0.5)\|\|_\infty$	0.2725	0.0883	0.0262	0.0070	0.0018
order	-	1.6258	1.7546	1.9016	1.9580
	$\Phi^2, \nu = 0.1$				
$N_x + 1$	10	20	40	80	160
$\|\|e(0.5)\|\|_\infty$	0.0406	0.0041	4.6935e-04	5.5158e-05	6.6163e-06
order	-	3.3079	3.1268	3.0890	3.0594
	$\Phi^3, \nu = 0.0015$				
$N_x + 1$	10	20	40	80	160
$\|\|e(0.5)\|\|_\infty$	0.0031	2.1985e-04	1.4234e-05		
order	-	3.8274	3.9490		

Table 9: Errors of the DG approximation at $T = 0.5$, linear advection equation in 2D, $u^0(x, y) = \sin(2\pi x) \cdot \sin(2\pi y)$, $N_x = N_y$. No entry: not computed

A.3 The inviscid Burgers' equation

In Section 1.4.1, it has been proven that for $t < \frac{2}{\pi}$, the solution using initial condition (28) is single valued (does not contain shocks). In this appendix, the accuracy of the numerical results is explored, for small times. Similarly to Appendix A.1, it is expected that the numerical solution is of order $k+1$ before the shock is formed. In Table 10, the order of the error can be seen. Indeed, these orders are approximately equal to $k+1$.

	Φ^0				
$N+1$	10	20	40	80	160
$\|\|e(0.1)\|\|_\infty$	0.1921	0.1014	0.0528	0.0267	0.0134
order	-	0.9215	0.9418	0.9856	0.9933
	Φ^1				
$N+1$	10	20	40	80	160
$\|\|e(0.1)\|\|_\infty$	0.0298	0.0081	0.0021	5.3640e-04	1.3501e-04
order	-	1.8786	1.9358	1.9796	1.9902
	Φ^2				
$N+1$	10	20	40	80	160
$\|\|e(0.1)\|\|_\infty$	0.0020	3.7663e-04	5.0363e-05	6.6787e-06	8.5945e-07
order	-	2.4084	2.9027	2.9147	2.9581
	Φ^3				
$N+1$	10	20	40	80	160
$\|\|e(0.1)\|\|_\infty$	1.0434e-04	1.1183e-05	7.2528e-07	4.5295e-08	2.8494e-09
order	-	3.2219	3.9466	4.0011	3.9906

Table 10: Errors of the DG approximation at $T = 0.1$, inviscid Burgers' equation, initial condition (28), $\nu = 0.05$

A.4 The Euler equations

In this appendix, the weak formulation of the Euler equations is computed (introduced in Section 1.4.2). The DG approximation on each element, I_j, is used, given by,

$$\mathbf{u}_h(x) = \sum_{\ell=0}^{k} \mathbf{u}_j^{(\ell)} \phi_\ell(\xi), \quad \xi = \frac{2}{\Delta x}(x - x_j), \text{ for } x \in I_j, j = 0, \ldots, N,$$

where the vector $\mathbf{u}_j^{(\ell)} = \left(u_j^{(1,\ell)}, u_j^{(2,\ell)}, u_j^{(3,\ell)}\right)^\top$ consists of the DG coefficients for $\mathbf{u}_h$ in I_j, belonging to $\phi_\ell, \ell = 0, \ldots, k$.
The weak formulation of equation (9) transforms to,

$$\frac{\Delta x}{2}\frac{d}{dt}\mathbf{u}_j^{(b,m)} = \int_{-1}^{1} (\mathbf{f}(\mathbf{u}_h)(\xi))^{(b)} \phi'_m(\xi) d\xi + \hat{\mathbf{F}}^{(b)}_{j-\frac{1}{2}} v^+_{h,j-\frac{1}{2}} - \hat{\mathbf{F}}^{(b)}_{j+\frac{1}{2}} v^-_{h,j+\frac{1}{2}},$$

where $\hat{\mathbf{F}}^{(b)}_{j\pm\frac{1}{2}}$ denotes the flux of $u_h^{(b)}(x)$ through the boundary $x_{j\pm\frac{1}{2}}$, $b = 1, 2, 3$. This means that the following equation should be solved:

$$\frac{d}{dt}\mathbf{u}_j^{(m)} = \frac{2}{\Delta x}\left(\int_{-1}^{1} (\mathbf{f}(\mathbf{u}_h)(\xi))\phi'_m(\xi) d\xi + \hat{\mathbf{F}}_{j-\frac{1}{2}} v^+_{h,j-\frac{1}{2}} - \hat{\mathbf{F}}_{j+\frac{1}{2}} v^-_{h,j+\frac{1}{2}}\right), \tag{74}$$

$\hat{\mathbf{F}}_{j\pm\frac{1}{2}} = \left(\hat{\mathbf{F}}^{(1)}_{j\pm\frac{1}{2}}, \hat{\mathbf{F}}^{(2)}_{j\pm\frac{1}{2}}, \hat{\mathbf{F}}^{(3)}_{j\pm\frac{1}{2}}\right)^\top$.
In this case, the fluxes through the boundaries are computed using a Roe linearization (one of the approximate Riemann solvers in LeVeque, [18]). This method uses the local Lax Friedrichs flux as defined in equations (11), but extends this to a more general formula. Therefore, several quantities should be defined, such as the total specific enthalpy, H, given by,

$$H = \frac{E + p}{\rho},$$

and the sound speed of the polytropic gas, which equals,

$$c = \sqrt{\frac{\gamma p}{\rho}}.$$

For element $I_j, j \in \{0, \ldots, N\}$, the so-called Roe averages are given by,

$$u_{\text{Roe}} = \frac{\sqrt{\bar{\rho}_j}\bar{u}_j + \sqrt{\bar{\rho}_{j+1}}\bar{u}_{j+1}}{\sqrt{\bar{\rho}_j} + \sqrt{\bar{\rho}_{j+1}}}, \qquad \text{velocity;} \tag{75a}$$

$$H_{\text{Roe}} = \frac{\sqrt{\bar{\rho}_j}\bar{H}_j + \sqrt{\bar{\rho}_{j+1}}\bar{H}_{j+1}}{\sqrt{\bar{\rho}_j} + \sqrt{\bar{\rho}_{j+1}}}, \qquad \text{total specific enthalpy;} \tag{75b}$$

$$c_{\text{Roe}} = \sqrt{(\gamma - 1)(H_{\text{Roe}} - \frac{1}{2}u^2_{\text{Roe}})}, \qquad \text{sound speed,} \tag{75c}$$

which are computed using the average values of ρ, u and H on the intervals I_j and I_{j+1}. Note that the DG method approximates the solution on each element $I_j (j = 0, \ldots, N)$, to be a sum of the scaled Legendre polynomials. Therefore, the approximate solution, $\mathbf{u}_h$, is smooth on each element I_j, which means that the Jacobian matrix, $\mathbf{f}'(\mathbf{u}_h)$, can be used to write,

$$0 = \mathbf{u}_{h,t} + \mathbf{f}(\mathbf{u}_h)_x = \mathbf{u}_{h,t} + \mathbf{f}'(\mathbf{u}_h)\mathbf{u}_{h,x}. \tag{76}$$

The matrix $f'(\mathbf{u}_h)$ is given by Leveque, [18], and computed using the Roe averages (75):

$$\mathbf{f}'(\mathbf{u}_h) = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2}(\gamma-3)u_{\text{Roe}}^2 & (3-\gamma)u_{\text{Roe}} & \gamma-1 \\ \frac{1}{2}(\gamma-1)u_{\text{Roe}}^3 - u_{\text{Roe}}H_{\text{Roe}} & H_{\text{Roe}} - (\gamma-1)u_{\text{Roe}}^2 & \gamma u_{\text{Roe}} \end{pmatrix}. \tag{77}$$

The matrix of eigenvalues of Jacobian (77) is given by,

$$\Lambda = \begin{pmatrix} \lambda^1 & 0 & 0 \\ 0 & \lambda^2 & 0 \\ 0 & 0 & \lambda^3 \end{pmatrix} = \begin{pmatrix} u_{\text{Roe}} - c_{\text{Roe}} & 0 & 0 \\ 0 & u_{\text{Roe}} & 0 \\ 0 & 0 & u_{\text{Roe}} + c_{\text{Roe}} \end{pmatrix}. \tag{78a}$$

The accompanying matrix of right eigenvectors is,

$$R = \begin{pmatrix} \mathbf{r}^1 & \mathbf{r}^2 & \mathbf{r}^3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ u_{\text{Roe}} - c_{\text{Roe}} & u_{\text{Roe}} & u_{\text{Roe}} + c_{\text{Roe}} \\ H_{\text{Roe}} - u_{\text{Roe}}c_{\text{Roe}} & \frac{1}{2}u_{\text{Roe}}^2 & H_{\text{Roe}} + u_{\text{Roe}}c_{\text{Roe}} \end{pmatrix}, \tag{78b}$$

with $\mathbf{f}'(\mathbf{u}_h)R = R\Lambda$. It should be noticed that the rows of the inverse matrix R^{-1} are the left eigenvectors of $\mathbf{f}'(\mathbf{u}_h)$, such that $R^{-1}\mathbf{f}'(\mathbf{u}_h) = \Lambda R^{-1}$. Therefore, the matrix of left eigenvectors is given by,

$$L = \begin{pmatrix} \mathbf{L}^1 \\ \mathbf{L}^2 \\ \mathbf{L}^3 \end{pmatrix} = R^{-1}. \tag{78c}$$

This means that equation (76) equals,

$$\mathbf{u}_{h,t} + \mathbf{f}'(\mathbf{u}_h)\mathbf{u}_{h,x} = \mathbf{u}_{h,t} + R\Lambda R^{-1}\mathbf{u}_{h,x} = 0.$$

Multiplying on the left with R^{-1}, gives the following uncoupled system of partial differential equations ($\mathbf{w}_h = R^{-1}\mathbf{u}_h$):

$$\left(R^{-1}\mathbf{u}_h\right)_t + \Lambda\left(R^{-1}\mathbf{u}_h\right)_x \equiv \mathbf{w}_{h,t} + \Lambda\mathbf{w}_{h,x} = 0,$$

which is in general hyperbolic (LeVeque, [18]), because the eigenvalues of Λ are real, and there is a complete set of eigenvectors R.
Roe uses the matrices R and L, to express the difference between $\mathbf{u}^-_{h,j+\frac{1}{2}}$ and $\mathbf{u}^+_{h,j+\frac{1}{2}}$ as:

$$\mathcal{W}^b_{j+\frac{1}{2}} = \mathbf{L}^b(\mathbf{u}^+_{h,j+\frac{1}{2}} - \mathbf{u}^-_{h,j+\frac{1}{2}})\mathbf{r}^b, b = 1, 2, 3.$$

The used flux, developed by Roe and using the local Lax Friedrichs flux, is now given by,

$$\hat{\mathbf{F}}_{j+\frac{1}{2}} = \frac{1}{2}\left(\mathbf{f}(\mathbf{u}_h)^-_{j+\frac{1}{2}} + \mathbf{f}(\mathbf{u}_h)^+_{j+\frac{1}{2}} - \sum_{b=1}^{3} a^b_{j+\frac{1}{2}} \mathcal{W}^b_{j+\frac{1}{2}}\right), j = 0, \ldots, N-1, \tag{79}$$

where

$$a^b_{j+\frac{1}{2}} = \max\left(\left|\lambda^b_j\right|, \left|\lambda^b_{j+1}\right|\right),$$

and the averages of velocity and sound speed are used for these eigenvalues:

$$\lambda^1_j = \bar{u}_j - \bar{c}_j,\ \lambda^2_j = \bar{u}_j,\ \lambda^3_j = \bar{u}_j + \bar{c}_j, j = 0, \ldots, N.$$

To check the accuracy of the code, the following initial conditions are used:

$$\rho(x,0) = 1 + 0.2\sin(x), \tag{80a}$$
$$u(x,0) = 1 = p(x,0), \tag{80b}$$

for $x \in [0, 2\pi]$, with periodic boundary conditions. Note that for the energy, the initial condition is given by,

$$E(x,0) = \frac{p(x,0)}{\gamma - 1} + \frac{1}{2}\rho(x,0)u(x,0)^2 = \frac{1}{\gamma - 1} + \frac{1}{2}(1 + 0.2\sin(x)). \tag{80c}$$

It should be noticed that these initial conditions make the Euler equations linear, since equations (33) transform into one scalar advection equation, given by $\rho_t + \rho_x = 0$. Therefore, the exact solution of system (34) with initial condition (80) is given by,

$$\rho(x,t) = 1 + 0.2\sin(x - t), \tag{81a}$$
$$u(x,t) = 1 = p(x,t), \tag{81b}$$

which means that,

$$E(x,t) = \frac{1}{\gamma - 1} + \frac{1}{2}(1 + 0.2\sin(x - t)), \tag{81c}$$

$x \in [0, 2\pi], t \geq 0$.
For the time stepping, the third order TVD Runge Kutta method is used. In Table 11, the accuracy of this method can be seen by inspecting the errors for $\rho(x,T)$ at $T = 0.5$, given by $||e(0.5)||_\infty = \max\{|\rho(x,0.5) - \rho_h(x,0.5)|, x \in [0,2\pi]\}$. Indeed, the errors are of order $k+1$, so the code works as expected. For $k = 3, N = 160$, there is a round off error, such that the order is less than 4.

	$\Phi^0, \nu = 0.1$				
$N+1$	10	20	40	80	160
$\|\|e(0.5)\|\|_\infty$	0.0645	0.0326	0.0163	0.0081	0.0041
order	-	0.9844	1.0020	1.0035	1.0024
	$\Phi^1, \nu = 0.1$				
$N+1$	10	20	40	80	160
$\|\|e(0.5)\|\|_\infty$	0.0098	0.0026	6.5629e-4	1.6525e-4	4.1417e-5
order	-	1.9272	1.9795	1.9897	1.9964
	$\Phi^2, \nu = 0.01$				
$N+1$	10	20	40	80	160
$\|\|e(0.5)\|\|_\infty$	5.1293e-4	6.2002e-5	7.6521e-6	9.5866e-7	1.1986e-7
order	-	3.0484	3.0184	2.9968	2.9997
	$\Phi^3, \nu = 0.001$				
$N+1$	10	20	40	80	160
$\|\|e(0.5)\|\|_\infty$	1.7263e-5	9.8646e-7	5.2153e-8	3.3757e-9	5.8041e-10
order	-	4.1293	4.2414	3.9495	2.5400

Table 11: Errors of the DG approximation for ρ at $T = 0.5$, Euler equations, initial condition (80)

B Examples and details about multiwavelets

This appendix contains an example of the theory of multiwavelets, as developed in Section 1.3. Furthermore, details are given about the multiwavelet decomposition and reconstruction.
As an example, the basis of scaling functions for V_0^4 is given by the scaled Legendre polynomials (see equations (3)), and visualized in Figure 59. The functions vanish outside the interval $[-1, 1]$.
The basis for V_1^4 can be seen in Figure 60a. Functions $\phi_{\ell 0}^1$ and $\phi_{\ell 1}^1, \ell = 0, \ldots 4$, vanish outside $[-1, 0]$ and $[0, 1]$, respectively.

The multiwavelet basis $\{\psi_0, \ldots, \psi_4\}$, for W_0^4 should satisfy ($\ell \in \{0, \ldots, 4\}$):

1. If $\psi_\ell \in W_0^4$ then $\psi_\ell \in V_1^4$, but $W_0^4 \neq V_1^4$ ($V_0^4 \oplus W_0^4 = V_1^4$, see equation (16));
2. $\langle \psi_\ell, \psi_\ell \rangle = 1$: the system is orthonormal;
3. $\langle \psi_\ell, \phi_i \rangle = 0$, for all functions $\phi_i \in V_0^4$, using equation (16);
4. It is possible to define every basis function, $\phi_{\ell j}^1 \in V_1^4, j = 0, 1$, using a linear combination of the basis functions $\phi_i \in V_0^4$ and $\psi_i \in W_0^4, i = 0, \ldots, 4$.

The multiwavelets that satisfy these requirements are developed by Alpert, [2], a good explanation is given in Hovhannisyan, [12]. In Table 12, this multiwavelet basis is given. The functions $f_i, i = 0, \ldots, p$, are given for $x \in (0, 1)$, and extended to the interval $(-1, 0)$ as an odd or even function, according to the formula $f_i(x) = (-1)^{i+p+1} f_i(-x)$. Outside the interval $(-1, 1)$, f_i is zero.

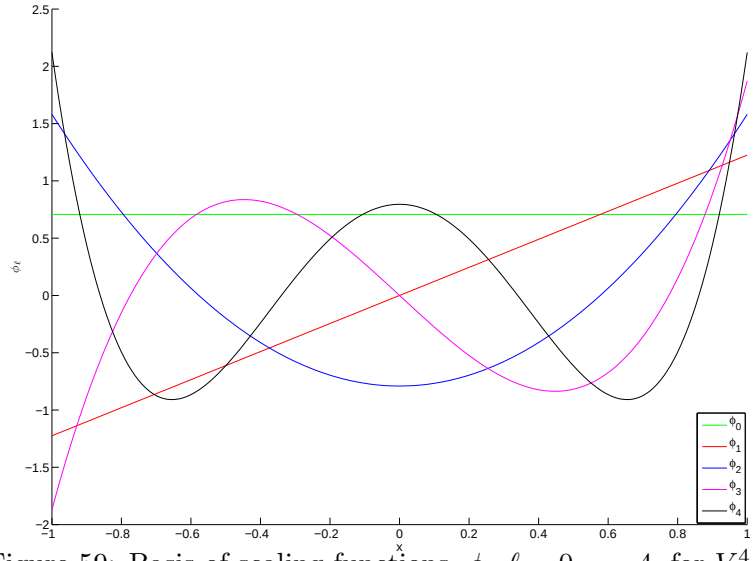
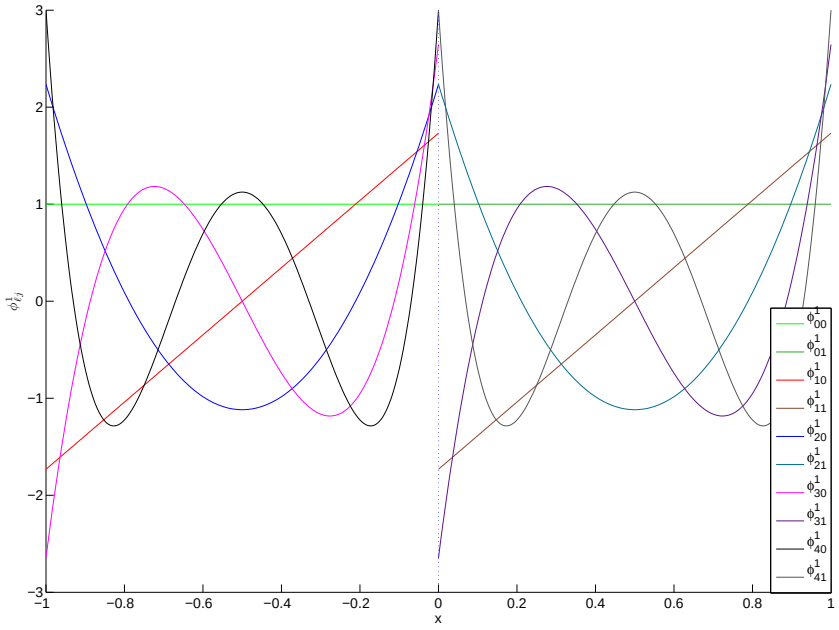


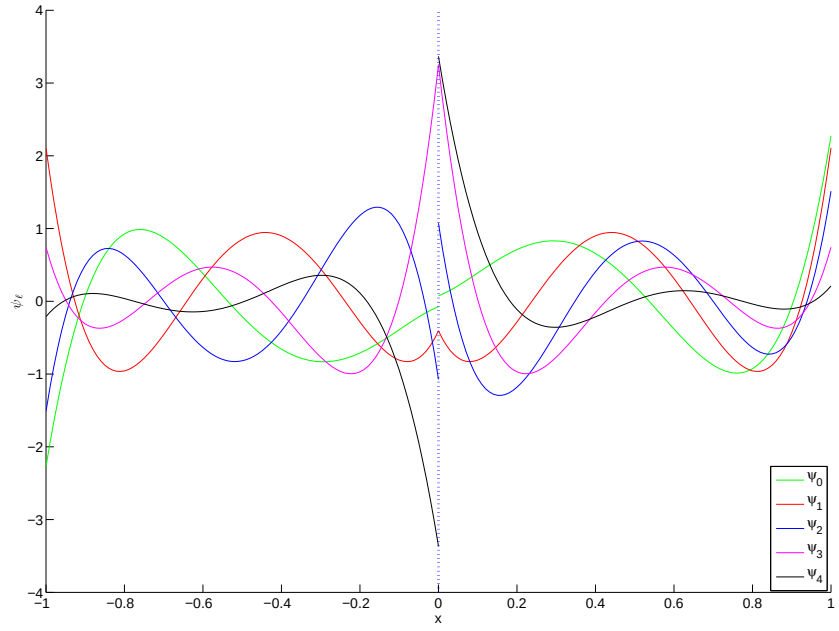
Figure 59: Basis of scaling functions, $\phi_\ell, \ell = 0, \ldots, 4$, for V_0^4

$p=0$		
$f_0(x)$	$=$	$\sqrt{\frac{1}{2}}$
$p=1$		
$f_0(x)$	$=$	$\sqrt{\frac{3}{2}}(-1+2x)$
$f_1(x)$	$=$	$\sqrt{\frac{1}{2}}(-2+3x)$
$p=2$		
$f_0(x)$	$=$	$\frac{1}{3}\sqrt{\frac{1}{2}}(1-24x+30x^2)$
$f_1(x)$	$=$	$\frac{1}{2}\sqrt{\frac{3}{2}}(3-16x+15x^2)$
$f_2(x)$	$=$	$\frac{1}{3}\sqrt{\frac{5}{2}}(4-15x+12x^2)$
$p=3$		
$f_0(x)$	$=$	$\sqrt{\frac{15}{34}}(1+4x-30x^2+28x^3)$
$f_1(x)$	$=$	$\sqrt{\frac{1}{42}}(-4+105x-300x^2+210x^3)$
$f_2(x)$	$=$	$\frac{1}{2}\sqrt{\frac{35}{34}}(-5+48x-105x^2+64x^3)$
$f_3(x)$	$=$	$\frac{1}{2}\sqrt{\frac{5}{42}}(-16+105x-192x^2+105x^3)$
$p=4$		
$f_0(x)$	$=$	$\sqrt{\frac{1}{186}}(1+30x+210x^2-840x^3+630x^4)$
$f_1(x)$	$=$	$\frac{1}{2}\sqrt{\frac{1}{38}}(-5-144x+1155x^2-2240x^3+1260x^4)$
$f_2(x)$	$=$	$\sqrt{\frac{35}{14694}}(22-735x+3504x^2-5460x^3+2700x^4)$
$f_3(x)$	$=$	$\frac{1}{8}\sqrt{\frac{21}{38}}(35-512x+1890x^2-2560x^3+1155x^4)$
$f_4(x)$	$=$	$\frac{1}{2}\sqrt{\frac{7}{158}}(32-315x+960x^2-1155x^3+480x^4)$

Table 12: Alpert's multiwavelet basis for $W_0^p, p=0,\ldots,4,\ x\in(0,1)$. The functions $f_i, i=0,\ldots,p$, are extended to the interval $(-1,0)$ as an odd or even function, according to the formula $f_i(x)=(-1)^{i+p+1}f_i(-x)$. Outside the interval $(-1,1)$, f_i is zero.



(a) Basis of scaling functions for V_1^4



(b) Multiwavelet basis $\psi_\ell, \ell = 0, \ldots, 4$, for W_0^4

Figure 60: Scaling functions on level 1, and multiwavelet basis for W_0^4

To effectively use multiwavelets in approximating a function, the signal is decomposed into a coarse approximation (scaling functions on level 0), and higher frequencies (multiwavelets on levels 0 to $n-1$), as described in Section 1.3.
Besides that, multiwavelet reconstruction can be used to compute $s^n_{\ell j}$ from the coefficients $s^0_{\ell 0}, d^m_{\ell j}$, $m = 0, \ldots, n$, $j = 0, \ldots, 2^m - 1$, $\ell = 0, \ldots, p$.
In the decomposition and reconstruction steps so-called quadrature mirror filter coefficients are needed. Two different types of coefficients are considered: lowpass filter coefficients (belonging to scaling functions), and highpass filter coefficients (belonging to multiwavelets).

Lowpass filter
Let basis function $\phi_\ell \in V^p_0$ be given, $\ell \in \{0, \ldots, p\}$. Note that $V^p_0 \subset V^p_1$ such that also $\phi_\ell \in V^p_1$. This means that it is possible to write, using equation (17), and the idea in Ruch, [20],

$$\phi_\ell(x) = P^p_1 \phi_\ell(x) = \sum_{r=0}^{p} \langle \phi_\ell, \phi^1_{r0} \rangle \phi^1_{r0}(x) + \sum_{r=0}^{p} \langle \phi_\ell, \phi^1_{r1} \rangle \phi^1_{r1}(x). \tag{82}$$

Using equation (15), this becomes,

$$\phi_\ell(x) = \sqrt{2} \left\{ \sum_{r=0}^{p} \langle \phi_\ell, \phi^1_{r0} \rangle \phi_r(2(x+1)-1) + \sum_{r=0}^{p} \langle \phi_\ell, \phi^1_{r1} \rangle \phi_r(2(x+1)-3) \right\} \tag{83a}$$

$$= \sqrt{2} \sum_{j=0}^{1} \sum_{r=0}^{p} h^{(j)}_{\ell r} \phi_r(2(x-j)+1), \tag{83b}$$

where $h^{(j)}_{\ell r} = \langle \phi_\ell, \phi^1_{rj} \rangle, j = 0, 1, r = 0, \ldots, p$.

Consider $h^{(0)}_{\ell r}$ for some $r \in \{0, \ldots, p\}$. By definition,

$$h^{(0)}_{\ell r} = \langle \phi_\ell, \phi^1_{r0} \rangle = \int_{-1}^{1} \phi_\ell(x) \phi^1_{r0}(x) dx = \sqrt{2} \int_{-1}^{1} \phi_\ell(x) \phi_r(2x+1) dx.$$

Note that $\phi_r(2x+1)$ is nonzero only if $(2x+1) \in [-1, 1]$ which means $x \in [-1, 0]$. This gives,

$$h^{(0)}_{\ell r} = \sqrt{2} \int_{-1}^{0} \phi_\ell(x) \phi_r(2x+1) dx = \frac{1}{\sqrt{2}} \int_{-1}^{1} \phi_\ell \left(\frac{y-1}{2} \right) \phi_r(y) dy \tag{84a}$$

$$\approx \frac{1}{\sqrt{2}} \sum_{m=0}^{q-1} w_m \phi_\ell \left(\frac{\hat{x}_m - 1}{2} \right) \phi_r(\hat{x}_m), \tag{84b}$$

where the last equation uses a Gauss-Legendre quadrature with points $\hat{x}_0, \ldots, \hat{x}_{q-1}$ and weights $w_0, \ldots, w_{q-1}$.

Similarly, it is found that,

$$h_{\ell r}^{(1)} = \sqrt{2}\int_0^1 \phi_\ell(x)\phi_r(2x-1)dx = \frac{1}{\sqrt{2}}\int_{-1}^1 \phi_\ell\left(\frac{y+1}{2}\right)\phi_r(y)dy,$$

such that $h_{\ell r}^{(1)}$ can be approximated using $\frac{1}{\sqrt{2}}\sum_{m=0}^{q-1} w_m \phi_\ell\left(\frac{\hat{x}_m+1}{2}\right)\phi_r(\hat{x}_m)$.
The coefficients $h_{\ell r}^{(j)}, j=0,1$, form the so-called lowpass scaling filter.

Highpass filter
The same approach can be used for the multiwavelets with respect to W_0^p. Let basis function $\psi_\ell \in W_0^p$ be given, $\ell \in \{0,\ldots,p\}$. Because $W_0^p \in V_1^p$ (see equation (16)), it holds that,

$$\psi_\ell(x) = \sqrt{2}\sum_{j=0}^{1}\sum_{r=0}^{p} g_{\ell r}^{(j)}\phi_r(2(x-j)+1), \tag{85}$$

where $g_{\ell r}^{(j)} = \langle \psi_\ell, \phi_{rj}^1\rangle, j=0,1$ (cf. equations (82) to (83)).
Therefore, the highpass filters are given as:

$$\begin{aligned} g_{\ell r}^{(0)} &\approx \tfrac{1}{\sqrt{2}}\textstyle\sum_{m=0}^{q-1} w_m \psi_\ell\left(\frac{\hat{x}_m-1}{2}\right)\phi_r(\hat{x}_m), \\ g_{\ell r}^{(1)} &\approx \tfrac{1}{\sqrt{2}}\textstyle\sum_{m=0}^{q-1} w_m \psi_\ell\left(\frac{\hat{x}_m+1}{2}\right)\phi_r(\hat{x}_m). \end{aligned}$$

The coefficients $h_{\ell r}^{(0)}$ and $g_{\ell r}^{(0)}$ differ from the definition that Alpert, [3] and Archibald, [4] use (but I think they are wrong).

Multiwavelet decomposition uses equations (15), (18a), and (83) for an arbitrary function $f \in L^2(-1,1)$:

$$\begin{aligned} s_{\ell j}^{n-1} &= \langle f, \phi_{\ell j}^{n-1}\rangle = \langle f, 2^{\frac{n-1}{2}}\phi_\ell(2^{n-1}(x+1)-2j-1)\rangle \\ &= \left\langle f, 2^{\frac{n-1}{2}}\sqrt{2}\sum_{r=0}^{p}\Big\{h_{\ell r}^{(0)}\phi_r(2(2^{n-1}(x+1)-2j-1)+1)\right. \\ &\quad \left. + h_{\ell r}^{(1)}\phi_r(2(2^{n-1}(x+1)-2j-1)-1)\Big\}\right\rangle \\ &= \sum_{r=0}^{p}\Big(h_{\ell r}^{(0)}\left\langle f, 2^{\frac{n}{2}}\phi_r(2^n(x+1)-2\cdot 2j-1)\right\rangle \\ &\quad + \; h_{\ell r}^{(1)}\left\langle f, 2^{\frac{n}{2}}\phi_r(2^n(x+1)-2(2j+1)-1)\right\rangle\Big) \\ &= \sum_{r=0}^{p}\left(h_{\ell r}^{(0)} s_{r,2j}^n + h_{\ell r}^{(1)} s_{r,2j+1}^n\right). \end{aligned} \tag{86}$$

Likewise, it holds that,

$$d_{\ell j}^{n-1} = \sum_{r=0}^{p}\left(g_{\ell r}^{(0)} s_{r,2j}^n + g_{\ell r}^{(1)} s_{r,2j+1}^n\right), \ell = 0,\ldots,p, j = 0,\ldots,2^{n-1}-1. \tag{87}$$

Thus, starting with $2^n(p+1)$ values for $s^n_{\ell j}$ and $d^n_{\ell j}$, the decomposition procedure can be applied repeatedly to compute the coefficients on coarser levels, $m = n-1, n-2, \ldots, 0$.

For multiwavelet reconstruction, the following matrices are defined:

$$H^{(b)} = \left\{ h^{(b)}_{i\ell} \right\}, G^{(b)} = \left\{ g^{(b)}_{i\ell} \right\}, \tag{88}$$

$b = 0, 1;\ i, \ell = 0, \ldots, p$, and,

$$S^m = \left(\mathbf{s}^m_0 \ldots \mathbf{s}^m_{2^m-1} \right), D^m = \left(\mathbf{d}^m_0 \ldots \mathbf{d}^m_{2^m-1} \right), m = 0, \ldots, n, \tag{89}$$

where,

$$\begin{aligned} \mathbf{s}^m_j &= \left(s^m_{0j} \cdots s^m_{pj} \right)^\top, \\ \mathbf{d}^m_j &= \left(d^m_{0j} \cdots d^m_{pj} \right)^\top, \end{aligned}$$

$j = 0, \ldots, 2^m - 1$.
Using this notation, the following relations hold (see Alpert [3]):

$$H^{(0)} H^{(0)^\top} + H^{(1)} H^{(1)^\top} = I, \tag{90a}$$

$$G^{(0)} G^{(0)^\top} + G^{(1)} G^{(1)^\top} = I, \tag{90b}$$

$$H^{(0)} G^{(0)^\top} + H^{(1)} G^{(1)^\top} = 0. \tag{90c}$$

$$G^{(0)} H^{(0)^\top} + G^{(1)} H^{(1)^\top} = 0. \tag{90d}$$

Equation (90a) can be proven as follows, using equation (83) and the orthonormality of the functions $\phi_0, \ldots, \phi_p$:

$$\delta_{i\ell} = \langle \phi_i, \phi_\ell \rangle = \sqrt{2} \sum_{r=0}^{p} \left(h^{(0)}_{ir} \langle \phi_r(2x+1), \phi_\ell(x) \rangle + h^{(1)}_{ir} \langle \phi_r(2x-1), \phi_\ell(x) \rangle \right). \tag{91}$$

The two inner products on the righthand side are computed separately, using equation (83) a second time, and the same reasoning as in equations (84):

$$\begin{aligned} \langle \phi_r(2x+1), \phi_\ell(x) \rangle &= \sqrt{2} \int_{-1}^{0} \phi_r(2x+1) \left(\sum_{q=0}^{p} \left(h^{(0)}_{\ell q} \phi_q(2x+1) + h^{(1)}_{\ell q} \phi_q(2x-1) \right) \right) dx \\ &= \sqrt{2} \sum_{q=0}^{p} h^{(0)}_{\ell q} \int_{-1}^{0} \phi_r(2x+1) \phi_q(2x+1) dx = \sqrt{2} \sum_{q=0}^{p} h^{(0)}_{\ell q} \frac{1}{2} \langle \phi_r, \phi_q \rangle \\ &= \frac{1}{\sqrt{2}} h^{(0)}_{\ell r}. \end{aligned}$$

The last relation used the orthonormality relation of the scaling functions.

The second inner product in equation (91) gives,

$$\langle \phi_r(2x-1), \phi_\ell(x) \rangle = \frac{1}{\sqrt{2}} h_{\ell r}^{(1)},$$

such that equation (91) becomes,

$$\delta_{i\ell} = \langle \phi_i, \phi_\ell \rangle = \sqrt{2} \sum_{r=0}^{p} \left(h_{ir}^{(0)} \frac{1}{\sqrt{2}} h_{\ell r}^{(0)} + h_{ir}^{(1)} \frac{1}{\sqrt{2}} h_{\ell r}^{(1)} \right) = \sum_{r=0}^{p} \left(h_{ir}^{(0)} h_{\ell r}^{(0)} + h_{ir}^{(1)} h_{\ell r}^{(1)} \right),$$

from which relation (90a) easily follows.
To prove relation (90b), note that equation (85) gives,

$$\delta_{i\ell} = \langle \psi_i, \psi_\ell \rangle = \sqrt{2} \sum_{r=0}^{p} \left(g_{ir}^{(0)} \langle \phi_r(2x+1), \psi_\ell(x) \rangle + g_{ir}^{(1)} \langle \phi_r(2x-1), \psi_\ell(x) \rangle \right).$$

The same approach as above can be used.
Relations (90c) and (90d) use the orthogonality relation from equation (16) to find,

$$\begin{aligned} 0 &= \langle \phi_i, \psi_\ell \rangle &&= \sqrt{2} \textstyle\sum_{r=0}^{p} \left(h_{ir}^{(0)} \langle \phi_r(2x+1), \psi_\ell(x) \rangle + h_{ir}^{(1)} \langle \phi_r(2x-1), \psi_\ell(x) \rangle \right), \\ 0 &= \langle \psi_i, \phi_\ell \rangle &&= \sqrt{2} \textstyle\sum_{r=0}^{p} \left(g_{ir}^{(0)} \langle \phi_r(2x+1), \phi_\ell(x) \rangle + g_{ir}^{(1)} \langle \phi_r(2x-1), \phi_\ell(x) \rangle \right). \end{aligned}$$

Introducing

$$U = \begin{pmatrix} H^{(0)} & H^{(1)} \\ G^{(0)} & G^{(1)} \end{pmatrix},$$

equations (90) give $UU^\top = I$. This means that U is an orthogonal matrix and satisfies $U^\top = U^{-1}$. This condition gives rise to an additional set of relations:

$$\begin{aligned} H^{(0)^\top} H^{(0)} + G^{(0)^\top} G^{(0)} &= I, \\ H^{(1)^\top} H^{(1)} + G^{(1)^\top} G^{(1)} &= I, \\ H^{(0)^\top} H^{(1)} + G^{(0)^\top} G^{(1)} &= 0. \\ H^{(1)^\top} H^{(0)} + G^{(1)^\top} G^{(0)} &= 0. \end{aligned}$$

Using the matrix notation as introduced in (88) and (89), the multiwavelet decomposition from equations (86) and (87) can be equivalently written as:

$$\mathbf{s}_j^{n-1} = H^{(0)} \mathbf{s}_{2j}^n + H^{(1)} \mathbf{s}_{2j+1}^n, \tag{92a}$$

$$\mathbf{d}_j^{n-1} = G^{(0)} \mathbf{s}_{2j}^n + G^{(1)} \mathbf{s}_{2j+1}^n. \tag{92b}$$

The multiwavelet reconstruction follows from multiplying equation (92a) on the left by $H^{(0)^\top}$, equation (92b) on the left by $G^{(0)^\top}$, and summing, such that,

$$\begin{aligned} H^{(0)^\top}\mathbf{s}_j^{n-1} + G^{(0)^\top}\mathbf{d}_j^{n-1} = \\ = H^{(0)^\top}\left(H^{(0)}\mathbf{s}_{2j}^n + H^{(1)}\mathbf{s}_{2j+1}^n\right) + G^{(0)^\top}\left(G^{(0)}\mathbf{s}_{2j}^n + G^{(1)}\mathbf{s}_{2j+1}^n\right) \\ = \left\{H^{(0)^\top}H^{(0)} + G^{(0)^\top}G^{(0)}\right\}\mathbf{s}_{2j}^n + \left\{H^{(0)^\top}H^{(1)} + G^{(0)^\top}G^{(1)}\right\}\mathbf{s}_{2j+1}^n \\ = \mathbf{s}_{2j}^n. \end{aligned}$$

The same approach leads to,

$$H^{(1)^\top}\mathbf{s}_j^{n-1} + G^{(1)^\top}\mathbf{d}_j^{n-1} = \mathbf{s}_{2j+1}^n,$$

from which it follows that,

$$s_{\ell,2j}^n = \sum_{r=0}^{p}\left(h_{r\ell}^{(0)}s_{rj}^{n-1} + g_{r\ell}^{(0)}d_{rj}^{n-1}\right), \tag{93a}$$

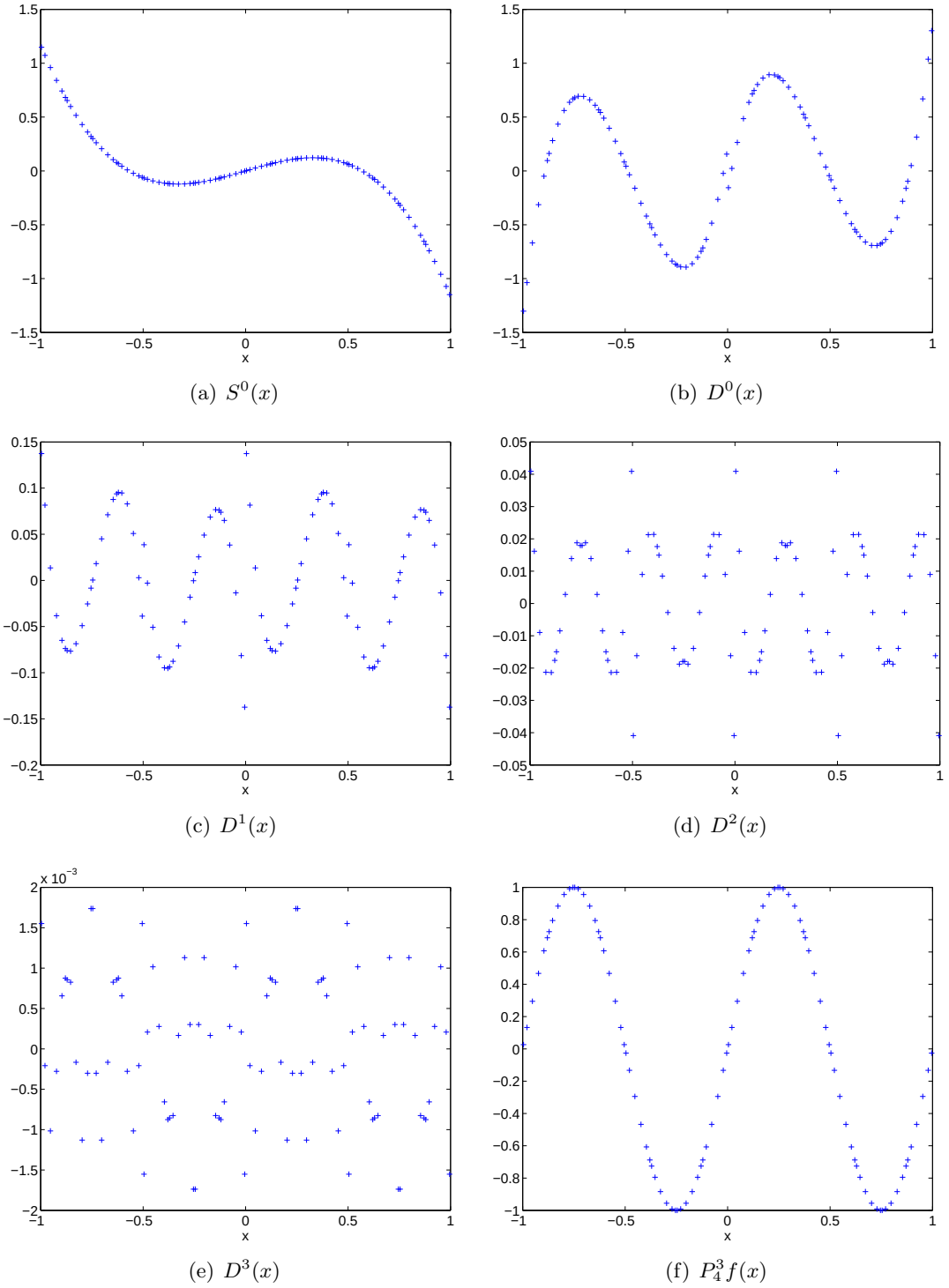
$$s_{\ell,2j+1}^n = \sum_{r=0}^{p}\left(h_{r\ell}^{(1)}s_{rj}^{n-1} + g_{r\ell}^{(1)}d_{rj}^{n-1}\right), \tag{93b}$$

$n \in \mathbb{N}, j = 0, \ldots, 2^{n-1} - 1$, and $\ell, r = 0, \ldots, p$ (multiwavelet reconstruction).

As an example, the function $f(x) = \sin(2\pi x), x \in [-1, 1]$ is decomposed into,

$$P_n^p f(x) = \sum_{\ell=0}^{p} s_{\ell 0}^0 \phi_\ell(x) + \sum_{m=0}^{n-1}\sum_{j=0}^{2^m-1}\sum_{\ell=0}^{p} d_{\ell j}^m \psi_{\ell j}^m(x) = S^0(x) + \sum_{m=0}^{n-1} D^m(x).$$

In Figure 61 the different parts of the projection $P_n^p f(x)$ are plotted. Figure 61a is a polynomial of degree p on the interval $[-1, 1]$ (a sum of scaling functions in level 0). In Figure 61b two piecewise polynomials on $[-1, 0]$ and $[0, 1]$ can be recognized (a sum of the multiwavelets in level 0), etc.



(a) $S^0(x)$ (b) $D^0(x)$

(c) $D^1(x)$ (d) $D^2(x)$

(e) $D^3(x)$ (f) $P_4^3 f(x)$

Figure 61: Decomposition procedure, $f(x) = \sin(2\pi x), p = 3, n = 4$ ($2^n = 16$ elements)

C Maximum principle limiter and multiwavelets

In this appendix, the maximum principle limiter of Section 1.5.6 (which is applied to DG coefficients) is compared to the multiwavelet decomposition. Because,

$$2^{-\frac{n}{2}} u_j^{(\ell)} = s_{\ell j}^n, j = 0, \ldots, N, \ell = 0, \ldots, k,$$

the maximum principle limiter computes,

$$\widetilde{s}_{0j}^n = s_{0j}^n, \text{ and } \widetilde{s}_{\ell j}^n = \theta s_{\ell j}^n, \ell = 1, \ldots, k,$$

where θ is defined in equation (41b).
Using decomposition (Appendix B), this can be rewritten as a limiter for the scaling function and multiwavelet coefficients. For the multiwavelet coefficients on level $n-1$, this gives,

$$\begin{aligned}
\widetilde{d}_{\ell j}^{n-1} &= \sum_{r=0}^{k} \left(g_{\ell r}^{(0)} \widetilde{s}_{r,2j}^n + g_{\ell r}^{(1)} \widetilde{s}_{r,2j+1}^n \right) \\
&= g_{\ell 0}^{(0)} s_{0,2j}^n + g_{\ell 0}^{(1)} s_{0,2j+1}^n + \sum_{r=1}^{k} \left(g_{\ell r}^{(0)} \theta s_{r,2j}^n + g_{\ell r}^{(1)} \theta s_{r,2j+1}^n \right) \\
&= \theta \sum_{r=0}^{k} \left(g_{\ell r}^{(0)} s_{r,2j}^n + g_{\ell r}^{(1)} s_{r,2j+1}^n \right) + (1-\theta) \left(g_{\ell 0}^{(0)} s_{0,2j}^n + g_{\ell 0}^{(1)} s_{0,2j+1}^n \right) \\
&= \theta d_{\ell j}^{n-1} + (1-\theta) \left(g_{\ell 0}^{(0)} s_{0,2j}^n + g_{\ell 0}^{(1)} s_{0,2j+1}^n \right).
\end{aligned}$$

Likewise,

$$\widetilde{s}_{\ell j}^{n-1} = \theta s_{\ell j}^{n-1} + (1-\theta) \left(h_{\ell 0}^{(0)} s_{0,2j}^n + h_{\ell 0}^{(1)} s_{0,2j+1}^n \right).$$

Unfortunately, for levels $n-2, \ldots, 0$, it is not easy the compute the limited coefficients:

$$\begin{aligned}
\widetilde{d}_{\ell j}^{n-2} &= \sum_{r=0}^{k} \left(g_{\ell r}^{(0)} \widetilde{s}_{r,2j}^{n-1} + g_{\ell r}^{(1)} \widetilde{s}_{r,2j+1}^{n-1} \right) \\
&= \sum_{r=0}^{k} \left(g_{\ell r}^{(0)} \left(\theta s_{r,2j}^{n-1} + (1-\theta) \left(h_{r0}^{(0)} s_{0,4j}^n + h_{r0}^{(1)} s_{0,4j+1}^n \right) \right) \right. \\
&\quad \left. + g_{\ell r}^{(1)} \left(\theta s_{r,2j+1}^{n-1} + (1-\theta) \left(h_{r0}^{(0)} s_{0,4j+2}^n + h_{r0}^{(1)} s_{0,4j+3}^n \right) \right) \right) \\
&= \theta d_{\ell j}^{n-2} + (1-\theta) \sum_{r=0}^{k} \left(g_{\ell r}^{(0)} (h_{r0}^{(0)} s_{0,4j}^n + h_{r0}^{(1)} s_{0,4j+1}^n) \right. \\
&\quad \left. + g_{\ell r}^{(1)} (h_{r0}^{(0)} s_{0,4j+2}^n + h_{r0}^{(1)} s_{0,4j+3}^n) \right),
\end{aligned}$$

Using these formulas, the computation of the limited multiwavelet coefficients would take lots of time, and this is therefore not used.

D Decomposition inside each element

Instead of using the multiwavelet decomposition on the interval $[a, b]$, it is possible to use this decomposition inside each element. Using this approach, it may be possible to avoid from getting Gibbs' phenomenon (see Section 2.1).

Therefore, each element $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}], i = 0, \ldots, N$, is divided into 2^n subintervals, with boundaries (using definition (14)),

$$\left[x_i + \frac{\Delta x}{2}\left(-1 + 2^{-n+1}j\right), x_i + \frac{\Delta x}{2}\left(-1 + 2^{-n+1}(j+1)\right)\right], j = 0, \ldots, 2^n - 1.$$

An example of these nested grids can be seen in Figure 62.

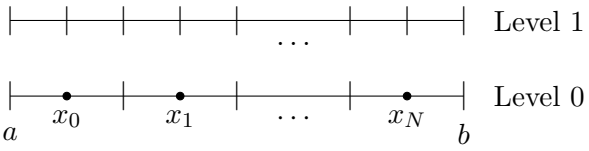


Figure 62: Example of nested grids for multiwavelet decomposition inside element I_i, $i = 0, \ldots, N$

Projecting $u^0(x)$ onto the space V_n^k gives on element I_i:

$$\begin{aligned} P_n^k u^0(x) &= \sum_{j=0}^{2^n-1} \sum_{\ell=0}^{k} s_{\ell j}^n(i) \phi_{\ell j}^n \left(\frac{2}{\Delta x}(x - x_i)\right), \\ &= 2^{\frac{n}{2}} \sum_{j=0}^{2^n-1} \sum_{\ell=0}^{k} s_{\ell j}^n(i) \phi_\ell \left(2^n \left(\frac{2}{\Delta x}(x - x_i) + 1\right) - 2j - 1\right), \end{aligned}$$

where, using equations (18),

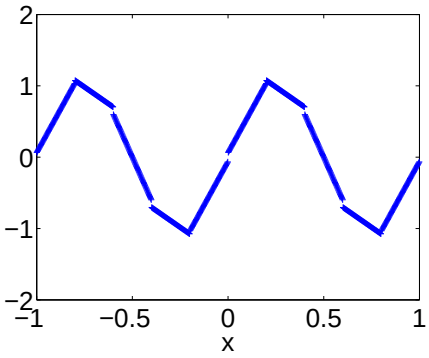
$$\begin{aligned} s_{\ell j}^n(i) = \left\langle u^0, \phi_{\ell j}^n \right\rangle &= \int_{-1}^{1} u^0 \left(x_i + \frac{\Delta x}{2}\xi\right) \phi_{\ell j}^n(\xi) d\xi \\ &= 2^{\frac{n}{2}} \int_{-1+2^{-n+1}j}^{-1+2^{-n+1}(j+1)} u^0 \left(x_i + \frac{\Delta x}{2}\xi\right) \phi_\ell(2^n(\xi+1) - 2j - 1) d\xi \\ &= 2^{-\frac{n}{2}} \int_{-1}^{1} u^0 \left(x_i + \frac{\Delta x}{2}\left(-1 + 2^{-n}(y + 2j + 1)\right)\right) \phi_\ell(y) dy. \end{aligned} \tag{94}$$

In Figure 63, the multiwavelet decomposition of the function $u^0(x) = \sin(2\pi x)$ can be seen. We use 10 elements ($N + 1 = 10$), and 8 subintervals ($n = 3, 2^n = 8$).

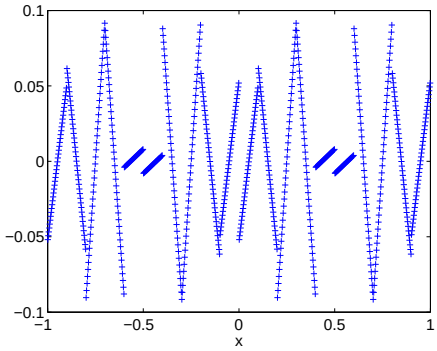
In practice, the DG approximation is used for the time stepping. For the multiwavelet decomposition of DG approximation, it holds that,

$$\begin{aligned} s^0_{\ell 0}(i) &= \int_{-1}^{1} u_h(x_i + \tfrac{\Delta x}{2} y)\phi_\ell(y)dy \\ &= \int_{-1}^{1} \left(\sum_{m=0}^{k} u_i^{(m)} \phi_m \left(\tfrac{2}{\Delta x} \left(x_i + \tfrac{\Delta x}{2} y \right) - x_i \right) \right) \phi_\ell(y)dy \\ &= \sum_{m=0}^{k} u_i^{(m)} \int_{-1}^{1} \phi_m(y)\phi_\ell(y)dy \\ &= u_i^{(\ell)}, \end{aligned}$$

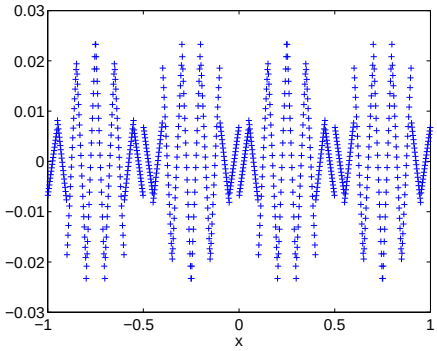
$\ell = 0, \ldots, k$. Therefore, the multiwavelet decomposition equals the DG approximation: all information the DG approximation contains, is found in the scaling function part, such that $d^m_{\ell j}(i) = 0, \forall \ell = 0, \ldots, k, m = 0, \ldots, n-1, j = 0, \ldots, 2^m - 1$, and $i = 0, \ldots, N$. Therefore, it is not possible to decompose the DG approximation into different levels of accuracy (see Figure 64 for the decomposition of the DG approximation, $u_h(x, 0)$, of $u^0(x) = \sin(2\pi x)$). Recall that this is possible for the approach of Section 1.3 and Appendix B.



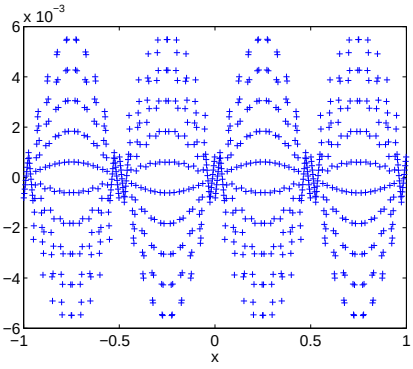
(a) $\sum_{\ell=0}^{k} s_{\ell 0}^{0}(i)\phi_{\ell}(\xi)$



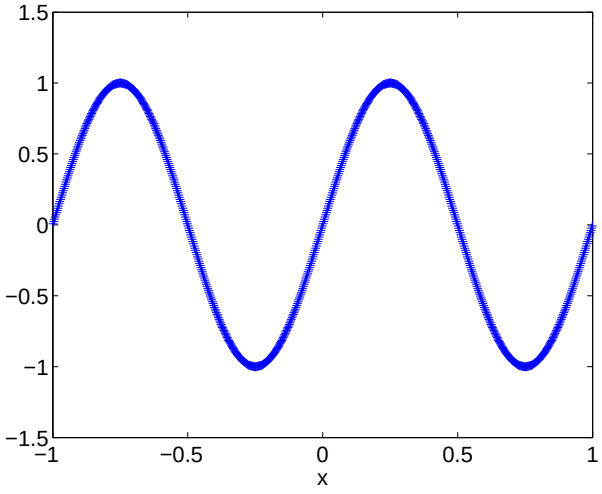
(b) $\sum_{\ell=0}^{k} d_{\ell 0}^{0}(i)\psi_{\ell 0}^{0}(\xi)$



(c) $\sum_{j=0}^{1}\sum_{\ell=0}^{k} d_{\ell j}^{1}(i)\psi_{\ell j}^{1}(\xi)$

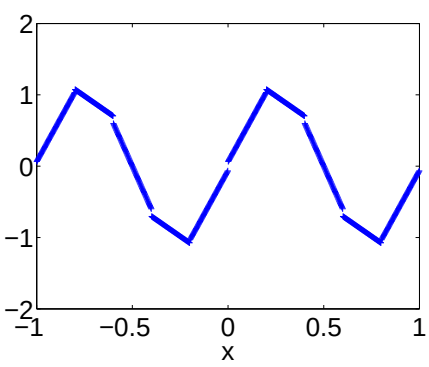


(d) $\sum_{j=0}^{3}\sum_{\ell=0}^{k} d_{\ell j}^{2}(i)\psi_{\ell j}^{2}(\xi)$

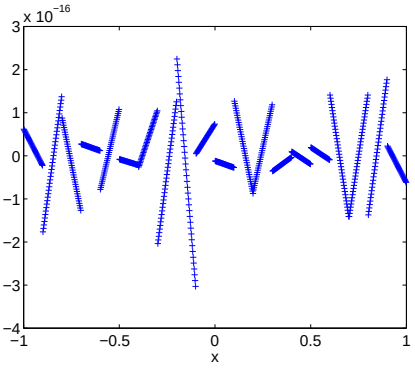


(e) Approximation, $\sum_{\ell=0}^{k} s_{\ell 0}^{0}(i)\phi_{\ell}(\xi) + \sum_{m=0}^{2}\sum_{j=0}^{2^m-1}\sum_{\ell=0}^{k} d_{\ell j}^{m}(i)\psi_{\ell j}^{m}(\xi)$

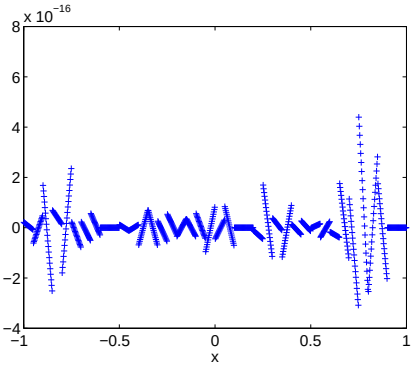
Figure 63: Decomposition procedure, $u^0(x) = \sin(2\pi x), k = 1, n = 3, N + 1 = 10$



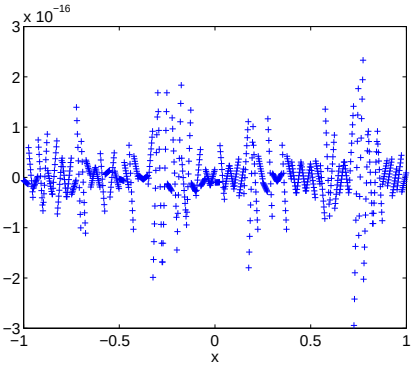
(a) $\sum_{\ell=0}^{k} s_{\ell 0}^{0}(i)\phi_{\ell}(\xi)$



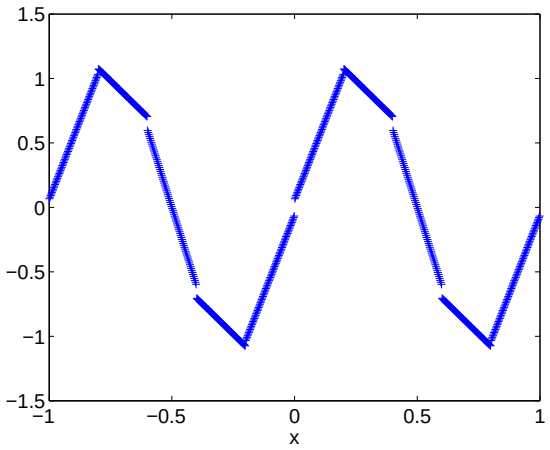
(b) $\sum_{\ell=0}^{k} d_{\ell 0}^{0}(i)\psi_{\ell 0}^{0}(\xi)$



(c) $\sum_{j=0}^{1}\sum_{\ell=0}^{k} d_{\ell j}^{1}(i)\psi_{\ell j}^{1}(\xi)$



(d) $\sum_{j=0}^{3}\sum_{\ell=0}^{k} d_{\ell j}^{2}(i)\psi_{\ell j}^{2}(\xi)$



(e) Approximation

Figure 64: Decomposition of DG approximation, $u_h(x,0)$, of $u^0(x) = \sin(2\pi x)$,$k = 1$, $n = 3, N+1 = 10$

E Counter example for statement (47) in Section 2.3.1

In this appendix, a counter example is given for statement (47) in Section 2.3.1.
For any function $f \in C(-1,1)$, the multiwavelet expansion on level $n \in \mathbb{N}$, with $k = 0$, is given by,

$$P_n^0 f(x) = s_{00}^0 \phi_0(x) + \sum_{m=0}^{n-1} \sum_{j=0}^{2^m-1} d_{0j}^m \psi_{0j}^m(x).$$

Assume that $d_{00}^0 = 0$. If relation (47) holds, then $d_{00}^1 = d_{01}^1 = 0$. Using definition (20), and Table 12,

$$d_{00}^0 = \int_{-1}^{1} f(x)\psi_0(x)dx = -\sqrt{\frac{1}{2}} \int_{-1}^{0} f(x)dx + \sqrt{\frac{1}{2}} \int_{0}^{1} f(x)dx = 0,$$

such that $d_{00}^0 = 0$ because $\int_{-1}^{0} f(x)dx = \int_{0}^{1} f(x)dx$.
For level 1,

$$d_{00}^1 = \int_{-1}^{0} f(x)\psi_{00}^1(x)dx = -\sqrt{\frac{1}{2}} \int_{-1}^{-\frac{1}{2}} f(x)dx + \sqrt{\frac{1}{2}} \int_{-\frac{1}{2}}^{0} f(x)dx,$$

which is only zero if $\int_{-1}^{-\frac{1}{2}} f(x)dx = \int_{-\frac{1}{2}}^{0} f(x)dx$.
But this does not depend on the value of d_{00}^0, and besides that, it is not automatically satisfied. Take for instance the function $f(x) = x^2$. Indeed, $d_{00}^0 = 0$, but,

$$d_{00}^1 = -\sqrt{\frac{1}{2}} \int_{-1}^{-\frac{1}{2}} x^2 dx + \sqrt{\frac{1}{2}} \int_{-\frac{1}{2}}^{0} x^2 dx = -\frac{3}{12}\sqrt{\frac{1}{2}}!$$

This counter example confirms that statement (47) is not true in general.