Aachen Institute for Advanced Study in Computational Engineering Science

Preprint: AICES-2010/11-1

22/November/2010

Analytical and numerical investigation of the influence of artificial viscosity in Discontinuous Galerkin methods on an adjoint-based error estimator

J. Schütz, G. May, S. Noelle

RWTH AACHEN UNIVERSITY

AICES

Financial support from the Deutsche Forschungsgemeinschaft (German Research Association) through grant GSC 111 is gratefully acknowledged.

List of AICES technical reports: http://www.aices.rwth-aachen.de/preprints

Analytical and numerical investigation of the influence of artificial viscosity in Discontinuous Galerkin methods on an adjoint-based error estimator

Jochen Schütz, Georg May, Sebastian Noelle

November 22, 2010

1 Introduction

1.1 Background

Recent years have seen tremendous progress in the numerical computation of hyperbolic conservation laws, both for the steady and unsteady case. Nevertheless, there is an ever increasing need for more complex simulations, which is always higher than the computational power available.
This is the reason for the development of error control in the context of numerical computations to distribute degrees of freedom as economically as possible.
Traditionally, e.g. in the context of elliptic equations, one has always tried to reduce the solution error in some given norm below some given treshold, i.e. the approach *take as little degrees of freedom as possible to achieve* $\|w - w_h\| < \varepsilon$, where w is the exact and w_h the approximated solution (see e.g. [1] for a good overview).
In engineering applications, however, one is often interested in only one or a few single numbers coming from the solution w, e.g. in aerodynamics these numbers could be lift or drag. Mathematically speaking, one is interested in $J(w)$, where J is a given, probably nonlinear functional (usually called *target functional*). This has motivated the aproach *take as little degrees of freedom as possible to achieve* $|J(w) - J(w_h)| < \varepsilon$.
Formally, one can split the error as

$$J(w) - J(w_h) = J'(w_h)(w - w_h) + O(\|w - w_h\|^2). \tag{1.1}$$

Assuming now that w_h is sufficiently close to w, one can neglect the last term, while the action of the derivative on $w - w_h$ can be (usually only approximately) calculated by the adjoint method. This yiels in principle an error estimator.
A lot of work has been put into the development and theoretical justification of adjoint methods supposed that the underlying solution is smooth ([7], [12]). Despite its importance for the use in the context of hyperbolic conservation laws, only a few publications have been concerned with the correct adjoint formulation in the case where there is a jump discontinuity in the underlying forward solution, see e.g. [5], [13], [22] and the references therein.

Recently, it has been observed that the discrete approximation to the dual can actually fail ([11], [13]) in the case of a shock in the underlying forward solution. The key observation of the authors of the aforementioned papers is that the loss of information within the discontinuity is too much to allow for giving precise initial data of the gradient of the objective function. It is in their setting possible to generate for every giving target functional a counterexample where the discrete adjoint approximation does not converge. This is not only true for discrete adjoint approximation, but also for continuous adjoint, because, ultimately, also continuous adjoint does rely on the information given by the approximate forward solution.
To circumvent this feature, the authors in [13] have proposed (and, for their very special setting, also proved) that an over-refinement of the shock-structure, i.e. giving viscosity of $O(h^\alpha)$, where h is the mesh-size and $\alpha < 1$, does lead to convergent adjoint equations.
We have been conducting experiments and came to the conclusion that this does also hold for the case when one approximates the one-dimensional steady-state Euler equation with a DG method - stabilized by explicit artificial diffusion - so, as expected, it is not a particular feature of the approach of [11] and [13]. We will incorporate these insights into our (primal and adjoint) discretization method.
Approximation if the underlying solution has a jump discontinuity is not only difficult in the case of the adjoint approach, but already in the context of the solution w. A lot of research has been put into the development of methods that were able to reliably capture the discontinuity (see e.g. [14], [18]). In the context of Finite Volume methods these are the so-called monotonic methods. Well-known Godunov theorem ([15]) states that no linear high-order scheme can be monotonic, i.e even for linear problems, one has to introduce non-linear shifts in the solution (e.g. flux-limiters or solution-dependent diffusion) to approximate the solution with a high-order method without getting spurious oscillations.
This work focusses around the use of asymptotic adjoint consistent high-order discontinuous Galerkin methods, stabilized by artificial diffusion, made adaptive by the adjoint approach. We are interested in how far the above-mentioned results can be translated to the case of steady-state Euler equations, and how adaptive meshing can be improved to guarantee both convergent adjoints and a highly accurate approximation of some user-defined functional.

1.2 Underlying primal equation: The quasi one-dimensional Euler equations

We consider steady Euler nozzle flow through a convergent-divergent duct of given geometry $A(x)$ ($x \in \Omega \subset \mathbb{R}$), where A is sufficiently smooth (at least $C^1(\Omega)$). The underlying equations are

$$f(w)_x + S(w) = 0 \qquad \forall x \in \Omega \tag{1.2}$$

$$Bw = g \qquad \forall x \in \partial\Omega \tag{1.3}$$

thereby denoting

$$w = (\rho, \rho u, E)^T \tag{1.4}$$

$$f = (\rho u, \rho u^2 + p, u(E+p))^T \tag{1.5}$$

$$S = \frac{A'}{A}(\rho u, \rho u^2, u(E+p))^T \tag{1.6}$$

with the usual meaning of ρ as density, u as velocity, p as pressure and E as energy.
B is a suitable boundary operator, e.g. prescribing Entropy and Enthalpy on the inflow and pressure on the outflow boundary and the equation is closed by using the equation of state for a polytropic gas in the form

$$p := (\gamma - 1)(E - 0.5\rho u^2) \tag{1.7}$$

. If the flow turns out to be supersonic, which can always be achieved by taking appropriate boundary conditions, and the exit pressure is not high enough to allow for an isentropic flow, we will have a discontinuous solution w with a single jump discontinuity at $x = \alpha$.
We chose (1.2)-(1.3) as a test problem, because there is an explicit solution w known, cf. [2].

2 Linearization and dual equation

As we have seen in (1.1), we are mainly interested in variations, i.e. (directional) gradients of the functional

$$J(w) := \int_\Omega p(w) dx. \tag{2.1}$$

We will for simplicity in this section only deal with the one-dimensional model problem (for more sophisticated cases, we refer to [5] and the references therein). We consider the linearization of the *augmented* functional (which we for simplicity also call J)

$$J(w) = \int_\Omega p(w)\ dx - v_\alpha^T \cdot [f(w)] \tag{2.2}$$

with $v_\alpha \in \mathbb{R}^3$ denoting a yet to be determined parameter and $[f(w)]$ denoting the jump in $f(w)$ at position $x = \alpha$.
Note that, for a solution w to (1.2) we have, due to Rankine-Hugeniot-Condition ([18]), that

$$J(w) = \int_\Omega p(w)\ dx \quad \forall v_\alpha \in \mathbb{R}^3, \tag{2.3}$$

which is the functional we are actually interested in. (This explains why we also called it J).
The linearization of J in a direction $\overline{w}$ that does not introduce perturbations at the boundary can be written as

$$J'(w)\overline{w} = \int_\Omega p'(w)\overline{w}\ dx - \overline{\alpha}[p(w)] - v_\alpha^T[f'(w)\overline{w}] - v_\alpha^T\overline{\alpha}[\frac{df(w)}{dx}] \tag{2.4}$$

with $\overline{\alpha}$ being the displacement of the shock position. Note that the first two terms stem from the linearized functional, and the last two terms from the linearized Rankine Hugeniot Condition.
Our goal is now to find a v such that

$$J'(w)\overline{w} = \int_\Omega v^T \cdot R(w)\ dx \tag{2.5}$$

where

$$R(w) := (f'(w)\overline{w})_x + S'(w)\overline{w}. \tag{2.6}$$

We therefore calculate

$$\int_\Omega v^T \cdot ((f'(w)\overline{w})_x + S'(w)\overline{w}) \ dx = -\int_\Omega \left(f'(w)^T v_x - S'(w)^T v\right)^T \overline{w} \ dx \tag{2.7}$$

$$+ \left[v^T f'(w)\overline{w}\right]_{x=0}^{x=1} - [v^T f'(w)\overline{w}]_{x=\alpha-}^{x=\alpha^+}. \tag{2.8}$$

If v is supposed to fulfill the dual equation

$$-f'(w)^T v_x + S'(w)^T v = p'(w) \qquad \forall x \in \Omega \tag{2.9}$$

$$v^T f'(w)\xi = 0 \qquad \forall x \in \partial\Omega, \forall \xi \in Ker(B'(w)), \tag{2.10}$$

then (2.7)-(2.8) simplifies to

$$\int_\Omega v^T \cdot ((f'(w)\overline{w})_x + S'(w)\overline{w}) \ dx = \int_\Omega p'(w)\overline{w} \ dx - [v^T f'(w)\overline{w}]_{x=\alpha-}^{x=\alpha^+}. \tag{2.11}$$

Plugging this into (2.4), we end up with

$$J'(w)\overline{w} = \int_\Omega v^T R(w) - \overline{\alpha}[p(w)] + [v^T f'(w)\overline{w}]_{x=\alpha-}^{x=\alpha^+} - v_\alpha^T[f'(w)\overline{w}] - v_\alpha^T\overline{\alpha}[\frac{df(w)}{dx}] \tag{2.12}$$

If v is now supposed to be continuous with the value $v(\alpha) = v_\alpha$, then terms 3 and 4 cancel.
If we furthermore define v_α to be such that

$$v_\alpha^T[f(w)_x] = -[p(w)] \tag{2.13}$$

at the shock location $x = \alpha$, then, because terms 2 and 5 in (2.12) cancel, we end up with the desired result (2.5).
For the Euler equations (1.2) one can explicitely calculate (2.13) as follows:
Due to the underlying equation (1.2), we have that

$$[f(w)_x] = -[S(w)] \tag{2.14}$$

and due to Rankine-Hugeniot, we have that

$$[f(w)] = 0 \tag{2.15}$$

which yields

$$[S(w)] = \frac{A'(\alpha)}{A(\alpha)}(0, [\rho u^2], 0). \tag{2.16}$$

Plugging all this information into (2.13), we get

$$v_2(\alpha)\frac{-A'(\alpha)}{A(\alpha)}[\rho u^2] = -[p(w)] \tag{2.17}$$

which yiels

$$v_2(\alpha) = \frac{A(\alpha)}{A'(\alpha)}\frac{[p(w)]}{[\rho u^2]}. \tag{2.18}$$

Again, thanks to Rankine-Hugeniot, we have

$$[\rho u^2] = -[p(w)] \tag{2.19}$$

which in all yields the internal adjoint boundary condition for the Euler equations,

$$v_2(\alpha) = -\frac{A(\alpha)}{A'(\alpha)}. \tag{2.20}$$

Note that (2.9) is a non-standard pde as it involves, at $x = \alpha$, the non-conservative product of a discontinuous function with v_x which may be only a measure. Special care has to be taken in defining the solution to this equation, we refer for further reading to ([21],[8])
Of course the internal boundary condition (2.20) is awkward in actual numerical calculations for several reasons, one being the uncertainty about α. This is why most authors do not at all consider this internal boundary condition but just calculate a solution to (2.9)-(2.10) in the hope of reaching (2.20) for free. This has been very successful in the context of low-order methods (cf. [12], [23], [24]), although Giles published, for the time-dependent scalar case, a simple counter example (cf. [11]).
We want to investigate in this work whether the use of artificial viscosity in combination with an adjoint procedure does enforce the boundary condition, and if not, whether it will make any difference in the course of the adaptation procedure.

3 Discretization

We discretize equation (1.2) by using a Discontinuous Galerkin method (cf. [9], [3]). In the last few years, these methods have gained quite a lot of attention, as they bring together the advantages of Finite Element methods (the built-in high-order) and the advantages of Finite Volume methods (the stability due to upwinding). Furthermore, even in the very high order context, DG methods stay local and degrees of freedoms between cells are only coupled over the cell boundary. This makes DG very well suited for parallel computations.
Nevertheless, due to Godunov's order barrier (cf. [15]), a high-order approximation has to be stabilized. Until now, this has been mostly done by limiting which is by now accepted as a standard method of supressing spurious oscillations ([14]).
Recently, several authors (e.g. [20], [10]) have suggested to explicitely add artificial diffusion terms into the method, as the viscous solution does enjoy better regularity properties. Of course the amount of viscosity added has to depend (nonlinearly) on the approximate solution w_h, it has to be high in the case of a discontinuity, and it has to be near by 0 in the case of a smooth solution.

The artificial viscosity approach means that we do actually approximate the *viscous equation*

$$\nabla \cdot f(w_\varepsilon) + S(w_\varepsilon) = \nabla \cdot (\varepsilon(w)\nabla w_\varepsilon) \qquad \forall x \in \Omega \tag{3.1}$$

$$Bw_\varepsilon = g \qquad \forall x \in \partial\Omega \tag{3.2}$$

Let us assume we are given a triangulation $\{\Omega_k\}_{k=1}^N$ which covers the whole domain Ω, i.e. $\bigcup_{k=1}^N \Omega_k = \Omega$. By Γ, we denote the set of edges, i.e. $\Gamma = \{\partial\Omega_k\}$, while by Γ_0 we denote those edges that lie in the interior of the domain Ω.
We can then define our ansatz-space

$$\Pi^p_{\text{dis}} := \{f \in L^2(\Omega) | f_{|\Omega_k} \in \Pi(\Omega_k)^p\} \tag{3.3}$$

where $\Pi(\Omega_k)^p$ is the space of vector-valued polynomials up to degree p.
We follow - for the convective term - the approach of Cockburn and Shu [9], while for the viscous term, we use a Bassi-Rebay 2 discretization [6].
The discretization procedure leaves us with a non-linear equation of the form

$$N_\varepsilon(w_h, v_h) = 0 \qquad \forall v_h \in \Pi^p_{\text{dis}}, \tag{3.4}$$

where we can write N_ε as

$$N_\varepsilon(w_h, v_h) = \sum_{k=1}^N \left(\int_{\Omega_k} -f(w_h) \cdot \nabla v_h + S(w_h) \cdot v_h \; dx + \int_{\partial\Omega_k} v_h^- g(w^+, w^-, n) \; d\sigma(x) \right) \tag{3.5}$$

$$+ B_\varepsilon(w_h, v_h). \tag{3.6}$$

g does denote any numerical flux function known from the Finite Volume community, e.g. a Lax-Friedrichs flux, a Roe flux etc. At the boundary, g is allowed (and in the actual implementation will be) different. $B_\varepsilon(w_h, v_h)$ is the viscous discretization, here given in the unified form of [3]

$$B_\varepsilon(w_h, v_h) = \sum_{k=1}^N \left(\int_{\Omega_k} \nabla_h w_h \cdot \varepsilon \nabla_h v_h dx \right) \tag{3.7}$$

$$+ \int_\Gamma [\![\widetilde{w} - w_h]\!] \cdot \{\varepsilon \nabla_h v_h\} - \{\widetilde{\sigma}\}[\![v_h]\!] \; d\sigma(x) \tag{3.8}$$

$$+ \int_{\Gamma_0} \{\widetilde{w} - w_h\}[\![\varepsilon \nabla_h v_h]\!] - [\![\widetilde{\sigma}]\!]\{v_h\} \; d\sigma(x). \tag{3.9}$$

$\widetilde{w} \equiv \widetilde{w}(w_h)$ and $\widetilde{\sigma} \equiv \widetilde{\sigma}(w_h)$ are approximations for w and $\varepsilon\nabla w$, respectively, on the boundary. Standard assumptions on these approximations are

- *consistency*, which means that in the case of a smooth function w, $\widetilde{w}(w)$ reduces to w on the edge, and $\widetilde{\sigma}(w)$ reduces to $\varepsilon\nabla w$ on the edge.
- *conservativity*, which means that both $\widetilde{w}$ and $\widetilde{\sigma}$ are single-valued on the edges.

See again [3] for details.

Artificial viscosity model implemented If one is interested in calculating solutions with discontinuities, one has to account for a suitable stabilization technique. We want to investigate the artificial viscosity approach.
In the one-dimensional case, we will for simplicity and demonstration purposes only deal with the constant viscosity approach.
Nevertheless, in 'real-world' applications, the viscosity coefficient in $\varepsilon(w)$ in (3.1) has to be chosen - similarly to a limiter function - such that it is only active in regions of non-smoothness of the solution. Furthermore, it should be local in order not to destroy the compactness of the discontinuous Galerkin method.
In [20], the authors have proposed to take a sensor based on the Legendre coefficients of the solution and then calculate an $\varepsilon(w)$ out of that. They have demonstrated very nice looking results, with shocks that lie inside one cell.
Though it is possible, it is not trivial to extend this approach to two or three dimensions. This is why we favour another approach and made experiments - in the spirit of Hartmann [16] - with ε defined as

$$\varepsilon(w) = \frac{\delta \cdot h^{\alpha}}{|\partial K|} \int_{\partial K} \nabla \cdot \mathbb{F}(w) d\sigma \tag{3.10}$$

with user-defined parameters δ and α, which are of course dependent on the polynomial order p implemented.
We know that for smooth parts of the solution, $\varepsilon(w) = O(h^{\alpha+p+1})$, while for discontinuities, $\varepsilon(w) = O(h^{\alpha})$. This makes ε suitable for the use in an artificial viscosity model.
Nevertheless, as this paper is purely one-dimensional, from now on we will assume ε to be constant. We expect the experiments we made to be very similar in two dimensions, but, due to limited ressources, the experiments could not be performed as these experiments needs a very high mesh resolution and rely on an advanced-known solution w.

Consistency and (Asymptotic) Adjoint Consistency Our discretization N_{ε} is called consistent iff for the exact solution w of (1.2) holds

$$N_{\varepsilon}(w, v_h) = 0 \qquad \forall v_h \in \Pi^{\infty}_{dis}. \tag{3.11}$$

Consistency is a quite natural concept which is necessary (but of course not sufficient) for the discretization to converge towards the solution w and has been acknowledged for a long time.
A rather new concept is the idea of (asymptotic) adjoint consistency (cf. [3], [17] and [19] where asymptotic adjoint consistency is probably first defined). This kind of consistency states that the discrete adjoint approximation - to be defined soon - is a consistent approximation of the exact adjoint solution. This has a lot of implications, e.g. the achievement of optimal L^2 order (via the *Nitsche-Trick*, cf. [3]) and the superconvergence of a certain class of (nonlinear) functionals (cf. [17]). Furthermore, by using adjoint consistent approximations, one is completely free in choosing whether to take the discrete or the continuous adjoint approach, both having their own advantages.
Let us define what we mean by adjoint consistency:

3.1 Definition. The discrete adjoint equations to (3.4) are

$$N'_\varepsilon(w_h)(z_h, v_h) - J'(w_h)z_h = 0 \qquad \forall z_h \in \Pi^{\overline{p}}_{dis}. \tag{3.12}$$

The discretization N_ε is said to be adjoint consistent given that for the exact adjoint solution v and exact primal solution w holds

$$N'_\varepsilon(w)(z_h, v) - J'(w)z_h = 0 \qquad \forall z_h \in \Pi^{\infty}_{dis}, \tag{3.13}$$

while it is asymptotic adjoint consistent iff

$$\lim_{h\to 0} \sup_{\|z_h\|=1} (N'_\varepsilon(w)(z_h, v) - J'(w)z_h) = 0 \qquad \forall z_h \in \Pi^{\infty}_{dis}. \tag{3.14}$$

holds. h is thereby denoting e.g. the diameter of the grid-cell.

We mainly focus on the asymptotic adjoint consistency concept, because by rather approximating the viscous equation (3.1) than the original equation (1.2), we cannot expect to get more than an asymptotic adjoint behaviour. This is already clear from continuous considerations.
We can show that our discretization is asymptotically adjoint consistent. (For the analysis, please see Appendix A.)
The discrete analogue to (2.5) is

$$J'(w)\overline{w} = N_\varepsilon(w_h, v_h). \tag{3.15}$$

Due to consistency, we know that $N_\varepsilon(w_h, v_h) = 0 \quad \forall v_h \in \Pi^p_{dis}$, so, in general, one should take $\overline{p} > p$ or think about anything like e.g. patch-wise interpolation (cf. [4]).
The adjoint consistency allows us to freely decide whether to take the continuous or the discrete adjoint aproach. We favour the continuous adjoint approach because we need also stabilization in the adjoint problem - unfortunately not at the same location where we would need it in the primal problem. Thus, the continuous adjoint approach gives us more freedom in how to take this dual stabilization.

Adaptation Procedure In our adaptation procedure, we start with a rather coarse grid and calculate a solution w^1_h to (3.4). We then calculate a dual solution v^1_h using the continuous adjoint approach, and calculate an error estimator based on (2.5) or its discrete version (3.15) which drives the adaptation algorithm. We then get a solution w^2_h from which we calculate a dual v^2_h etc...
We calculate solutions as long as the error estimator lies above some threshold or until we reach a prescribed maximum number of degrees of freedom.
To avoid 'extremal' refinement steps (like e.g. refine only one cell, or refine all), we use the so-called *fixed-fraction criterium* ([7]) which refines a fixed fraction q of cells (we usually use $q = \frac{1}{3}$) that have the largest error indicator.

4 Why can it be a problem approximating the adjoint solution?

As we have already mentioned, unless one gives enough diffusion - which means smearing out the shock - one can not expect to get a convergent adjoint solution in the case when a shock is present in the forward

solution.
We will, in a similar way as has been done in [13] motivate that this over-refinement of the shock-structure comes up naturally, and that - if not using it - one can for every given target functional construct a counterexample in which the adjoint solution does not converge.
Let us therefore look at the information the dual algorithm gets from the target functional under consideration: It is just the midvalue linearization $\xi'(u_h)$. In a smooth case, this is approximately the right thing to do, in the non-smooth case, information is irreversibly lost! The solution of the dual equation does depend on the jump in ξ (this can for example be seen in the preceding section, (2.13)). So it is necessary that with the use of $\xi'(u_h)$, the algorithm can 'guess' the jump in ξ.
Now imagine that there is only a fixed number of points within the shock, and, for the sake of simplicity,

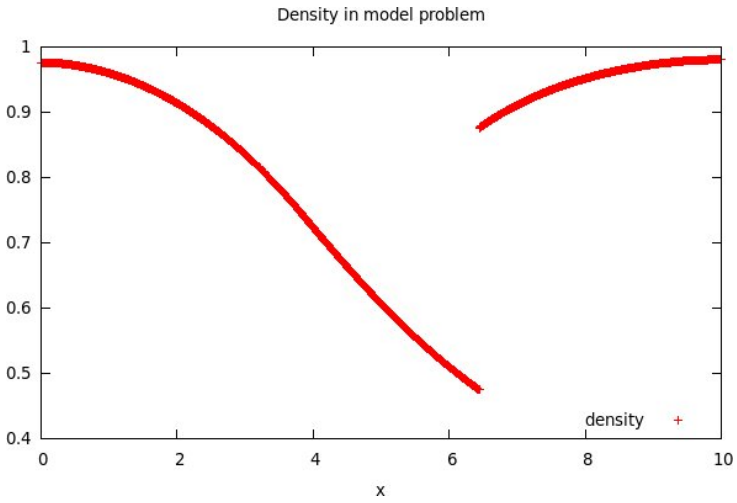

Figure 1:
Density function in our testcase

imagine that pressure does only depend on one variable, say density ρ (which is of course not true at all, but the idea carries over to the multivariate case easily, just that visualization is not that simple any more). From Fig. 1 we can conclude that in our demonstration case, the density has a jump from about 0.45 to 0.9. Fig. 2 plots the density against the pressure for more or less arbitrary values of E and ρu, and it plots another function which has - at the indicated sample points - nearly the same gradient with respect to ρ. Nevertheless, as one can see, it has a completely different jump. Now assume that the indicated number of sample points are not increasing (which is the case if one resolves the shock with viscosity of size $O(h)$) - why should the adjoint procedure know which of the indicated functions is the right one?!
We will demonstrate in the on-going work that there is indeed no linear dependence on the mesh size for the adjoint to converge.
Note that - to make the above example rigorous - one has to consider p as a function of ρ, ρu and E in the domain where ρ and E jump (due to Rankinge Hugeniot, ρu is continuous). At the given sampling points, one has to find a function that has the exact gradient values, while it is different elsewhere and especially has a different jump.

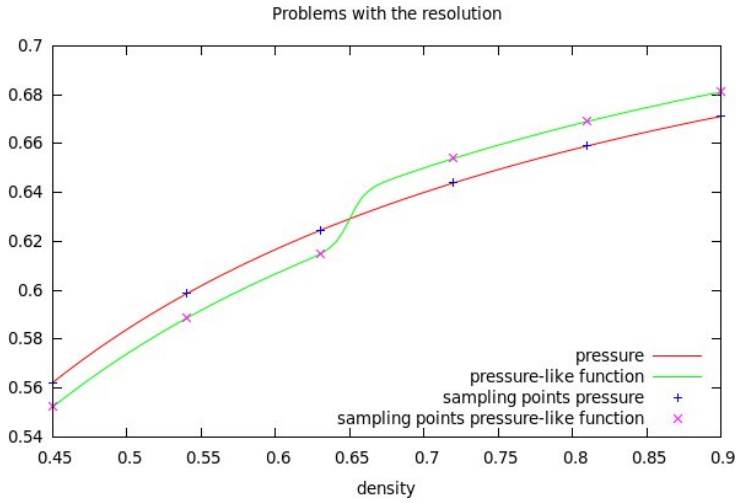

Figure 2:
Two functions with different jumps, but same gradient at indicated sampling points

5 Numerical results

The h^α viscosity approach As mentioned in the introduction, the convergence towards the dual solution can actually fail unless one gives enough diffusion ([11], [13]). We have - in the 1D Euler case - been conducting experiments and have found out that this is also true in our setting.
Let us consider a one-dimensional test-case with a free-stream Machnumber of 0.5 and an exit pressure of $p_{out} = 1.0$. For our smooth geometry, this creates a shock at position $x = 5.25...$ (which can, thanks to the explicit solution, be calculated in advance). Both primal and dual solution are calculated with polynomial order once $p = 1$ and once $p = 3$ on the same mesh.
To demonstrate our findings, we assume that the viscosity $\varepsilon(w) \equiv \varepsilon$ is constant throughout our domain Ω. As a measure of error, we use the relative deviation of the dual solution v in fulfilling (2.20). Figure 3 plots the amount of viscosity used versus the error in the dual solution.
There are some remarks in order about the plot: Notice the right hand side where all the graphs lie above each other. This means that both primal and dual solution are perfectly grid-converged and that the only source of error comes from the viscosity term. The interesting point is always when the graphs 'leaves' the common region (on the left hand side of the plot) because this is where discretization errors begin to dominate the overall error. What can be seen is that these occurences do not appear linearly with respect to the mesh-size but that they eventually occur earlier and earlier.
These findings motivate the conclusion that one shoud take viscosity constant to h^α, where α is slightly less than unity, to ultimately guarantee convergence of the adjoint solution.
We demonstrate the advantage in taking the viscosity proportional to $h^{0.8}$ in Figure 4. When giving viscosity proportional to h, the dual solution does not converge at all, while when adding diffusion as $O(h^{0.8})$, we have a perfectly acceptable convergence history.
What one can also see in comparing the two pictures in Figure 3 is that the higher-order solution needs a little less (about half) the amount of viscosity to accurately resolve the dual solution on the same mesh. It remains to prove these findings similar to [13]. Unfortunately, this will be - due to the more complicated structur of the Euler equations - a non-trivial task as already mentioned in [13].

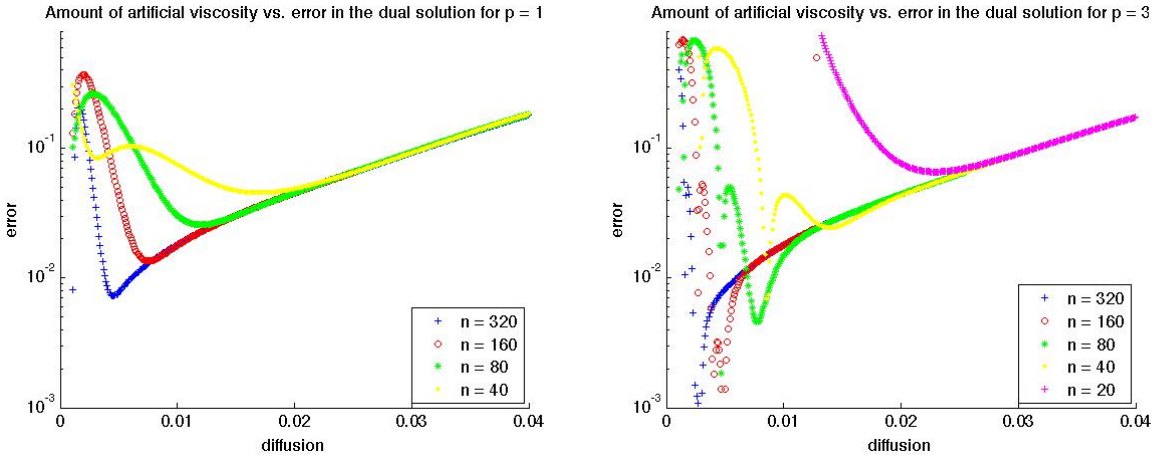

Figure 3:
Amount of artificial viscosity versus error in the adjoint solution for $p = 1$ and $p = 3$.

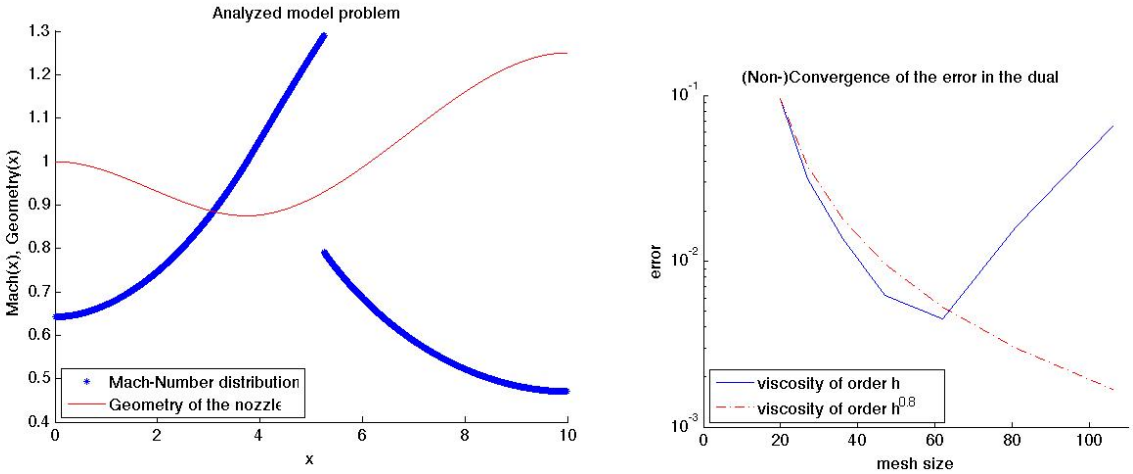

Figure 4:
Plot of the model problem and convergence history of the adjoint solution for $p = 3$.

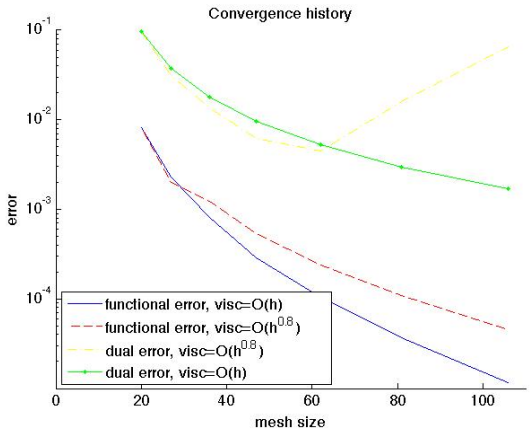

Figure 5:
Convergence history of the adjoint solution for $p = 3$.

Adaptivity by the adjoint without enforcing the h^α approach - a problem? Despite the very negative results in the preceeding subchapter, we want to demonstrate that - for using adaptivity - it does not seem necessary at all to enforce the dual boundary condition.
Let us therefore look at Figure 5. We ran a test with the same starting parameters as the one for Fig. 4, with the difference that we now did not refine uniform, but we did refine where the adjoint error estimation did tell us to. The constant viscosity was chosen to be of size $O(h)$ and $O(h^{0.8})$, respectively, where h is the minimum mesh-size.
Of course we are not primarily interested in whether the adjoint internal boundary condition (2.20) does converge or whether not, but we are interested in how far the primal functional J converges. It can be seen clearly from Fig. 5, that even if the dual does not converge towards the boundary condition (and, in so far, the dual does not converge properly), we still get an adaptation criterion that seems to be very reasonable. Note that, furthermore, due to the viscosity of size $O(h)$, the 'standard' approach what concerns viscosity seems to converge mush faster in terms of the functional J (which is - after all - not very surprising). The actual surprising part is that as an adaptation criterion, the standard approach seems to be as good as the over-refinement approach.
This in some way demonstrates the capability of the adjoint approach even if one does not enforce or even achieve the internal dual boundary condition (2.20). (This is what actually most authors do - just neglect (2.20). Our findings could justify their approach.)

6 Appendix A - Asymptotic Adjoint Consistency Analysis

In this section, we give a brief analysis why our discretization is asymptotic adjoint consistent.
For notational simplicity, we will only consider the one-dimensional case, although all the results can be straightforward extended to the two-dimensional case.
We proceed in two steps: First, we only investigate the part coming from the convection (this part is

actually adjoint consistent), then we investigate the part coming from the artificial diffusion. Asymptotic Consistency Analysis in the literature did always assume that the underlying solutions are all smooth. We also will in this section suppose this (although it does not really fit in the overall concept of the paper).

Adjoint Consistency of the convective part

Recall that the convective part is given as

$$N_{conv}(w_h, v_h) = \sum_{k=1}^{N} \left(\int_{\Omega_k} -f(w_h) \cdot \nabla(v_h) + S(w_h) \cdot v_h \; dx + \int_{\partial\Omega_k} v_h^- g(w^+, w^-, n) \; d\sigma(x) \right). \tag{6.1}$$

At the boundary $\{x_1, x_N\}$, we have to replace w^+ or w^- by some boundary operator $w_b = w_b(w^-)$ or $w_b = w_b(w^+)$, respectively. Furthermore, we underline that g at the boundary does not have to be the same as g in the interior. On the contrary, we have to choose g at the boundary in a very special way to achieve adjoint consistency.

The derivative of N_{conv} with respect to the first argument, evaluated at w, in some direction z then is

$$\begin{aligned}
N'_{conv}(w)(z, v) = &- \int_\Omega (f'(w)z)^T \cdot v_x \; dx + \int_\Omega (S'(w)z)^T \cdot v \; dx && (6.2)\\
&+ \sum_{i=2}^{N-1} \left(\frac{\partial g}{\partial w_1}^T z^+ v^- + \frac{\partial g}{\partial w_2}^T z^- v^- - \frac{\partial g}{\partial w_1}^T z^+ v^+ - \frac{\partial g}{\partial w_2}^T z^- v^+ \right) && (6.3)\\
&+ \frac{\partial g}{\partial w_1}^T z(x_N)^+ v(x_N)^- + \frac{\partial g}{\partial w_2}^T w_b' z(x_N)^- v(x_N)^- && (6.4)\\
&- \frac{\partial g}{\partial w_1}^T w_b' z(x_1)^+ v(x_1)^- - \frac{\partial g}{\partial w_2}^T z(x_1)^- v(x_1)^- && (6.5)\\
&=: I + II + III + IV && (6.6)
\end{aligned}$$

For simplicity, we did opress the argument at g, which is - due to the assumed continuity of w - always (w, w).

Now, we take into account the defining equations for v (see 2.9) and we get (still under the assumption that also v is smooth)

$$I = -\int_\Omega z^T f'(w)^T v_x \; dx + \int_\Omega z^T S'(w)^T v \; dx \overset{(2.9)}{=} \int_\Omega p'(w) z dx \tag{6.7}$$

Furthermore, we have - due to the assumed continuity of v, that $v^+ = v^- \equiv v$ and so

$$II = \sum_{i=2}^{N-1} \left(\frac{\partial g}{\partial w_1} z^+ + \frac{\partial g}{\partial w_2} z^- - \frac{\partial g}{\partial w_1} z^+ - \frac{\partial g}{\partial w_2} z^- \right) z = 0. \tag{6.8}$$

Third and fourth term are analogous, but need some modification in the numerical flux function as we will see:

$$III = (z(x_n)^-)^T \left(\left(\frac{\partial g}{\partial w_1}^T \right) + w_b'^T \left(\frac{\partial g}{\partial w_2} \right)^T \right) v \tag{6.9}$$

Now assume that one has

$$g(a,b) = f(b) \tag{6.10}$$

at the boundary, then one gets that

$$III = 0 \tag{6.11}$$

for *every* choice of w_b', due to the dual boundary condition (2.10) and the fact that (6.10) implies

$$\frac{\partial g}{\partial u_2} = \frac{\partial f}{\partial u}. \tag{6.12}$$

Note however that this does usually not hold when employing a standard Riemannsolver g.
Similar, one gets

$$IV = 0 \tag{6.13}$$

when modifying g at the left boundary such that it just depends on the first argument.
This completes the adjoint consistency analysis for the convective term. We mention that an analogous approach has been carried out in [17] with a similar conclusion.

Adjoint consistency of the diffusive part Recall that the diffusive part is given as in (3.7)-(3.8) and recall the consistency and conservativity conditions.
Assuming that ε does only depend on w, i.e. $\varepsilon \equiv \varepsilon(w)$, the Frechet derivative of (3.7) for fixed v in direction of z reads

$$B_\varepsilon'(w)(z,v) = \int_\Omega \nabla z \cdot \varepsilon \nabla v dx + \int_\Omega \nabla w \varepsilon'(w) z \nabla v dx \tag{6.14}$$

$$+ \int_\Gamma [\![\widetilde{z} - z]\!] \cdot \{\varepsilon \nabla v\} + \int_\Gamma [\![\widetilde{w} - w]\!] \{\varepsilon'(w) z \nabla v\} - \int_\Gamma \{\widetilde{\sigma}'(w) z\} [\![v]\!] d\sigma \tag{6.15}$$

$$+ \int_{\Gamma_0} \{\widetilde{z} - z\} \cdot [\![\varepsilon \nabla v]\!] + \{\widetilde{w} - w\} \cdot [\![\varepsilon'(w) z \nabla v]\!] - [\![\widetilde{\sigma}'(w) z]\!] \{v\} d\sigma \tag{6.16}$$

An integration by parts yields:

$$\int_\Omega \nabla z \cdot \varepsilon \nabla v dx \tag{6.17}$$

$$= -\int_\Omega z \nabla \cdot (\varepsilon \nabla v)\, dx + \int_\Gamma [\![z]\!] \cdot \{\varepsilon \nabla v\} d\sigma + \int_{\Gamma_0} \{z\} \cdot [\![\varepsilon \nabla v]\!] d\sigma \tag{6.18}$$

$$\tag{6.19}$$

which in all yields

$$(6.14) = -\int_\Omega z\nabla\cdot(\varepsilon\nabla v)\,dx + \int_\Omega \nabla w\varepsilon'(w)z\nabla v dx \tag{6.20}$$

$$+\int_\Gamma [\![\widetilde{z}]\!]\cdot\{\varepsilon\nabla v\} + \int_\Gamma [\![\widetilde{w}-w]\!]\{\varepsilon'(w)z\nabla v\} - \int_\Gamma \{\widetilde{\sigma}'(w)z\}[\![v]\!]d\sigma \tag{6.21}$$

$$+\int_{\Gamma_0} \{\widetilde{z}\}\cdot[\![\varepsilon\nabla v]\!] + \{\widetilde{w}-w\}\cdot[\![\varepsilon'(w)z\nabla v]\!] - [\![\widetilde{\sigma}'(w)z]\!]\{v\}d\sigma \tag{6.22}$$

We now *assume* that w and v are sufficiently smooth, which in this case means continuous. Furthermore, we exploit the consistency and conservativity of the fluxes to end up with

$$(6.14) = -\int_\Omega z\nabla\cdot(\varepsilon\nabla v)\,dx + \int_\Omega \nabla w\varepsilon'(w)z\nabla v dx \tag{6.23}$$

We can now exploit Cauchy-Schwartz to get

$$(6.23) \leq \|z\|_{L^2}\cdot\|\varepsilon\nabla v\|_{L^2} + \|\nabla w\|_{L^2}\|\varepsilon'(w)z\nabla v\|_{L^2} \tag{6.24}$$

For fixed $\|z\|$, this then turns to 0 because the ε should go to zero in a reasonably fast way to be a suitable artificial viscosity function.

References

[1] Mark Aisnworth and J. Tinsley Oden. A posteriori error estimation in finite element analysis. *Comput. Methods Appl. Mech. Engrg.*, 142:1–88, 1997.

[2] John D. Anderson. *Fundamentals of Aerodynamics*. McGraw-Hill, 3. edition, 2001.

[3] Douglas N. Arnold, Franco Brezzi, Bernardo Cockburn, and L. Donatella Marini. Unified analysis of discontinuous galerkin methods for elliptic problems. *SIAM J. Numer. Anal*, 39(5):1749–1779, 2002.

[4] Wolfgang Bangerth and Rolf Rannacher. *Adaptive Finite Element Methods for Differential Equations*. Birkhäuser, 2003.

[5] Claude Bardos and Olivier Pironneau. Data assimilation for conservation laws. *Methods and Applications of Analysis*, 12(2):103–134, June 2005.

[6] F. Bassi, S. Rebay, G. Mariotti, S. Pedinotti, and M. Savini. A high-order accurate discontinuous finite element method for inviscid and viscous turbomachinery flows. *Proceedings of 2nd European Conference on Turbomachinery, Fluid Dynamics and Thermodynamics, R. Decuypere and . Dibelius, eds.*, pages 99–108, 1997.

[7] Roland Becker and Rolf Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numerica*, 10:1–102, 2001.

[8] F. Bouchut and F. James. One-dimensional transport equations with discontinuous coefficients. *Nonlinear Analysis, Theory, Methods and Applications*, 32(7):891–933, 1998.

[9] Bernardo Cockburn and Chi-Wang Shu. The runge-kutta local projection p^1-discontinuous-galerkin finite element method for scalar conservation laws. *IMA Preprint*, 388, 1988.

[10] G.Barter and David L. Darmofal. Shock capturing with higher-order, pde-based artificial viscosity. *Proc. of the 18th AIAA CFD Conference*, 2007.

[11] Michael B. Giles. Discrete adjoint approximations with shocks. *Proc. of the 9th. International Conference on Hyperbolic Problems*, 2002.

[12] Michael B. Giles and Niles A. Pierce. An introduction to the adjoint approach in design. *Flow, Turbulence and Combustion*, 65:393–415, 2000.

[13] Mike Giles and Stefan Ulbrich. Convergence of linearised and adjoint approximations for discontinuous solutions of conservation laws. *Technical Report TU Darmstadt and Oxford University*, 2008.

[14] Edwige Godlewski and Pierre-Arnaud Raviart. *Hyperbolic Systems of Conservation Laws*. Ellipses, 1991.

[15] Sergei Godunov. A difference scheme for numerical solution of discontinuous solution of hydrodynamic equations. *Math. Sbornik*, 47:271–306, 1959.

[16] Ralf Hartmann. Adaptive discontinuous galerkin methods with shock-capturing for the compressible navier-stokes equations. *Int. J. Numer. Meth. Fluids*, 51:1131–1156, 2006.

[17] Ralf Hartmann. Adjoint consistency analysis of discontinuous galerkin discretizations. *SIAM J. Numer. Anal*, 45(6):2671–2696, 2007.

[18] Randall J. LeVeque. *Numerical Methods for Conservation Laws*. Birkhäuser, 1990.

[19] Todd A. Oliver and David L. Darmofal. Analysis of dual consistency for discontinuous galerkin discretizations of source terms. *SIAM J. Numer. Anal*, 2008.

[20] Per-Olof Persson and Jaime Peraire. Sub-cell shock capturing for discontinuous galerkin methods. *Proc. of the 44th AIAA Aerospace Sciences Meeting and Exhibit*, 2006.

[21] Eitan Tadmor. Local error estimates for discontinuous solutions of nonlinear hyperbolic equations. *SIAM J. Numer. Anal*, 28(4):891–906, 1991.

[22] Stefan Ulbrich. Optimal control of nonlinear hyperbolic conservation laws with source terms. *Technical Report TU München*, 2001.

[23] David A. Venditti and David L. Darmofal. Adjoint error estimation and grid adaptation for functional outputs: Application to quasi-one-dimensional flow. *Journal of Computational Physics*, (164):204–227, 2000.

[24] David A. Venditti and David L. Darmofal. Grid adaptation for functional outputs: Application to two-dimensional inviscid flows. *Journal of Computational Physics*, (176):40–69, 2002.