RWTH Aachen
MathCCES
Wintersemester 2014/15

CES Seminar WS 2014/15

A new class of discontinuous Petrov-Galerkin methods

Oliver Henrichs
oliver.henrichs@rwth-aachen.de
May 15, 2015

Contents

1 Introduction

I present a discontinuous Petrov-Galerkin (DPG) method to numerically compute solutions to boundary value problems using semi-automatically computed test spaces tailored for optimal stability properties, developed by Leszek Demkowicz and Jay Gopalakrishnan. The corresponding papers are [DEGO1] and [DEGO2]. In both papers the authors formulate the finite element method, while in the latter paper strategies actual computation of optimal test spaces is explained, too.
The first question that arises is, what does discontinuous Petrov-Galerkin mean? Standard Galerkin methods use the same function space for both trial and test space on each element, with both spaces being continuous at least on element level. This method uses a different approach. The method is discontinuous because the trial functions are continuous on element level, but may incorporate discontinuities at element interfaces. It is a Petrov-Galerkin method, because the trial and test spaces are different. Additionally, and in contrast to standard DG-methods, the method introduces flux unkowns at element interfaces. The novelty in this method lies in the incorporation of several approaches, which, individually, have been used before. Discontinuities within single elements have been used before, e.g. in [BELY01], but not solely on test functions to obtain stability; flux unknowns have been e.g. used in hybridized DG methods [GOPA09], to solve first for the interfaces, and later compute the internal variables locally. However, these methods used stabilizing parameters, and p-independent stability is yet to be proven for them. In contrast, the DPG method presented here gains its stability solely from the choice of the test space. The main advantage of the described method lies in its rapid h-convergence and provable p-convergence on very general meshes, which is an improvement to standard DG-methods. Here, p describes the order of the approximating functions, while h determines the size of mesh elements, or, in other words the number of elements in the grid.
To keep the text at a reasonable length, I leave out most proofs, but give references for interested readers. Therefore I restrict myself to the most important lemmas, theorems and subsequently conclusions which can be drawn from them. In the light of this I will cover the development of the discontinuous Petrov-Galerkin method in the one element case detailed, and will, for the multi-element case, briefly cover its equations and error estimates. Following that I will introduce the concept of finding optimal test functions in practice and will conclude with a 1D example case to verify the method.

2 Motivation

First it is to be shown that one can indeed use the test spaces only for stabilization purposes, completely ignoring its approximation properties. The following well-known theorem by Babuška [BABU70] is rewritten in [DEGO1, p.2-3]:
Let X, $X_h \subset X$ and V_h be Banach Spaces in $\mathbb{R}$ and $a_h(\cdot,\cdot)$ be a bilinear form on $X \times V_h$. Suppose the exact solution $U \in X$ as well as the discrete solution $U_h \in X_h$ satisfy

$$a_h(U - U_h, v_h) = 0 \qquad \text{for all } v_h \in V_h. \tag{1}$$

If a_h is continuous, i.e. there is a $C_1 > 0$ such that

$$a_h(w, v_h) \leq C_1 ||w||_X ||v_h||_{V_h} \qquad \text{for all } w \in X,\ v_h \in V_h \tag{2}$$

and if the inf-sup condition, i.e. there is a $C_2 > 0$, such that

$$C_2 ||w_h||_X \leq \sup_{v_h \in V_h} \frac{a_h(w_h, v_h)}{||v_h||_{V_h}} \tag{3}$$

holds, then the following well known error estimate can be proven:

Theorem 1.1.

$$||U - U_h||_X \leq (1 + \frac{C_1}{C_2}) \inf_{w_h \in X_h} ||U_h - w_h||_X. \tag{4}$$

Thus, the error depends on the approximation qualities of the trial space X_h and on the ratio of the continuity and inf-sup constants $\frac{C_1}{C_2}$. This means that in the process of designing test and trial function spaces the trial functions must be designed such that $||U_h - w_h||_X$ is minimized, while the test functions only appear implicitely in the stability constant $1 + \frac{C_1}{C_2}$. This stability constant will be minimized by the test spaces that have been developed in the papers on which this seminar work bases.

3 The discontinuous Petrov-Galerkin method

The method developed in [DEGO1] is explained by discussing the solution of the following problem: Given a polyhedral domain Ω as well as functions f and g, find a finite element approximation to the solution u of

$$\vec{\beta} \cdot \vec{\nabla} u = f \qquad \text{in } \Omega \tag{5}$$

$$u = g \qquad \text{on } \partial_{\text{in}}\Omega \tag{6}$$

where $\vec{\beta}$ is the (constant) flow direction and $\partial_{\text{in}}\Omega$ is the corresponding global inflow boundary defined by

$$\partial_{\text{in}}\ \Omega = \{\vec{x} \in \partial\Omega : \vec{\beta} \cdot \vec{n}(\vec{x}) < 0\}. \tag{7}$$

The general ideas and study of element spaces is best done on a simple example. In this section $\Omega = K$ denotes a single triangle or a single line between two points, with in- and outflow boundaries

$$\partial_{out} K = \{\vec{x} \in \partial K : \vec{\beta} \cdot \vec{n} > 0\}, \tag{8}$$

$$\partial_{in} K = \{\vec{x} \in \partial K : \vec{\beta} \cdot \vec{n} < 0\}, \tag{9}$$

where $\vec{n}$ denotes the unit outward normal. The following section first introduces the DPG method, shows that it has a unique solution and finds error estimates to proof its stability and convergence.

3.1 Coordinate system alignment

The first thing that needs to be done is to introduce a new coordinate system. It is aligned with the flow direction $\vec{e}_\beta$, such that one can discriminate between the flow direction and all other directions. This enables us to easily figure out functions that are constant along flow direction, which comes in handy later. Thus in [DEGO1] an orthonormal completion of $\vec{e}_\beta$ is introduced, such that for any $\vec{x}$

$$\vec{x} = x_1 \vec{e}^{(1)} + ... + x_N \vec{e}^{(N)} \tag{10}$$

$$= \eta_1 \vec{e}_\beta + \eta_2 \vec{e}_\beta^{(2)} + ... + \eta_N \vec{e}_\beta^{(N)} \tag{11}$$

Now any function ζ defined on $\partial_{\text{out}} K$ can be rewritten as $\zeta(\eta_2, \eta_3, \cdots, \eta_N)$ independent of η_1 and subsequently be extended from $\partial_{\text{out}} K$ into K such that it is constant along the $\vec{\beta}$-direction. This extension from the outward boundary is called ε_{out}, e.g.

$$\varepsilon_{\text{out}} \zeta(\eta_1, \eta_2, \cdots, \eta_N) = \zeta(\eta_2, \eta_3, \cdots, \eta_N)$$

for all η_1.

3.2 The discontinuous Petrov-Galerkin method on one element

The discontinuous Petrov-Galerkin method on one element computes a polynomial approximation to the solution u of the transport equation (5) with $\Omega = K$. Starting point of all analysis is multiplying (5) by a test function v and integrate it by parts:

$$\begin{aligned}(f, v) &= (\vec{\beta} \cdot \vec{\nabla} u, v)_K \\ &= -(u, \vec{\beta} \cdot \vec{\nabla} v)_K + \langle \vec{\beta} \cdot \vec{n} u, v \rangle_{\partial_K} \\ &= -(u, \vec{\beta} \cdot \vec{\nabla} v)_K + \langle \Phi, v \rangle_{\partial_{\text{out}} K} + \langle \vec{\beta} \cdot \vec{n} g, v \rangle_{\partial_{\text{in}} K}\end{aligned} \tag{12}$$

where $\Phi = \vec{\beta} \cdot \vec{n} u$ is the new unknown outflux variable on $\partial_{\text{out}} K$. Now the DPG method approximates the pair (u, Φ) on the spaces

$$M_p(\partial_{\text{out}} K) = \{\mu : \mu|_F \in P_p(F), \text{ for all faces } F \text{ of } K \text{ contained in } \partial_{\text{out}} K\}, \tag{13}$$

$$V_p(K) = \eta_1 P_p(K) + \varepsilon_{\text{out}}(M_{p+1}(\partial_{\text{out}} K)). \tag{14}$$

where P_p are polynomials of degree p on their respective domains and

$$\eta_1 P_p(K) = \{\eta_1 p_p : p_p \in P_p(K)\}.$$

Note that the trial functions of the faces F are one order higher than the trial functions within the elements. This choice of spaces is motivated by three arguments which will ultimately guarantee the solvability of the method. The first argument is, that the two constituent spaces of $V_p(K)$, namely $\eta_1 P_p(K)$ and $\varepsilon_{\text{out}}(M_{p+1}(\partial_{\text{out}} K))$, are linearly independent. The second argument claims that $V_p(K)$ can be written as the direct sum of both function spaces. The third argument is, that for any $\mu \in M_{p+1}(\partial_{\text{out}} K)$ and any w in $P_p(K)$, there is a unique $v \in V_p(K)$ satisfying

$$\vec{\beta} \cdot \vec{\nabla} v = w, \tag{15}$$

$$v|_{\partial_{\text{out}} K} = \mu. \tag{16}$$

Thus, one can always design a test function $v \in V_p(K)$ to fit desired properties of both trial functions u_p and Φ_{p+1}. The interested reader finds proofs for the three arguments in [DEGO1, p.6]. Next, the spectral DPG method can be introduced. It defines approximations $(u_p, \Phi_{p+1}) \in P_p(K) \times M_{p+1}(\partial_{\text{out}} K)$, satisfying

$$-(u_p, \vec{\beta} \cdot \vec{\nabla} v)_K + \langle \Phi_{p+1}, v \rangle_{\partial_{\text{out}} K} = (f, v)_K - \langle \vec{\beta} \cdot \vec{n} g, v \rangle_{\partial_{\text{in}} K} \tag{17}$$

for all $v \in V_p(K)$. Defining

$$a((u_p, \Phi_{p+1}), v) \overset{\text{def}}{=} -(u_p, \vec{\beta} \cdot \vec{\nabla} v)_K + \langle \Phi_{p+1}, v \rangle_{\partial_{\text{out}} K} \tag{18}$$

$$b(v) \overset{\text{def}}{=} (f, v)_K - \langle \vec{\beta} \cdot \vec{n} g, v \rangle_{\partial_{\text{in}} K} \tag{19}$$

$$X_p(K) \overset{\text{def}}{=} P_p(K) \times M_{p+1}(\partial_{\text{out}} K), \tag{20}$$

one can rewrite equation (17) into: Find $(u_p, \Phi_{p+1}) \in X_p(K)$, such that

$$a((u_p, \Phi_{p+1}), v) = b(v) \qquad \forall v \in V_p(K).$$

Now one can see that the test space $V_p(K)$ indeed differs from the trial space $X_p(K)$. Thus the presented method is a Petrov-Galerkin method. With the arguments for the choice of the function spaces in mind, Gopalakrishnan and Demkowicz proved the following in [DEGO1, p.7]:

Proposition 2.1.
There is a unique solution to equation (17).

Proof. The system described by equation (17) is square, because $X_p(K)$ and $V_p(K)$ have the same dimension. Thus, it only needs to be shown that if the right hand side of equation (17) vanishes, only the trivial solution fulfills the equation. Indeed, due to equation (15) and (16) one can choose a test function v such that $\vec{\beta} \cdot \vec{\nabla} v = -u_p$ and $v|_{\partial_{out} K} = \Phi_{p+1}$. With that choice, the left hand side of equation (17) degenerates into the norms of u_p and Φ_{p+1}, which therefore have to be 0, because the equation has to hold for all v.

In addition it can be shown that the solution of the spectral method is the best-approximation on its function space X_p, see [DEGO1, p.8].

Now the stability of the method, i.e. its sensitivity with respect to changes in the input data, is analyzed. In order to do this with the new method and later with its application on a multi-element mesh, the authors

introduce local solution operators

$$\mathcal{Q}_p : L^2(\partial_{\text{in}} K) \longmapsto M_{p+1}(\partial_{\text{out}} K) \tag{21}$$

$$\mathcal{U}_p : L^2(\partial_{\text{in}} K) \longmapsto P_p(K). \tag{22}$$

The local solutions $\mathcal{Q}_p\phi, \mathcal{U}_p\phi$ to a function $\phi \in L^2(\partial_{\text{in}} K)$ with $f = 0$ in equation (17) can be defined subsequently as:

$$-(\mathcal{U}_p\phi, \vec{\beta} \cdot \vec{\nabla} v)_K + \langle \mathcal{Q}_p\phi, v \rangle_{\partial_{\text{out}} K} = \langle \phi, v \rangle_{\partial_{\text{in}} K} \tag{23}$$

for all $v \in V_p(K)$. With the definition of the following norms, using $h_{\beta,K}$, the length of the longest line segment in the element K in β-direction, and $\vec{n}_F$, the unit normal to a face F of element K:

$$||\psi||^2_{\frac{1}{\vec{\beta}},\mathscr{F}_h} = \sum_{F \subset \mathscr{F}_h} \frac{1}{|\vec{\beta} \cdot \vec{n}_F|} ||\psi||^2_F \tag{24}$$

$$||v||^2_{\vec{\beta},\mathscr{F}_h} = \sum_{F \subset \mathscr{F}_h} |\vec{\beta} \cdot \vec{n}_F| \, ||v||^2_F. \tag{25}$$

For subcollections $\mathscr{F}_h$ of all faces of K where $\vec{\beta} \neq 0$, the following stability estimates are stated in [DEGO1, p.8]:

$$||\mathcal{Q}_p\phi||^2_{\frac{1}{\vec{\beta}},\partial_{\text{out}} K} \leq ||\phi||^2_{\frac{1}{\vec{\beta}},\partial_{\text{in}} K} \tag{26}$$

$$||\mathcal{U}_p\phi||^2_K \leq \frac{h_{\beta,K}}{\vec{\beta}} ||\phi||^2_{\frac{1}{\vec{\beta}},\partial_{\text{in}} K}. \tag{27}$$

Similar norms and stability estimates are defined for the case $f \neq 0$, in which the right hand side of the stability estimate includes the norms of the function f over K.

3.3 The discontinuous Petrov-Galerkin method on a multi-element mesh

The results in this section have been developed in [DEGO1, p.10-18]. The same analysis as in the previous section is being carried out for a mesh $\mathcal{J}_h$, using just slightly different, now mesh dependent, norms defined on subcollections of mesh elements. These norms incorporate $h_{\vec{\beta},K}$, the length of the longest line segment of an element in β-direction, which puts mild restrictions on the mesh. The method reads

$$a_h((u_h, \Phi_h), v_h) = (f, v_h)_\Omega - \langle \vec{\beta} \cdot \vec{g}, v_h \rangle_{\partial_{\text{in}}\Omega} \tag{28}$$

where u_h, Φ_h and v_h are now defined piecewise on each element equivalent to their one element spaces in equations (14) respectively (20). The bilinear form a_h now entails sums over all mesh elements

$$a_h((u_h, \phi_h), v_h) = \sum_{K \in \mathcal{J}_h} \left[\langle \Phi_h, v_h \rangle_{\partial_{\text{out}} K} - \langle \Phi_h, v_h \rangle_{\partial_{\text{in}} K \setminus \partial_{\text{in}}\Omega} - (u_h, \vec{\beta} \cdot \vec{\nabla} v_h)_K \right]. \tag{29}$$

Further comments on the restrictions of the mesh enable the formulation of a Poincaré inequality, which in turn yields an L^2 error estimate for the solution u. The mesh must be dividable into layers (see [DEGO1, p.13-14]), with each of them having a certain layer-width

$$d_l = \max\{h_{\beta,K} : K \in \text{layer}\}.$$

These are assumed to sum up to less than a mesh dependent constant $L_\Omega \geq \sum_{l=0}^{n} d_l$. If furthermore two neighbouring elements never differ in $h_{\beta,K}$ by more than a factor C_1, then the following can be proven (see [DEGO1, p.14-16]):

$$||\psi||^2_{h,\frac{1}{\vec{\beta}},\mathcal{J}_h} \leq C_0 L_\Omega^2 |||\psi|||^2_{\frac{1}{h},\frac{1}{\vec{\beta}},\mathcal{J}_h} \tag{30}$$

for all outflux functions ψ, where C_0 is independent of mesh and polynomial degree p. The norm on the right shows the sum over all mesh elements of the differences between outflow values of ψ and the inflow values of ψ transported to the outflow boundary on each element. The left hand side norm represents the solution on the outflow boundary. Thus, the inequality bounds the norm of the outflow fluxes by the norm of the inflow fluxes and serves as a stability estimate.

Additionally, an L^2-error estimate can be proven for the solution u_h:

$$||u - u_h||_\Omega^2 \leq \frac{4}{\vec{\beta}}\epsilon(\Phi) + 2 \inf_{\omega_h \in W_h} ||u - \omega_h||_\omega^2. \tag{31}$$

W_h is the piecewise-polynomial trial space within the elements, and $\epsilon(\phi)$ is the sum of left and right side of equation (30). Similar results for the error in the flux function Φ_h in L^2-like norms can be found in [DEGO1, p.16-18].

3.4 Implementation aspects and remarks

The method allows computation of the fluxes Φ_h first and to only subsequently solve for the u_h locally. The reason is, that the finite element space V_h is split into two parts, namely $\varepsilon_{\text{out}}(M_h)$ and its linear independent remainder $\eta_1 P_p(K)$ with chosen degree p. Enumerating the basis for V_h such that the ε_{out} come first will therefore shrink a_h in equations (28) and (29) to an equation no longer depending on u. Thus, the system has a block-triangular structure

$$\begin{pmatrix} A & 0 \\ B & C \end{pmatrix} \begin{pmatrix} \Phi \\ U \end{pmatrix} = F. \tag{32}$$

Furthermore, note that only the functions in $\varepsilon_{\text{out}}(M_h)$ may have discontinuities. Therefore there is no need to integrate the basis functions of u_h with discontinuous functions within the elements. Additionally, there is a way of aligning the basis of M_h in such a way, that the submatrix A is square diagonal, giving way for solving Φ_h by backsubstitution. This advantage may outweigh the cost of solving additional equations due to treating the outflow values as unknowns.

4 Choice of optimal test functions

The following section will cover the concept of choosing an optimal test space which was presented in the second paper published by Demkowicz and Gopalakrishnan [DEGO2, p.3-5]. In addition, a brief subsection sums up the findings into a procedure that incorporates the necessary steps to apply the DPG method.

4.1 Concept of optimal test functions

The general concept can be shown looking at a general variational boundary value problem: Find $u \in U$, such that

$$a(u, v) = l(v) \ \forall v \in V, \tag{33}$$

where a is continuous and fulfills the inf-sup condition. Then Babuška's theorem guarantees the existence of a unique solution and gives an error estimate.
First an energy norm is being defined:

$$||u||_E \overset{\text{def}}{=} \sup_{||v||_V = 1} a(u, v). \tag{34}$$

It is equivalent to the norm on U. The next very important definition will be the equation needed to generate the test functions. The map $T : U \to V$ is defined to be the solution to

$$(Tu, v)_V = a(u, v), \ \forall v \in V. \tag{35}$$

Now it can be proven that the supremum in the energy norm (34) is attained by $v = Tu$ and therefore

$$||u||_E = \sup_{||v||_V=1} a(u,v) = \sup_{||v||_V=1} (Tu,v)_V \text{ by (35)} \tag{36}$$

$$= (Tu, \frac{Tu}{||Tu||_V})_V \tag{37}$$

$$= \frac{1}{||Tu||_V}(Tu,Tu)_V. \tag{38}$$

Given a trial space

$$U_h = \text{span}\{e_j : j = 1,\dots,h\} \tag{39}$$

the optimal test space is defined by

$$V_h = \text{span}\{Te_j : j = 1,\dots,h\}, \tag{40}$$

with V_h being optimal in the sense, that it has the best possible ratio of continuity to stability constant for U in the energy norm. Therefore the choice of this test space yields a solution which, for a given trial space U_h, is the best approximation of u in the energy norm according to Babuška's theorem.

4.2 Practical application

In order to apply the described procedure, four steps have to be performed:

1. Given a boundary value problem, develop a scheme that allows inter-element discontinuities in the test space.
2. Choose a trial subspace U_h with good approximation qualities. Typically these are standard piecewise polynomial spaces.
3. Approximately compute optimal test functions locally by looking for $T_h : U_h \to V_h$ such that

   $$(T_h U_h, \tilde{v}_h)_V = a(u_h, \tilde{v}_h), \ \ \forall \tilde{v}_h \in \tilde{V}_h \tag{41}$$

   with T_h being injective on U_h. Here V_h is a suitable space of discontinuous functions to represent the optimal test space.

4. Finally, the now symmetric positive definite matrix system has to be solved. It is positive definite, because T_n is injective, and symmetric because

   $$\begin{aligned} a(e_j,t_i) &= (T_h e_j, t_i)_V \\ &= (T_h e_j, T_h e_i)_V \\ &= (T_h e_i, T_h e_j)_V \\ &= a(e_i,t_j). \end{aligned} \tag{42}$$

For practical purposes one has to acknowledge that the choice of the inner product will greatly affect the computation of the test function space. It might be they are analytically computable, but usually approximations to the test functions have to be computed.

5 Application and implementation

This section covers the application of the Petrov-Galerkin method to the above 1D pure convection problem, first in the spectral, afterwards in the multi-element case. Both problems are example cases in the papers ([DEGO1], [DEGO2]) on which this seminar work builds and shall serve to confirm the theory which has been developed previously. Both problems will be tackled with their own procedure, which differs a bit in each case. This is because the multi-element case, having inner element boundaries, a different bilinear form a as well as a

different inner product $(.,.)_V$, calls for different test functions. Finally, DG and DPG solutions of both problems will be compared and discussed individually.

The 1D pure convection example is the 1D equivalent of equations (5) and (6), i.e.

$$\beta \cdot u_x = f \quad \text{in } \Omega \tag{43}$$
$$u = u_0 \quad \text{on } \Gamma_{\text{in}} \tag{44}$$

with $\beta = 1$ and $\Omega = (0, 1)$. We choose f and u_0 such that

$$u(x) = \arctan(\alpha(x - x_0)) \tag{45}$$

with choices $\alpha = 100$ and $x_0 = 1$, as in the example in [DEGO1, p.20-21]. The plot shows that the exact solution develops, from an almost flat curve for most of the interval $(0, 1)$, towards a steep slope at $x = 1$.

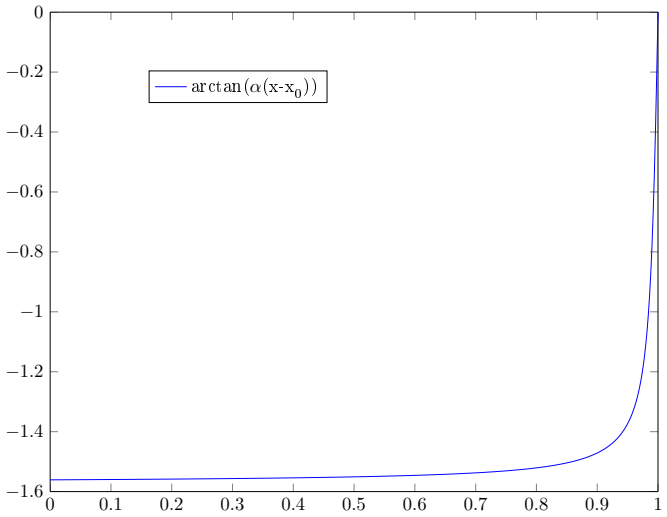

Figure 1: Exact solution of test case.

5.1 Spectral 1D pure convection

In the spectral case, the whole domain Ω is represented by a single element K.

5.1.1 Procedure

Following the procedure for practical application, first a discontinuous Galerkin scheme has to be formulated. Then, a suitable trial space must be chosen. Finally, the test space will be calculated according to (34). In this example the Galerkin scheme in equation (17) is suitable, but reduces, with only one 1D-element $K = \Omega = (0, 1)$, to the following shorter equation:

$$\int_K -uv_x + \int_{\partial K \setminus \Gamma_{\text{in}}} \Phi v = \int_K fv + \int_{\partial K \cap \Gamma_{\text{in}}} u_0 v \tag{46}$$

$$\Leftrightarrow \quad \int_0^1 -uv_x + \Phi v(1) = \int_0^1 fv + u_0 v(0). \tag{47}$$

Here the faces reduce to points and therefore the flux functions on these faces, Φ and u_0, reduce to real numbers. Now a trial function space for the elements must be chosen. The faces are point values while the trial functions

within the element will be Legendre polynomials of degree p, i.e.

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \\ P_p(x) &= \frac{(2p-1)xP_{p-1}(x) - (p-1)P_{p-2}(x)}{p}. \end{aligned}$$

Thus, the trial space is defined by

$$U_p = \mathcal{P}_p(K) \times \mathbb{R}.$$

Finally, the test space functions have to be calculated. Note that this is the point where the DPG method differs from standard DG methods, which use the same function space for both trial and test functions. The chosen inner product reads

$$(v, w)_V = \int_0^1 v'w' + v(1)w(1). \tag{48}$$

In conjunction with the bilinear form

$$a((u, \Phi), v) = \int_0^1 -uv' + \Phi v(1) \tag{49}$$

from equation (47), the determining equation for the test function υ_Φ corresponding to the flux function Φ reads

$$(\upsilon_\Phi, \delta_\upsilon)_V = \int_0^1 \upsilon'_\Phi \delta'_\upsilon + \upsilon_\Phi(1)\delta_\upsilon(1) \stackrel{!}{=} \delta_\upsilon(1) = a((0,1), \delta_\upsilon) \qquad \forall\, \delta_\upsilon \in H^1(K). \tag{50}$$

Usually solving this problem requires the solution to this problem has to be approximated by e.g. another finite element method. However, with our choice of inner product and due to the fact that the problem is 1D, its solution becomes analytically computable. It is

$$\upsilon_\Phi = 1. \tag{51}$$

Similarily the test function υ_u corresponding to the interior trial function u satisfies

$$(\upsilon_u, \delta_\upsilon)_V = \int_0^1 \upsilon'_u \delta'_\upsilon + \upsilon_u(1)\delta_\upsilon(1) \stackrel{!}{=} -\int_0^1 u\delta'_\upsilon = a((u,0), \delta_\upsilon) \qquad \forall\, \delta_\upsilon \in H^1(K), \tag{52}$$

leading to

$$\upsilon_u(x) = \int_x^1 u(s)ds. \tag{53}$$

Therefore the test functions are integrals of the Legendre polynomials that build the trial space. The resulting test space is

$$V_p = \text{span}\{\upsilon_u, \upsilon_\Phi\}. \tag{54}$$

5.1.2 Results

Figure 2 shows the error in u of both the DPG and the DG method for the spectral case. As one can see, the DPG method has a lower error by approximately a factor of 5 while the error converges at the same rate with both methods. Figures 3 and 4 show both the solution to the transport problem with the DG and with the DPG

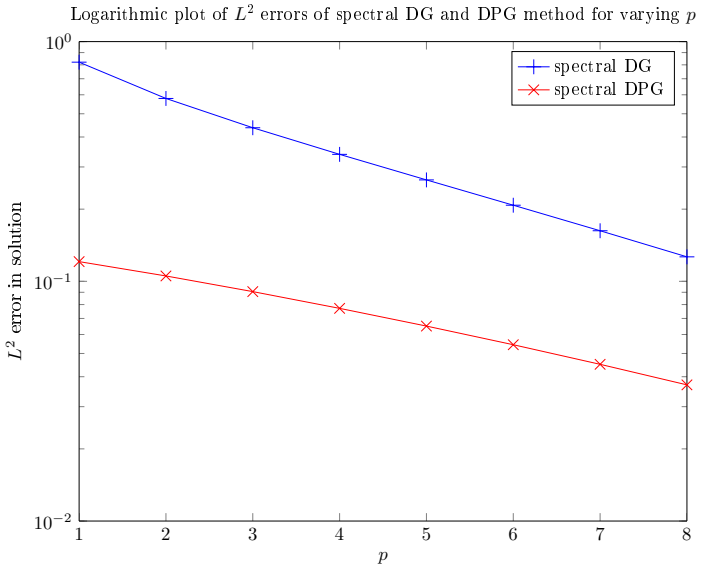

Figure 2: The DPG method shows lower error than the DG method.

method. The DG solution oscillates far more, mainly because its value on the right face is fixed to the value of the exact solution, making it less flexible. The DPG solution on the other hand shows far less oscillations which is caused by the disconnection of the flux value at $x = 1$ from the inner element function.
This is an almost perfect copy of the results shown in [DEGO1, p.21-22], which shows the correctness of their findings.

5.2 Multielement 1D pure convection

The multi-element version of the pure convection problem can, for the most part, be treated in the same way as the spectral problem. The difference stems from the fact, that the mulitelement version of the convection problem lives on the domain $\Omega = (0, 1)$ split into multiple elements Ω_i defined on the intervals (x_i, x_{i+1}). Thus, on inner element faces, the treatment of fluxes has to be dealt with.

5.2.1 Procedure

The inner product now sums over the elements and uses, same as in the spectral case, only the flux at the outflow face:

$$(v, \omega)_V = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} v_x \omega_x + \alpha_i v^{\mathrm{up}}(x_i)\omega^{\mathrm{up}}(x_i). \tag{55}$$

The α_i are scaling factors which can be chosen freely. The bilinear form sums over the elements, too, but has two unknown fluxes on inner elements, one in upwind and one in downwind direction:

$$a((u, \Phi), v) = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} -u v_x + \Phi_i v^{\mathrm{up}}(x_i) - \phi_{i-1} v^{\mathrm{dn}}(x_{i-1}). \tag{56}$$

In this setting $v^{\mathrm{up}}(x_i)$ is the test function defined on the element $[x_i, x_{i+1}]$ evaluated at x_i, while $v^{\mathrm{dn}}(x_{i-1})$ is the test function defined on the element $[x_{i-2}, x_{i-1}]$ evaluated at x_{i-1}. The discontinuous Galerkin scheme

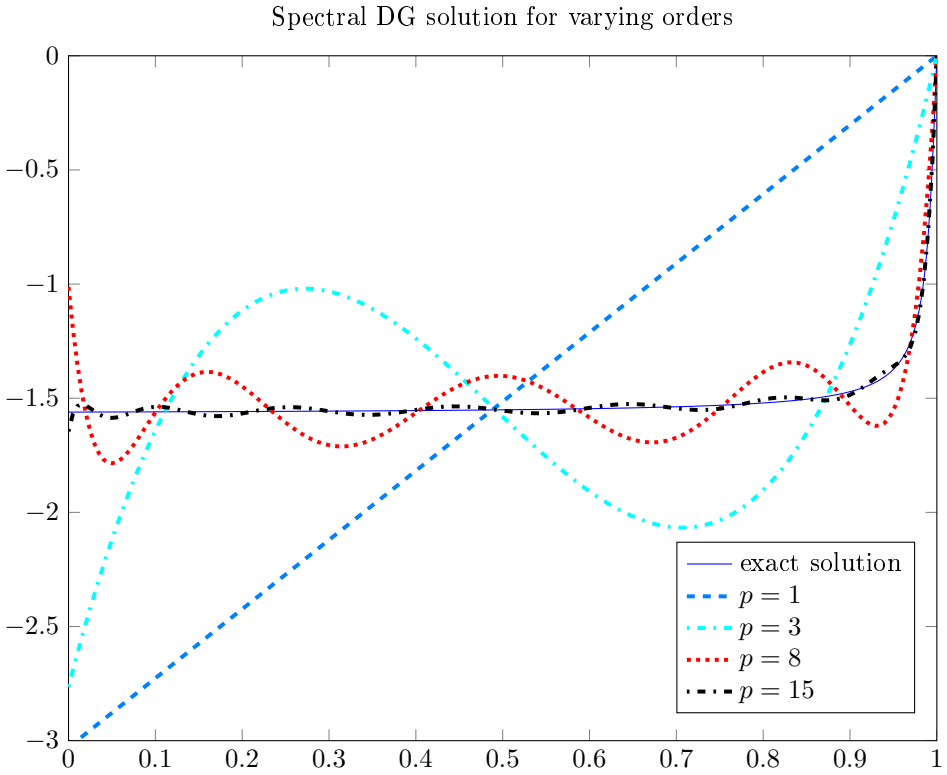

Figure 3: Strong oscillations develop in the DG solution.

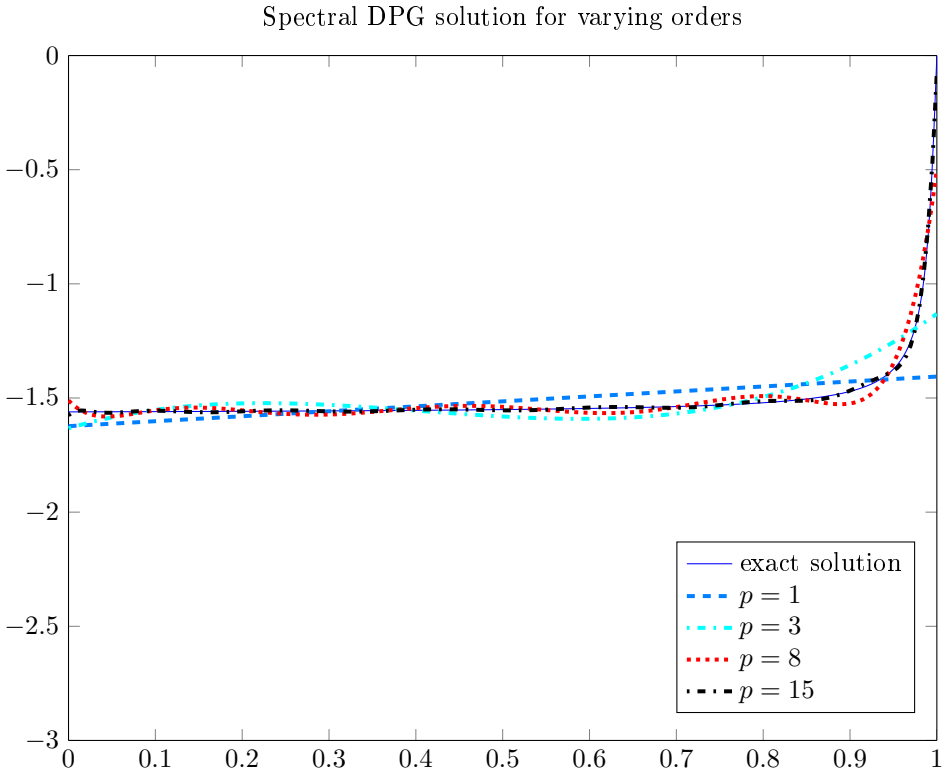

Figure 4: Oscillations are barely visible in the DPG solution.

therefore reads

$$a((u,\Phi),v) = \sum_{i=1}^{n} \int_{x_{i-1}}^{x_i} fv. \tag{57}$$

Now the test function space has to be calculated. It consists of the test functions corresponding to the Legendre polynomials on the inner elements and the test functions on the faces. Contrary to the spectral case there are n outflux faces corresponding to the n elements. The test functions corresponding to the trial functions within the elements are calculated using the new inner product and bilinear form in (55) and (56). The result, similar to the spectral case, reads

$$v_u(x) = \int_{x}^{x_i} u(s)ds. \tag{58}$$

For the fluxes, the case is a little more complicated. Their testfunctions are now supported on two adjacent elements and have to be treated as such.

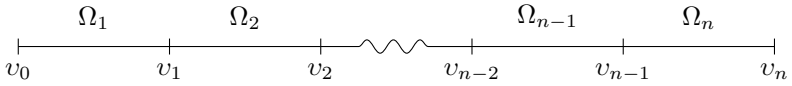

Figure 5: Elements and test functions for faces.

Figure 5 displays the principle. The test function v_1 is dependent on the elements Ω_1 and Ω_2, dependencies of the other inner test functions are accordingly. There are two exception to this rule. One is, that v_0 is on Γ_{in} and therefore does not need to be calculated. The other is v_n, which depends solely on Ω_n. Applying the defining equation (35) to the chosen inner product (c.f. 55) and the bilinear form (c.f. 56) evaluated in one of the flux functions Φ_i, one gets

$$(v_i,\delta_v)_V = \left\{ \begin{array}{ll} \int_{x_{i-1}}^{x_i} v_i'\delta_v' + \alpha_i v_i^{\text{up}}(x_i)\delta_v^{\text{up}}(x_i) & = \delta_v^{\text{up}}(x_i) \\ \int_{x_i}^{x_{i+1}} v_i'\delta_v' + \alpha_{i+1} v_i^{\text{up}}(x_{i+1})\delta_v^{\text{up}}(x_{i+1}) & = -\delta_v^{\text{dn}}(x_i) \end{array} \right\} = a((0,e_i),\delta_v) \tag{59}$$

for all $\delta_v \in H^1(x_{i-1},x_i)$, respectively $\delta_v \in H^1(x_i,x_{i+1})$. The result is

$$v_i(x) = \begin{cases} \frac{1}{\alpha_i} & ,x \in (x_{i-1},x_i) \\ x - x_{i+1} - \frac{1}{\alpha_{i+1}} & ,x \in (x_i,x_{i+1}) \\ 0 & ,\text{else.} \end{cases} \tag{60}$$

Therefore the whole test space contains

$$V_h = \text{span}\{v_u : \forall u \in \mathcal{P}^p(K) \text{ and} \tag{61}$$

$$v_i : \forall i = 1,...,n\}. \tag{62}$$

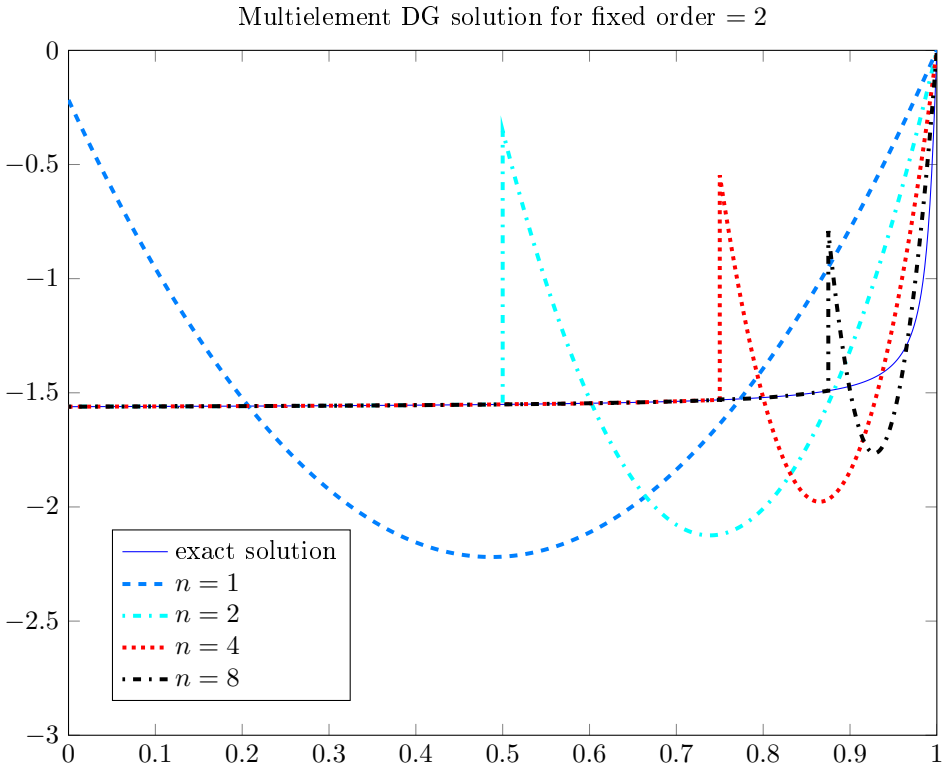

Figure 6: The DG method is bound to $f(1) = 0$.

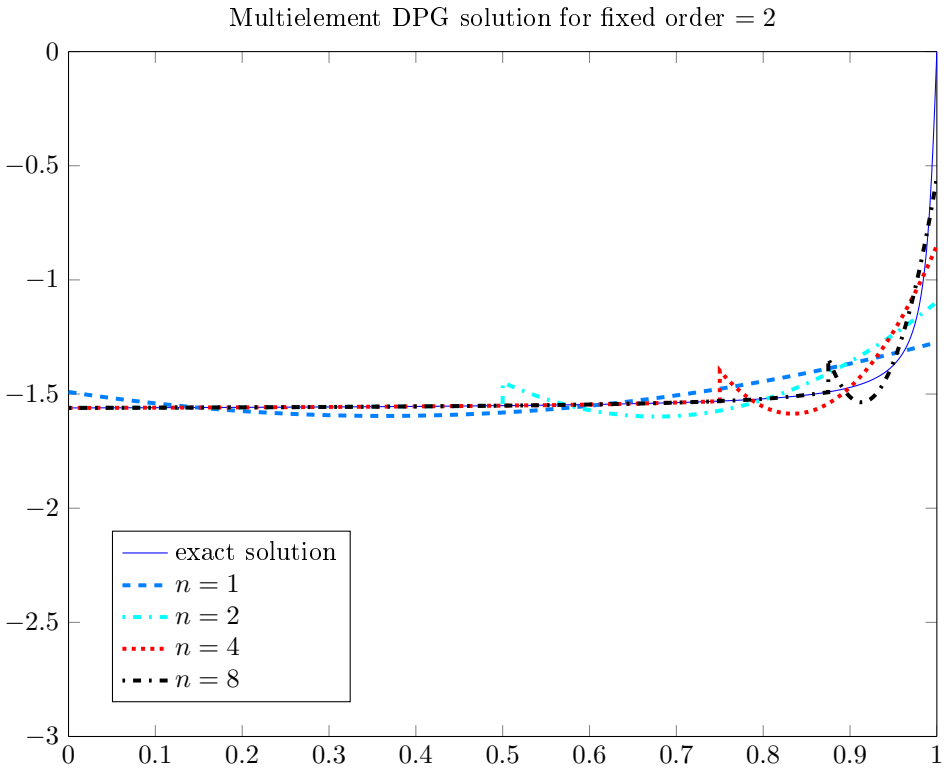

Figure 7: The DPG method is not bound to $f(1) = 0$.

5.2.2 Results

Figures 6 and 7 show results for a fixed order $p = 2$ and a varying number of elements n. Figure 6 displays the DG solution, while figure 7 displays the DPG solution. The solutions show a behaviour similar to the spectral case. The convection problem disconnects at the interfaces, with each element being an almost separate spectral case. Since the exact solution is constant over almost the whole domain, and only introduces a steep slope towards $x = 1$, the elements not affected by the slope behave nicely for both DG and DPG solution. For the elements close to $x = 1$ the DPG solution reduces the error, for it, as in the spectral example, is not bound to the right hand boundary value $f(1) = 0$.

6 Conclusion

A discontinuous Petrov-Galerkin method has been discussed. The implemented method shows, that once the test functions to a given trial space have been correctly calculated, the method achieves lower errors and quicker, more stable convergence than a comparable discontinuous Galerkin solver. The major tasks to set up the Petrov-Galerkin scheme are both calculation of the test functions and incorporation of the flux formulation at the interfaces. This additional work though is justified by the improvement in the results. Further work could entail implementation of the convection-diffusion equation that is being shown in [DEGO2, p.15ff]. Also a 2D example case for the pure convection equation is a possible extension, see [DEGO1, p.21-23].

References

[DEGO1] Demkowicz L., Gopalakrishnan J., "A Class of Discontinuous Petrov-Galerkin Methods. Part I: The Transport Equation" (2010). *Computer Methods in Applied Mechanics and Engineering, Vol. 199, pp.1558-1572.*

[DEGO2] Demkowicz L.,Gopalakrishnan J., "A Class of Discontinuous Petrov-Galerkin Methods. Part II: Optimal Test Functions" (2011). *Numerical Methods for Partial Differential Equations, 27, Issue 1, pp.70-105.*

[BABU70] Babuška I., "Error-bounds for finite element method" (1971). *Numerische Mathematik, 16, pp.322-333.*

[BELY01] Belytschko T., Moes N., Usui S., Parimi C., "Arbitrary discontinuities in finite elements" (2001). *International Journal for Numerical Methods in Engineering, 50, pp.993-1013.*

[GOPA09] Cockburn B., Gopalakrishnan J., Lazarov R., "Unified hybridization of discontinuous Galerkin, mixed and continuous Galerkin methods for second order elliptic problems" (2009). *SIAM Journal on Numerical Analysis, 47, pp.1319-1365.*