ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

High Order Discontinuous Galerkin Method

Semester Project
of
Benjamin Stamm

Directed by:
Prof. A. Quarteroni
Dr. E. Burman

Section of Mathematics, EPFL, Lausanne

June 24, 2004

Contents

1 Motivation

Standard continuous Galerkin-based finite element methods have poor stability properties when applied to transport-dominated flow problems, so excessive numerical stabilization is needed. In contrast, the Discontinuous Galerkin method is known to have good stability properties when applied to first order hyperbolic problems.

2 Outline

In this semester project we will consider the Discontinuous Galerkin method.

In section 3, the transport-reaction problem is presented. It mentions also the hypothesis under which uniqueness and existence of the variational formulation in L^2 is guaranteed. A stability result follows.

In section 4, the pure transport problem is presented as well as a stability result for this problem.

Section 5 introduces the Discontinuous Galerkin method followed by a convergence analysis. The main result of this project is the proof of the convergence theorem.

Section 6 deals with numerical results. In the framework of this semester project a Matlab code is developed for the computation of a numerical approximation. The method is applied to a simple test case with known solution. Finally, the results are analysed.

Section 7 is the conclusion of the Discontinuous Galerkin method.

In sections 8 and 9 we give a rudimentary introduction to orthogonal polynomials and numerical integration.

3 The Transport-Reaction Problem

In this section, the transport-reaction problem is studied with non constant coefficients. The following problem is considered:

Find $u : \Omega \to \mathbb{R}$ such that:

$$u_\beta + \mu u = f \quad \text{in} \quad \Omega \tag{1}$$

$$u = g \quad \text{on} \quad \Gamma_- \tag{2}$$

where $u_\beta = \beta \cdot \nabla u$ denotes the derivative in the β-direction. Γ_- is defined by $\Gamma_- = \{x \in \partial\Omega : n(x) \cdot \beta < 0\}$, where $n(x)$ is the outward normal unit vector at the point x. Analogously Γ_+ is defined by $\Gamma_+ = \{x \in \partial\Omega : n(x) \cdot \beta \geq 0\}$. β is a vector field such that $\beta \in [W^{1,\infty}(\Omega)]^d$, and $\mu \in L^\infty(\Omega)$, $f \in H^1(\Omega)$, $g \in H^{\frac{1}{2}}(\Gamma_-)$. Let be

$$W_\Omega = \{w \in L^2(\Omega) : \beta \cdot \nabla w \in L^2(\Omega)\} \subset L^2(\Omega)$$

with norm

$$\|u\|^2_{W,\Omega} = \|u\|^2_{L^2(\Omega)} + \|\beta \cdot \nabla u\|^2_{L^2\Omega}$$

W is a Hilbert space.

3.1 Uniqueness and existence of the variational formulation in $L^2(\Omega)$

Multiplying (1) by a smooth test function $v : \Omega \to \mathbb{R}$ and integrating on the domain Ω leads to:

Find $u \in W(K)$

$$\int_\Omega (\mu u + \beta \cdot \nabla u) v = \int_\Omega f v \qquad \forall\, v \in L^2(\Omega)$$

Let V be the space

$$V = \{w \in W_\Omega : w|_{\Gamma_-} = g\} \subset W_\Omega \tag{3}$$

The bilinear form $a : W_\Omega \times L^2(\Omega) \to \mathbb{R}$ and the linear form $F : L^2(\Omega) \to \mathbb{R}$ are defined by

$$\begin{aligned} a(u,v) &= \int_\Omega (\mu u + \beta \cdot \nabla u) v \qquad \forall u \in W_\Omega, \forall v \in L^2(\Omega) \\ F(v) &= \int_\Omega v f \qquad\qquad \forall v \in L^2(\Omega) \end{aligned}$$

Then, the variational formulation in $L^2(\Omega)$ is:

Find $u \in V$ such that:

$$a(u,v) = F(v) \qquad \forall v \in L^2(\Omega) \tag{4}$$

It can be shown that under the hypothesis that there exists a constant μ_0 such that

$$\mu(x) - \frac{1}{2} \nabla \cdot \beta(x) \geq \mu_0 > 0 \quad \text{a.e. in} \quad \Omega \tag{5}$$

the conditions of the Nečas Theorem are satisfied. That implies that there exists a unique solution of (4). In addition, condition (5) is also necessary for uniqueness and existence. For more details see [1].

3.2 Stability

In this section, a stability result for the Transport-Reaction Problem on the whole domain is developed.

Lemma 3.1 (stability for the transport-reaction problem) *If it exists a constant μ_1 such that $\mu(x) \geq \mu_1 > 0\ \forall\, x \in \Omega$, then the following stability result is given:*

$$\mu_1 \|u\|^2_{L^2(\Omega)} + \int_{\Gamma_+} |\beta \cdot n| u^2 \leq \frac{1}{\mu_1} \|f\|^2_{L^2(\Omega)} + \int_{\Gamma_-} |\beta \cdot n| g^2$$

PROOF. Let us take equation (1), multiply it by u and integrate over Ω. We get

$$(u_\beta, u)_\Omega + (\mu u, u)_\Omega = (f, u)_\Omega$$

Then, the following integration by parts is used

$$\begin{aligned} (u_\beta, u)_\Omega &= -(u, u_\beta)_\Omega + (\beta \cdot n u, u)_{\partial\Omega} \\ &= -(u_\beta, u)_\Omega + (|\beta \cdot n| u, u)_{\Gamma_+} - (|\beta \cdot n| g, g)_{\Gamma_-} \end{aligned}$$

so that

$$(\mu u, u)_\Omega + \frac{1}{2}(|\beta \cdot n| u, u)_{\Gamma_+} = (f, u)_\Omega + \frac{1}{2}(|\beta \cdot n| g, g)_{\Gamma_-}$$

Now, the Cauchy-Schwarz inequality is applied to $(f, u)_\Omega$. This leads to

$$2(\mu u, u)_\Omega + (|\beta \cdot n| u, u)_{\Gamma_+} = 2\|f\|_\Omega \|f\|_\Omega + (|\beta \cdot n| g, g)_{\Gamma_-}$$

Finally the Young inequality with $\varepsilon = \frac{1}{2\mu_1}$ is applied. Then

$$\mu_1 \|u\|^2_{L^2(\Omega)} + \int_{\Gamma_+} |\beta \cdot n| u^2 \leq \frac{1}{\mu_1}\|f\|^2_{L^2(\Omega)} + \int_{\Gamma_-} |\beta \cdot n| g^2$$

□

4 The pure Transport Problem

In this section, the pure Transport Problem is considered:

Find $u : \Omega \to \mathbb{R}$ such that

$$\begin{aligned} u_\beta &= f \quad \text{in} \quad \Omega \subset \mathbb{R}^d && (6) \\ u &= g \quad \text{on} \quad \Gamma_- && (7) \end{aligned}$$

Then the variational formulation of this problem is:

Find $u \in W_\Omega$ such that:

$$\begin{aligned} (u_\beta, v)_\Omega &= (f, v)_\Omega \quad \forall v \in L^2(\Omega) && (8) \\ u &= g \quad \text{on} \quad \Gamma_- \end{aligned}$$

4.1 Stability

Lemma 4.1 (stability for the pure Transport Problem) *In the case of the pure transport problem, i.e. if $\mu \equiv 0$ and under the hypothesis that there exists a vector function $\gamma \in \left(L^\infty(K)\right)^d$ such that it exists a constant γ_1 which satisfies*

$$\vec{\gamma} \cdot \vec{\beta} \geq \gamma_1 > 0$$

then, the following stability result is given

$$\begin{aligned} &\gamma_1 \|e^{-\vec{\gamma}\cdot\vec{x}} u\|^2_{L^2(\Omega)} + (|\beta \cdot n| e^{-\vec{\gamma}\cdot\vec{x}} u, e^{-\vec{\gamma}\cdot\vec{x}} u)_{\Gamma_+} \\ &\leq \frac{1}{\gamma_1}\|e^{-\vec{\gamma}\cdot\vec{x}} f\|^2_{L^2(\Omega)} + (|\beta \cdot n| e^{-\vec{\gamma}\cdot\vec{x}} g, e^{-\vec{\gamma}\cdot\vec{x}} g)_{\Gamma_-} \end{aligned}$$

PROOF. Let be $\tilde{u}(\vec{x}) = e^{-\vec{\gamma}\cdot\vec{x}} u(\vec{x})$ and note that problem (6) is equivalent to the following problem:

Find $\tilde{u} : \Omega \to \mathbb{R}$ such that

$$\begin{aligned} \vec{\beta} \cdot \nabla \tilde{u} + \nu \tilde{u} &= e^{-\vec{\gamma}\cdot\vec{x}} f \quad \text{in} \quad \Omega \\ \tilde{u} &= e^{-\vec{\gamma}\cdot\vec{x}} g \quad \text{on} \quad \Gamma_- \end{aligned}$$

where $\nu = \vec{\gamma} \cdot \vec{\beta}$. This is a transport-reaction problem and thanks to the hypothesis, Lemma (3.1) can be applied, so that we get the result.

□

5 Discontinuous Galerkin Method

5.1 Notations

Let us first introduce some notations. Consider an element K. K can be arbitrary a simplex or a parallelepiped , then ∂K is split into

$$\begin{aligned} \partial K_- &= \{x \in \partial K : n(x) \cdot \beta < 0\} \\ \partial K_+ &= \{x \in \partial K : n(x) \cdot \beta \geq 0\} \end{aligned}$$

and we have that $\partial K = \partial K_- \cup \partial K_+$.

$$\begin{aligned} v^- &= \lim_{s \to 0^-} v(x + s\beta) \\ v^+ &= \lim_{s \to 0^+} v(x + s\beta) \end{aligned}$$

$$[v] = v^+ - v^-$$

$$\hat{v}(x) = \begin{cases} v^+ & x \in \partial K_- \\ v^- & x \in \partial K_+ \end{cases}$$

Let be h_K the diameter of the element K and $h = \max_K h_K$.
In addition, $(u, v)_\Lambda$ denotes the usual L^2-scalar product on Λ.

5.2 The method

In this section, the principle of the Discontinuous Galerkin method is presented. Let us consider the transport-reaction equations. We will begin by studying the problem restricted to one element, meaning a simplex or parallelepiped. On the element K, the following problem is considered:

Find $u : \Omega \to \mathbb{R}$ such that

$$\begin{aligned} u_\beta + \mu u &= f \quad \text{in} \quad K && (9) \\ u &= u^- \quad \text{on} \quad \partial K_- && (10) \end{aligned}$$

where u^- is a given function on ∂K_-. The associated variational formulation is:

Find $u \in W_K$

$$\begin{aligned} (u_\beta + \mu u, v)_K &= (f, v)_K \qquad \forall v \in \ L^2(K) \\ u &= u^- \quad \text{on} \quad \partial K_- \end{aligned}$$

Using integration by parts leads to:

$$(u_\beta, v)_K + (\mu u, v)_K = -(u, v_\beta)_K + (\beta \cdot n\hat{u}, \hat{v})_{\partial K} + (\mu u, v)_K = (f, v)_K$$

where n is the outward unit vector of K. Boundary conditions are imposed weakly, so that the problem is:

Find $u \in W_K$ such that:

$$(\mu u, v)_K - (u, v_\beta)_K + (\beta \cdot n u^-, v^-)_{\partial K_+} = (f, v)_K + (|\beta \cdot n| u^-, v^+)_{\partial K_-} \qquad \forall v \in \ W_K$$

Note that u^- on ∂K_- is the given function of the boundary condition, but that u^- on ∂K_+ is the solution on ∂K_+.

Now, a Galerkin approximation is used. That means that the space W_K is replaced by the finite dimensional space V_h which satisfies $V_h \subset W_K$. The formulation is the following:

Find $u_h \in V_h$ such that:

$$\begin{aligned}&(\mu u_h, v_h)_K - (u_h, v_{h\beta})_K + (\beta \cdot n u_h^-, v_h^-)_{\partial K_+}\\ =\quad &(f, v_h)_K + (|\beta \cdot n| u_h^-, v_h^+)_{\partial K_-} \qquad \forall v \in\ V_h\end{aligned}$$

where u_h^- is an approximation of u^- in the space V_h. Then using integration by parts for $(u_h, v_{h\beta})_K$ a second time leads to:

Find $u_h \in V_h)$ such that:

$$a_K(u_h, v_h) = (f, v_h)_K \qquad \forall v \in V_h \tag{11}$$

where u_h^- is given on the inflow boundary ∂K_- and

$$a_K(u_h, v_h) = (u_{h\beta} + \mu u_h, v_h)_K + (|\beta \cdot n| [u_h], v_h^+)_{\partial K_-}$$

Choosing a basis $\{\varphi_i\}$ for the space V_h and writing $u_h = \sum_j u_j \varphi_j(x)$ leads to a algebraic system:

$$A\vec{u} = \vec{f} \tag{12}$$

where

$$\begin{aligned}A_{i,j} &= (\varphi_{j,\beta} + \mu\varphi_j, \varphi_i)_K + (|\beta \cdot n| \varphi_j^+, \varphi_i^+)_{\partial K_-}\\ \vec{f_i} &= (f, \varphi_i)_K + (|\beta \cdot n| u^-, \varphi_i^+)_{\partial K_-}\end{aligned}$$

To solve the problem on one element, one needs only to know the the inflow boundary data and f. Because the data is given on the inflow boundary of the whole domain, one can find an order of elements such that the problem can be solved element by element. For an example see Fig. 1. Instead of solving a big system of linear equations, many small systems of linear equations has to be solved when using Discontinuous Galerkin.

Now, we would like to show existence and uniqueness of the approximation u_h. Since the problem can be uncoupled, it is sufficient to prove well posedness for one element K. The following lemma shows the result in the case of the Transport-Reaction problem.

Lemma 5.1 *If $\mu \in L^\infty(\Omega)$ is such that*

$$\mu(x) \geq \mu_1 > 0 \qquad \forall\, x \in \Omega$$

then for each element $K \in \tau_h$, it exists a unique solution of the linear system (12)

$$A\vec{u} = \vec{f}$$

for every constant vector $\vec{f}$.

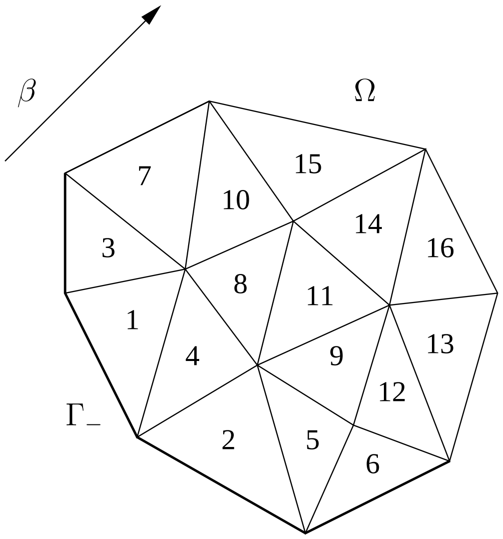

Figure 1: Example of an order of triangles for the Discontinuous Galerkin Method

PROOF. First, the test function $v_h \in V_h$ of the variational formulation is chosen equal to u_h. Using partial integration, we obtain

$$(u_{h\beta}, u_h)_K = -(u_h, u_{h\beta})_K + (\beta \cdot n u_h^+, u_h^+)_{\partial K_-} + (\beta \cdot n u_h^-, u_h^-)_{\partial K_+} \tag{13}$$

so that

$$2\,(u_{h\beta}, u_h)_K = (\beta \cdot n u_h^+, u_h^+)_{\partial K^-} + (\beta \cdot n u_h^-, u_h^-)_{\partial K^+} \tag{14}$$

Secondly, the left hand side of the variational formulation is considered.

$$\begin{aligned} & 2\,(u_{h\beta}, u_h)_K + 2\,(|\beta \cdot n| u_h^+, u_h^+)_{\partial K_-} \\ = \;& -(|\beta \cdot n| u_h^+, u_h^+)_{\partial K_-} + (|\beta \cdot n| u_h^-, u_h^-)_{\partial K_+} \\ & +2\,(|\beta \cdot n| u_h^+, u_h^+)_{\partial K_-} \\ = \;& (|\beta \cdot n| u_h^-, u_h^-)_{\partial K_+} + (|\beta \cdot n| u_h^+, u_h^+)_{\partial K_-} \\ = \;& \int_{\partial K} |\beta \cdot n| \hat{u_h}^2 d\sigma \end{aligned} \tag{15}$$

so that

$$\begin{aligned} & 2\,(u_{h\beta} + \mu u_h, u_h)_K + 2\,(|\beta \cdot n| u_h^+, u_h^+)_{\partial K_-} \\ & = \int_{\partial K} |\beta \cdot n| \hat{u_h}^2 d\sigma + 2 \int_K \mu u_h^2 dx \end{aligned}$$

Using the same basis of the space V_h as above, this is equivalent to the following equation:

$$\vec{u} A \vec{u} = \frac{1}{2} \int_{\partial K} |\beta \cdot n| u_h^2 d\sigma + \int_K \mu \hat{u_h}^2 dx \tag{16}$$

If $A\vec{u} = 0$ then $\vec{u} A \vec{u} = 0$ and as consequence

$$0 \geq \frac{1}{2} \int_{\partial K} |\beta \cdot n| \hat{u_h}^2 d\sigma + \mu_1 \int_K u_h^2 dx \geq \|u_h\|^2_{L^2(K)} \geq 0 \tag{17}$$

That implies that $u_h \equiv 0$ and $\vec{u} = \vec{0}$. This proves the regularity of the matrix A as well as the uniqueness and existence of the approximation u_h.

□

In the case of the pure Transport Problem, we have the following result.

Lemma 5.2 *We assume that one of the following conditions are satisfied.*

- $V_h = P^1(K)$
- $V_h = P^2(K)$ *and each side* Γ *of the the triangle K satisfies*

$$|\beta \cdot n(x)| > C \qquad \forall x \in \Gamma$$

then, there exists a unique solution of the linear system (12), derived from the pure transport problem,

$$A\vec{u} = \vec{f}$$

for every constant vector $\vec{f}$.

PROOF. Taking equation (16) with $\mu = 0$ leads to:

$$\vec{u}A\vec{u} = \frac{1}{2}\int_{\partial K} |\beta \cdot n| \hat{u_h}^2 d\sigma \tag{18}$$

The uniqueness and existence of u_h follows of the regularity of the matrix A. Let be $\vec{u}$ so that $A\vec{u} = \vec{0}$, then

$$\int_{\partial K} |\beta \cdot n| \hat{u_h}^2 d\sigma = 0 \tag{19}$$

If it can be shown that $\vec{u} = \vec{0}$, resp. $u_h = 0$, A is a regular matrix and u_h exists and is unique for each $\vec{f}$. But concluding that $u_h = 0$ depends now from the space V_h.

- If $V_h = P^1(K)$, then $u_h = 0$, also if β is parallel to one side of the element K.
- If $V_h = P^2(K)$ it is necessary that each side Γ of the triangle K satisfies

$$|\beta \cdot n(x)| > C \qquad \forall x \in \Gamma \tag{20}$$

  where $C > 0$. This is a supplementary condition of the triangulation.

□

Remark: Figure 2 shows the interpolation points for the different spaces.
Remark: If $V_h = P^N(K)$ with $N \geq 3$, there is no control over the basis functions with support in the interior of the triangle. For example suppose that $N = 3$, then one interpolation point p_{int} is in the interior of the triangle. Let ψ be the function such that

$$\psi(p_{int}) = 1 \qquad \text{and} \qquad \psi(x) = 0 \quad \forall x \in \partial K$$

Then

$$\int_{\partial K} |\beta \cdot n| \psi^2 = 0 \quad \text{but} \quad \psi(x) \neq 0 \; \forall x \in int(K)$$

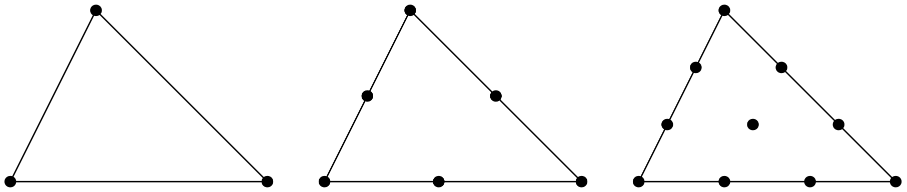

Figure 2: The interpolation points for $P^1(K)$, $P^2(K)$ and $P^3(K)$

5.3 Convergence Analysis

In this section, a global convergence result for the Discontinuous Galerkin method is developed under the hypothesis that $\mu = 1$. Note that in this section C and C_K denotes generic constants taking different values. We will not pay attention to the explicit form ot these constants but note that they are independent of h and N, unless otherwise noted. We start from the local problem defined by (11). Then, the sum over all elements $K \in \tau_h$ is taken and. This describes the global problem:

Find $u_h \in W_h^N$ such that

$$\sum_{K \in \tau_h} a_K(u_h, v_h) = (f, v_h)_\Omega + (|\beta \cdot n| g, v_h^+)_{\Gamma_-} \qquad \forall v_h \in W_h^N$$

where

$$W_h^N = \{w \in L^2(\Omega) | \; w_{|_K} \in V_h \;, \; \forall K \in \tau_h\}$$

In fact, this is equivalent to

Find $u_h \in W_h^N$ such that

$$a(u_h, v_h) = (f, v_h)_\Omega + (|\beta \cdot n| g, v_h^+)_{\Gamma_-} \qquad \forall v_h \in \; W_h^N \tag{21}$$

where

$$a(u_h, v_h) = (\mu u_h + u_{h,\beta}, v_h)_\Omega + \sum_K (|\beta \cdot n| [u_h], v_h^+)_{\partial K_- \setminus \Gamma} + (|\beta \cdot n| u_h^+, v_h^+)_{\Gamma_-}$$

Note that the result is developed for $\mu = 1$. Taking $v_h = u_h$ and integrating by parts, we obtain

$$\begin{aligned} a(u_h, u_h) &= \|u_h\|^2_{L^2(\Omega)} + \sum_K (u_{h,\beta}, u_h)_K + \sum_K (|\beta \cdot n| [u_h], v_h^+)_{\partial K_- \setminus \Gamma} \\ &\quad + (|\beta \cdot n| u_h^+, v_h^+)_{\Gamma_-} \\ &= \|u_h\|^2_{L^2(\Omega)} + \sum_K (|\beta \cdot n| [u_h], u_h^+)_{\partial K_- \setminus \Gamma} + (|\beta \cdot n| u_h^+, u_h^+)_{\Gamma_-} \\ &\quad + \sum_K \frac{1}{2} \Big((|\beta \cdot n| u_h^-, u_h^-)_{\partial K_+} - (|\beta \cdot n| u_h^+, u_h^+)_{\partial K_-} \Big) \end{aligned}$$

Now, the following equality is used

$$(a - b)a = \frac{1}{2}(a^2 + (a - b)^2 - b^2)$$

so that

$$(|\beta\cdot n|[u_h], u_h^+)_{\partial K_-\setminus\Gamma} = \frac{1}{2}\int_{\partial K_-\setminus\Gamma} |\beta\cdot n|\big((u_h^+)^2 + [u_h]^2 - (u^-)^2\big)$$

then

$$\begin{aligned} a(u_h,u_h) &= \|u_h\|^2_{L^2(\Omega)} + \sum_K \Big[\frac{1}{2}\int_{\partial K_-\setminus\Gamma} |\beta\cdot n|\Big((u_h^+)^2 + [u_h]^2 - (u_h^-)^2\Big) \\ &\quad + \frac{1}{2}\int_{\partial K_-\setminus\Gamma} |\beta\cdot n|(u_h^-)^2 - \frac{1}{2}\int_{\partial K_-\setminus\Gamma} |\beta\cdot n|(u_h^+)^2\Big] \\ &\quad + \frac{1}{2}\int_\Gamma |\beta\cdot n|\hat{u_h}^2 \\ &= \|u_h\|^2_{L^2(\Omega)} + \sum_K \Big(\frac{1}{2}\int_{\partial K_-\setminus\Gamma} |\beta\cdot n|[u_h]^2\Big) + \frac{1}{2}\int_\Gamma |\beta\cdot n|\hat{u_h}^2 \end{aligned}$$

This motivates us to define the following norm

$$|||u_h|||^2 = \|u_h\|^2_{L^2(\Omega)} + \sum_K \Big(\frac{1}{2}\int_{\partial K_-\setminus\Gamma} |\beta\cdot n|[u_h]^2\Big) + \frac{1}{2}\int_\Gamma |\beta\cdot n|\hat{u_h}^2$$

and as consequence

$$|||u_h|||^2 = a(u_h,u_h).$$

We immediately conclude that

$$|||u_h|||^2 = (f,u_h)_\Omega + (|\beta\cdot n|g, u_h^+)_{\Gamma_-}.$$

Existence and uniqueness of the discrete solution follows. For $u \in H^1(\Omega)$ the numerical schema is consistent. This is expressed in the form of the Galerkin orthogonality in the following lemme.

Lemma 5.3 (Galerkin orthogonality)

$$a(u_h - u, v_h) = 0 \qquad \forall\, v_h \, \in W_h^N$$

where u is the exact solution and u_h solution of (21).

PROOF. First, since u_h is solution of (21), it satisfies

$$a(u_h,v_h) = (f,v_h)_\Omega + (|\beta\cdot n|g, v_h^+)_{\Gamma_-} \qquad \forall v_h \in\ W_h^N \tag{22}$$

Secondly, with u the exact solution, we have

$$\begin{aligned} (\mu u + u_\beta, v_h)_\Omega &= (f,v_h)_\Omega \qquad \forall v_h \in\ W_h^N \\ (u-g, v_h)_{\Gamma_-} &= 0 \qquad \forall v_h \in\ W_h^N \end{aligned}$$

In addition, $u \in H^1(\Omega)$ and as consequence $u \in H^1(K)\ \ \forall K \in \tau_h$. So that the trace is well defined and hence

$$\int_{\partial K} |\beta\cdot n|[u]v \leq C\int_{\partial K}[u]v = 0.$$

So that

$$a(u,v_h) = (f,v_h)_\Omega + (|\beta\cdot n|g, v_h^+)_{\Gamma_-} \qquad \forall v_h \in\ W_h^N \tag{23}$$

Taking the difference between (22) and (23) leads to

$$a(u_h - u, v_h) = 0 \qquad \forall\, v_h \, \in W_h^N.$$

□

Now, the local projection $P_N : L^2(K) \to W_h^N$ is introduced. For all $u \in L^2(K)$, there exists $P_N u \in W_h^N \cap C^0(K)$ such that

$$\int_K (u - P_N u) v_h = 0 \quad \forall v_h \in W_h^N.$$

Let be h_K the diameter of the element K defined by $h_K = \max_{x,y \in K} |x - y|$. Then h is defined by $h = \max_K h_K$ In addition, we refer to [2] for the following Lemma's.

Lemma 5.4 (Local Projection Error on the boundary ∂K) *Let $K \in \tau_h$ and suppose that $u \in H^k(K)$ for some integer $k \geq 1$. Then, for any integer $s, 1 \leq s \leq min(N+1,k)$ and $N \geq 1$, we have that*

$$\|u - P_N u\|_{L^2(\partial K)} \leq C_K(d,s) \frac{h_K^{s-\frac{1}{2}}}{(N+1)^{s-\frac{1}{2}}} |u|_{H^s(K)}$$

Lemma 5.5 (Local Projection Error) *Let $K \in \tau_h$ and suppose that $u \in H^k(K)$ for some integer $k \geq 1$. Then, for any integer $s, 1 \leq s \leq min(N+1,k)$ and $N \geq 1$, we have that*

$$\|u - P_N u\|_{L^2(K)} \leq C_K \frac{h_K^s}{N^s} |u|_{H^s(K)}$$

We first state and proof two approximation lemmas.

Lemma 5.6 *If there exists a constant $C(\beta)$ such that $|\beta \cdot n| \leq C(\beta)$, then*

$$|||\eta||| \leq \sum_K C_K(d,s,\beta) \frac{h_K^{s-\frac{1}{2}}}{N^{s-\frac{1}{2}}} |u|_{H^s(K)}$$

PROOF Lemma(5.6)

$$|||\eta|||^2 = \|\eta\|^2_{L^2(\Omega)} + \sum_K \frac{1}{2} \int_{\partial K_- \setminus \Gamma} |\beta \cdot n| [\eta]^2 + \sum_K \frac{1}{2} \int_{\partial K \cap \Gamma} |\beta \cdot n| \hat{\eta}^2$$

Considering each term of this sum:

- $\|\eta\|^2_{L^2(\Omega)} \leq \sum_K \left(C \frac{h_K^s}{N^s} |u|_{H^s(K)} \right)^2$ by Lemma(5.5)
- $$\begin{aligned} \sum_K \frac{1}{2} \int_{\partial K \cap \Gamma} |\beta \cdot n| \hat{\eta}^2 &\leq \sum_K \frac{1}{2} \int_{\partial K} |\beta \cdot n| \hat{\eta}^2 \\ &\leq C(\beta) \sum_K \|\eta\|^2_{L^2(\partial K)} \leq \sum_K \left(C_K(d,s,\beta) \frac{h_k^{s-\frac{1}{2}}}{(N+1)^{s-\frac{1}{2}}} |u|_{H^s(K)} \right)^2 \end{aligned}$$
- $$\begin{aligned} \sum_K \frac{1}{2} \int_{\partial K_- \setminus \Gamma} |\beta \cdot n| [\eta]^2 &\leq \sum_K \frac{1}{2} \int_{\partial K_- \setminus \Gamma} |\beta \cdot n| \left(\eta^{+2} + \eta^{-2} \right) \\ &\leq 2 \sum_K \int_{\partial K} |\beta \cdot n| \hat{\eta}^2 \leq C(\beta) \sum_K \|\eta\|^2_{L^2(\partial K)} \\ &\leq \sum_K \left(C_K(d,s,\beta) \frac{h_k^{s-\frac{1}{2}}}{(N+1)^{s-\frac{1}{2}}} |u|_{H^s(K)} \right)^2 \end{aligned}$$

So that

$$\begin{aligned}
|||\eta||| &\leq \Big(\sum_K \big(C_K(d,s,\beta)\frac{h_k^{s-\frac{1}{2}}}{(N+1)^{s-\frac{1}{2}}}|u|_{H^s(K)}\big)^2 \\
&\quad +\big(C_K(d,s,\beta)\frac{h_k^{s-\frac{1}{2}}}{N^{s-\frac{1}{2}}}|u|_{H^s(K)}\big)^2\Big)^{\frac{1}{2}} \\
&\leq \Big(\sum_K \big(C_K(d,s,\beta)\frac{h_k^{s-\frac{1}{2}}}{N^{s-\frac{1}{2}}}|u|_{H^s(K)}\big)^2\Big)^{\frac{1}{2}} \leq \sum_K C_K(d,s,\beta)\frac{h_k^{s-\frac{1}{2}}}{N^{s-\frac{1}{2}}}|u|_{H^s(K)}
\end{aligned}$$

□ Lemma(5.6)

Lemma 5.7 *If there exists a constant $C(\beta)$ such that $|\beta\cdot n| \leq C(\beta)$, then*

$$\Big(\sum_K \int_{\partial K} |\beta\cdot n|(\eta^-)^2\Big)^{\frac{1}{2}} \leq \sum_K C_K(d,s,\beta)\frac{h_K^{s-\frac{1}{2}}}{(N+1)^{s-\frac{1}{2}}}|u|_{H^s(K)}$$

PROOF Lemma(5.7)

$$\begin{aligned}
\Big(\sum_K \int_{\partial K} |\beta\cdot n|(\eta^-)^2\Big)^{\frac{1}{2}} &\leq \Big(\sum_K C\|\hat{\eta}\|^2_{L^2(\partial K)}\Big)^{\frac{1}{2}} \\
&\leq \Big(\sum_K (C_K(d,s,\beta)\frac{h_K^{s-\frac{1}{2}}}{(N+1)^{s-\frac{1}{2}}}|u|_{H^s(K)})^2\Big)^{\frac{1}{2}} \\
&\leq \sum_K C_K(d,s,\beta)\frac{h_K^{s-\frac{1}{2}}}{(N+1)^{s-\frac{1}{2}}}|u|_{H^s(K)}
\end{aligned}$$

□ Lemma(5.7)

Theorem 5.8 (Global Convergence) *Suppose that $u \in H^k(K)$ for some integer $k \geq 1$ and there exists a constant $C(\beta)$ such that $|\beta\cdot n| \leq C(\beta)$. Then, for any integer $s, 1 \leq s \leq min(N+1,k)$ and $N \geq 1$, we have that*

$$|||u-u_h||| \leq C(d,s,\beta)\frac{h^{s-\frac{1}{2}}}{N^{s-\frac{1}{2}}}|u|_{H^s(\Omega)}$$

PROOF. Let be $\eta = P_N u - u$ and $\xi = P_N u - u_h$. Then

$$|||u-u_h||| \leq |||u-P_N u||| + |||P_N u - u_h||| = |||\eta||| + |||\xi|||$$

Let us show the following inequality

$$|||\xi||| \leq \sum_K C_K(d,s,\beta)\frac{h_K^{s-\frac{1}{2}}}{(N+1)^{s-\frac{1}{2}}}|u|_{H^s(K)}$$

The Galerkin Orthogonality is used for developing

$$\begin{aligned}
|||\xi|||^2 &= a(P_N u - u_h, P_N u - u_h) + a(u_h - u, P_N u - u_h) = a(\eta,\xi) \\
&= \sum_K (\eta,\xi)_K + \sum_K (\eta_\beta,\xi)_K + \sum_K (|\beta\cdot n|[\eta],\xi^+)_{\partial K_-\setminus\Gamma} \\
&\quad +(|\beta\cdot n|\eta^+,\xi^+)_{\Gamma_-}
\end{aligned}$$

Observe that $\xi = P_N u - u_h \in W_h^N$ and as consequence

$$(\eta, \xi)_K = 0$$

Using integration by parts of $(\eta_\beta, \xi)_K$:

$$(\eta_\beta, \xi)_K = (\eta, \xi_\beta)_K + \sum_K (|\beta \cdot n|\hat{\eta}, \hat{\xi})_{\partial K}$$

and the fact that $\xi_\beta \in W_h^N$ and in consequence that $(\eta, \xi_\beta)_K = 0$ leads to

$$\begin{aligned}
|||\xi|||^2 &= \sum_K (|\beta \cdot n|[\eta], \xi^+)_{\partial K_- \setminus \Gamma} + (|\beta \cdot n|\eta^+, \xi^+)_{\Gamma_-} \\
&\quad + \sum_K (|\beta \cdot n|\eta^-, \xi^-)_{\partial K_+} - \sum_K (|\beta \cdot n|\eta^+, \xi^+)_{\partial K_-} \\
&= \sum_K (|\beta \cdot n|\eta^-, \xi^-)_{\partial K_+} - \sum_K (|\beta \cdot n|\eta^-, \xi^+)_{\partial K_- \setminus \Gamma} \\
&= \sum_K (|\beta \cdot n|\eta^-, \xi^- - \xi^+)_{\partial K_- \setminus \Gamma} + (|\beta \cdot n|\eta^-, \xi^-)_{\Gamma_+}
\end{aligned}$$

First, the Cauchy-Schwarz inequality is applied to

$$(|\beta \cdot n|\eta^-, \xi^-)_{\Gamma_+} \leq \Big(\int_{\Gamma_+} |\beta \cdot n|(\eta^-)^2\Big)^{\frac{1}{2}} \Big(\int_{\Gamma_+} |\beta \cdot n|(\xi^-)^2\Big)^{\frac{1}{2}}$$

Observing that

$$\Big(\int_{\Gamma_+} |\beta \cdot n|(\xi^-)^2\Big)^{\frac{1}{2}} \leq |||\xi|||$$

and

$$\Big(\int_{\Gamma_+} |\beta \cdot n|(\eta^-)^2\Big)^{\frac{1}{2}} \leq \Big(\sum_K \int_{\partial K} |\beta \cdot n|(\eta^-)^2\Big)^{\frac{1}{2}}$$

Lemma(5.7) can be applied, so that

$$\begin{aligned}
(|\beta \cdot n|\eta^-, \xi^-)_{\Gamma_+} &\leq \Big(\int_{\Gamma_+} |\beta \cdot n|(\eta^-)^2\Big)^{\frac{1}{2}} \Big(\int_{\Gamma_+} |\beta \cdot n|(\xi^-)^2\Big)^{\frac{1}{2}} \\
&\leq \Big(\sum_K C_K(d,s,\beta) \frac{h_K^{s-\frac{1}{2}}}{(N+1)^{s-\frac{1}{2}}} |u|_{H^s(K)}\Big) |||\xi|||
\end{aligned}$$

Secondly, using Lemma(5.7) and applying the Cauchy-Schwarz inequality a second time leads to

$$\begin{aligned}
\sum_K (|\beta \cdot n|\eta^-, \xi^- - \xi^+)_{\partial K_- \setminus \Gamma} &\leq \Big(\sum_K \int_{\partial K_-} |\beta \cdot n|(\eta^-)^2\Big)^{\frac{1}{2}} \Big(\sum_K \int_{\partial K_-} |\beta \cdot n|[\xi]^2\Big)^{\frac{1}{2}} \\
&\leq \Big(\sum_K \int_{\partial K} |\beta \cdot n|(\eta^-)^2\Big)^{\frac{1}{2}} |||\xi||| \\
&\leq \Big(\sum_K C_K(d,s,\beta) \frac{h_K^{s-\frac{1}{2}}}{(N+1)^{s-\frac{1}{2}}} |u|_{H^s(K)}\Big) |||\xi|||
\end{aligned}$$

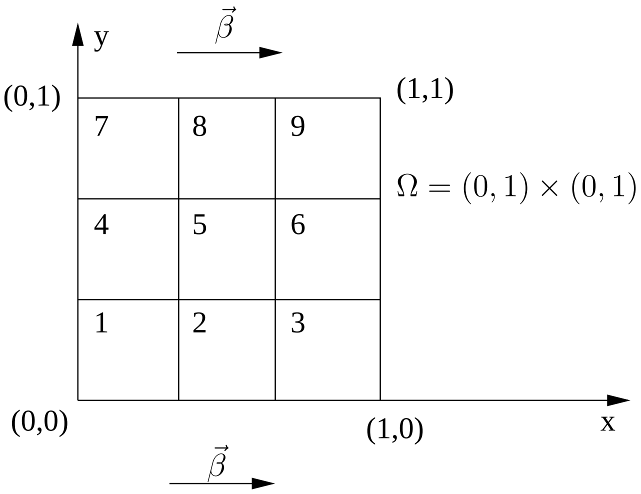

Figure 3: Order to compute the approximation on the elements

so that

$$|||\xi|||^2 \leq \Big(\sum_K C_K(d,s,\beta)\frac{h_K^{s-\frac{1}{2}}}{(N+1)^{s-\frac{1}{2}}}|u|_{H^s(K)}\Big)\ |||\xi|||$$

and

$$|||\xi||| \leq \sum_K C_K(d,s,\beta)\frac{h_K^{s-\frac{1}{2}}}{(N+1)^{s-\frac{1}{2}}}|u|_{H^s(K)}$$

In addition, by lemma(5.6)

$$|||\eta||| \leq C_K(d,s,\beta)\frac{h_K^{s-\frac{1}{2}}}{N^{s-\frac{1}{2}}}|u|_{H^s(K)}$$

so that

$$\begin{aligned}
|||u-u_h||| &\leq |||\xi||| + |||\eta||| \\
&\leq \sum_K C_{1,K}(d,s,\beta)\frac{h_K^{s-\frac{1}{2}}}{(N+1)^{s-\frac{1}{2}}}|u|_{H^s(K)} + C_{2,K}(d,s,\beta)\frac{h_K^{s-\frac{1}{2}}}{N^{s-\frac{1}{2}}}|u|_{H^s(K)} \\
&\leq \sum_K C_K(d,s,\beta)\frac{h_K^{s-\frac{1}{2}}}{N^{s-\frac{1}{2}}}|u|_{H^s(K)} \\
&\leq C(d,s,\beta)\frac{h^{s-\frac{1}{2}}}{N^{s-\frac{1}{2}}}|u|_{H^s(\Omega)}
\end{aligned}$$

□ Theorem(5.8)

6 Numerical Results

In this section, the Discontinuous Galerkin method is tested numerically. The principal aim is to verify the convergence result, developed in the previous section.

6.1 The code

The Discontinuous Galerkin method is implemented as Matlab code. You will find it in Appendix A.
On each element, the linear system (12) is solved. As finite dimensional space V_h, the polynomial space $Q^N(K)$, the set of all tensor-product polynomials on K of degree N in each coordinate direction is chosen. As basis of the polynomial space $Q^N(K)$, the tensor product of the Legendre's polynomials is chosen. This means that each basis $\varphi_i(x, y)$ is the product of a Legendre's polynomial in x and one in y, each of maximum degree N. The Legendre's polynomials belong to Jacobi polynomials, they are orthogonal for the non weighted L^2-scalar product. For more details about the Legendre's polynomials, see section(8).
For computing the matrix A and the vector $\vec{f}$, symoblic calculation is used. In fact, Matlab can use the Maple commands. This allows us to calculate the integrals exactly and no numerical integration is needed. For more details about Matlab using Maple, see the Matlab documentation. To solve the linear system, the GMRes-Algorithm is used.
Then, in the case of the test problems, $|||u - u_h|||$ is calculated where u denotes the exact solution defined by (24) and u_h is the corresponding approximation.

6.2 Test Problem

A two-dimensional problem is chosen. The domain is the rectangle $\Omega = (0, 1) \times (0, 1) \subset \mathbb{R}^2$ and as exact solution, the following function is taken

$$u(x, y) = \big((x - 1) + e^{-x}\big)y. \tag{24}$$

$\vec{\beta}$ is chosen as $\vec{\beta} = (1 \;\; 0)^T$, so that the differential equation is

$$\vec{\beta} \cdot \nabla u(x, y) + u(x, y) = xy = f(x, y)$$

Ω is divided in N^2 sub-rectangles (the elements), where N is the number of subintervals of $[0, 1]$. The order to compute the approximation on the elements is illustrated in Figure(3). The boundary condition on Γ_- is

$$u(\vec{x}) = 0 \qquad \forall\, \vec{x} \in \Gamma_-$$

It is obvious that $u \in C^\infty(\Omega)$ and $u \in H^k(\Omega)\ \forall k \geq 0$. Due to this regularity of the solution u, $s = \min(N + 1, k) = N + 1$ and the convergence result is in this case

$$|||u - u_h||| \leq C(d, N, \beta)\frac{h^{N+\frac{1}{2}}}{N^{N+\frac{1}{2}}}|u|_{H^{N+1}(\Omega)} \tag{25}$$

Note that the constant C is also depending on N. In figure(4) we plot the logarithm of the DG-norm against the logarithm of h resp. $\frac{1}{N}$ for each fixed N resp. h. These two graphics are commented in the two following subsections.

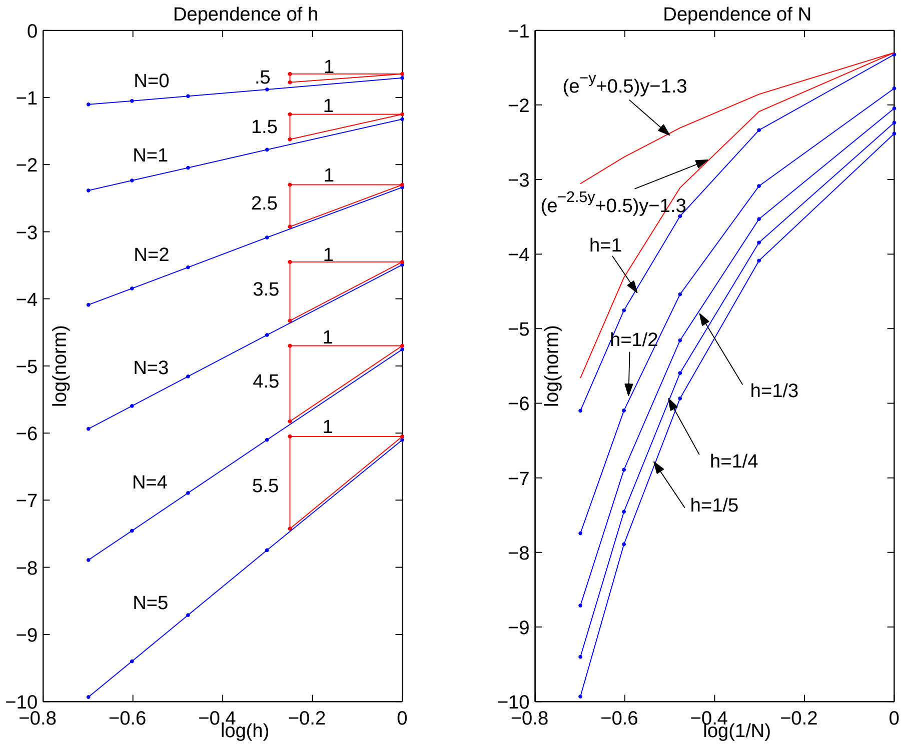

Figure 4: h- resp. N-refinement in the case of the test problem

6.2.1 h-refinement

In this case, we study the behavior of the error while varying h for each fixed N. Fixing N in (25) leads to the following convergence in h:

$$|||u - u_h||| \leq C_N h^{N+\frac{1}{2}}$$

In the case of a such behavior of the error, means

$$E = C_N h^{N+\frac{1}{2}} \tag{26}$$

we expect a straight line in the log-log diagram. In fact, taking the logarithm of (26) leads to

$$\log(E) = (N + \frac{1}{2}) \log(h) + \log(C_N)$$

Effecting the transformation of variables

$$y = \log(h) \Leftrightarrow h = e^{-y} \tag{27}$$

leads to

$$\log(E) = (N + \frac{1}{2})y + \log(C_N).$$

Therefor the expected straight line. As you can see in the first graphic of figure(4) we observe exactly a such behavior, i.e. the slope of the straight line varies for each N. So that $|||u - u_h|||$ converge at the rate $O(h^{N+\frac{1}{2}})$ as h tends to zero for each fixed N. As you can observe, the bound is attained, means the error estimation is optimal.

6.2.2 N-refinement

In this case, h is fixed. So that the convergence result gets

$$|||u-u_h||| \leq C_{h,N} N^{-(N+\frac{1}{2})} |u|_{H^{N+1}(\Omega)}$$

We neglect the N-dependence of the constant $C_{h,N}$ and the variation of $h^{N+\frac{1}{2}}$. Assuming the following error behavior

$$E = CN^{-(N+\frac{1}{2})} |u|_{H^{N+1}(\Omega)} \tag{28}$$

we expect the following behavior in the log-log diagram

$$\log(E) = \big(e^{-y} + \frac{1}{2}\big)y + \log(C|u|_{H^{N+1}(\Omega)}\big) \tag{29}$$

In fact, taking the logarithm of (28) leads to

$$\log(E) = \big(N + \frac{1}{2}\big)\log(N^{-1}) + \log(C|u|_{H^{N+1}(\Omega)}\big)$$

Effecting the same transformation of variables as above, i.e. effecting the transformation defined by (27), leads to (29). This curve is plotted in the second graphic of figure(4).
Motivated by this calculus, we model the error by the following function.

$$f(y) = (e^{-C_1 y} + \frac{1}{2})y + C_2.$$

As you can remark, a supplementary constant C_1 is added and the N-dependence of $|u|_{H^{N+1}(\Omega)}$ is neglected. In the second graphic of figure(4), a such function with $C_1 = 2.5$ and $C_2 = -1.3$ is also plotted.
As predicted, we can also observe the exponential decrease of the error. Note that the exponential decrease is only due to the regularity of the solution. For example, in the case where $u \in H^1(\Omega)$, we have that $s = min(N+1, 1) = 1, \forall N \geq 0$ and hence the convergence result becomes

$$|||u-u_h||| \leq C(d,\beta)\frac{h^{\frac{1}{2}}}{N^{\frac{1}{2}}}|u|_{H^1(\Omega)} \tag{30}$$

i.e. a convergence rate of $O\big((N^{-1})^{\frac{1}{2}}\big)$ of $|||u-u_h|||$ as N^{-1} tends to zero.

6.3 Conclusions of the numerical tests

In both cases, the h-refinement and N-refinement, the error estimations are optimal.
In the case where the solution is sufficient regular, it is more useful to refine in N than in h, because of the exponential decrease of the error in the N-refinement. In the case of a weak regularity of the solution, we have no numerical results, but a behavior of the error described in (30) is expected.

7 Conclusion

In this section, the advantages and disadvantages of the Discontinuous Galerkin method are discussed. An advantage is that the global problem can be decoupled in a lot of small problems. For each element, a linear system has to be solved.
In addition to that, it is also easy to use spectral methods. Since the approximation does not have to be continuous, one can take an arbitrary basis of polynomials for each element. But the fact that the approximation is discontinuous implies also a larger amount of unknowns. One should compare in a further project the Continuous and Discontinuous Galerkin method in the case of the same degree of freedom.
The fact that the approximation is discontinuous can be considered as an advantage or a disadvantage. If the exact solution is continuous, then the approximation will be discontinuous, but converge to the continuous solution. On the other hand if the exact solution is discontinuous, and with a good choice of the computational mesh one may catch better the effects of the disconinuous solution than with the Continuous Galerkin method.
It should also be mentioned, that no oscillations are observed.

8 Introduction to Legendre's polynomials

This section is a short introduction to the family of Legendre's polynomials. The proofs of the theorems are not presented.

Definition 8.1 *Legendre's polynomial of degree N is the polynomial defined by*

$$P_N(x) = \frac{1}{N!2^N}\frac{d^N}{dx^N}(x^2-1)^N \qquad \forall x \in [-1,1]$$

Proposition 8.2 *P_N has exactly N different real zeros all included in* $(-1,1)$

Definition 8.3 *The zeros of the Legendre's polynomial are called the Gauss points.*

Theorem 8.4 1. *The Legendre's polynomials $P_1,\ldots,P_N$ build a basis of $P^N([-1,1])$*

2. *The Legendre's polynomials satisfies*

$$\int_{-1}^{1} P_n(x)P_m(x)dx = \begin{cases} 0 & \text{if} \quad n \neq m \\ \frac{2}{2n+1} & \text{if} \quad n = m \end{cases}$$

Note that if an orthogonal basis of $P^N([-1,1])$ is given, it is simple to construct an orthogonal basis of $P^N([a,b])$ by a simple affine transformation.

Theorem 8.5 *If $f : [a,b] \to \mathbb{R}$ is a continuous function and let be $\{\Phi_0,\ldots,\Phi_N\}$ a set of orthonormal polynomials on $[a,b]$ where Φ_k is of degree k. Let be*

$$c_j = \int_a^b f(x)\Phi_j(x)dx$$

Then, the polynomial defined by

$$P = \sum_{j=0}^{N} c_j\Phi_j(x)$$

is such that

$$\|f-P\|_{L^2([a,b])} \leq \|f-Q\|_{L^2([a,b])} \qquad \forall Q \in \ P^N([a,b])$$

9 Introduction to numerical integration

This section is an introduction to numerical integration. Like in the previous section, only the results are presented.

Problem Let be $f \in C([a,b])$. One would like to approximate the functional

$$I(f) = \int_a^b f(x)\rho(x)dx \qquad \forall f \in C([a,b]) \tag{31}$$

where ρ is a weight function. Let be $a \leq x_0 < x_1 < \ldots < x_N \leq b$ $N+1$ points in $[a,b]$.

Definition 9.1

$$I_{N+1}(f) = \sum_{j=0}^{N} f(x_j)\omega_j \tag{32}$$

is called a quadrature formula where $\omega_j \in \mathbb{R}$ are called the weights. $\{x_0, \ldots, x_N\}$ are called the integration points.

Definition 9.2 *A quadrature formula J is called exact of degree N if*

- $J(f) = I(f) \qquad \forall f \in P^N$
- $J(x^{N+1}) \neq I(x^{N+1})$

Proposition 9.3 *Let be I defined by (31) and I_{N+1} by (32). This formula is exact of degree $2N+1$ if*

- *$\omega_i = \int_a^b L_i(x)\rho(x)$ where L_i is the i-th component of the Lagrange basis*
- *The integration points are the zeros of P_{N+1}, where P_{N+1} belongs to the family of orthogonal polynomials on $[a,b]$.*

Proposition 9.4 *Let be $x_0, \ldots, x_N$ $N+1$ distinct points and*

$$\omega_j = \int_a^b L_j(x)\rho(x)dx$$

Then I_{N+1} is at least exact of degree N for all $f \in C^{N+1}$.

Acknowledgements

I like to thank to Prof. Quarteroni and Erik Burman who gave me the opportunity to realize this semester project. A special thanks to Erik, who was always there for questions and spent a lot of time for explaining and discussing.

References

[1] A. Ern, J.-L. Guermond: *Elément finis: théorie, applications, mise en œuvre*, Springer (2002)

[2] P. Houston, Ch. Schwab, E. Süli: *Discontinuous hp-finite element methods for advection-diffusion-reaction problems*, 2002
http://e-collection.ethbib.ethz.ch/show?type=incoll&nr=301

10 Appendix A

```
function Norms=DG(a,b,Nx,Ny,mu,N,f,gcoeff)

% test arguments
s=size(gcoeff);
if (Ny ˜= s(1))
    errordlg(’N+1 ˜= length(gcoeff)’);
end
if (N+1 ˜= s(2))
    errordlg(’N+1 ˜= length(gcoeff)’);
end

% load maple orthogonal polynomials
maple(’with’,’orthopoly’)

% define constants
Beta=[1 0]’;
hx=a/Nx;
hy=b/Ny;
Nel=Ny*Nx;
if N<2
    numb=1;
else
    numb=5;
end
h1 = hx/numb;
h2 = hy/numb;
counter=1;

% Get order of triangles and other bouandary informations
% seq(1,:) = numberof element
% seq(2,:) = x-coordinate
% seq(3,:) = y-coordinate
% seq(4,:) = number of the element which is on the left (if on the boundary=-1)
% seq(5,:) = if on the bottom-boundary
% seq(6,:) = if on the top-boundary
% seq(7,:) = if on the right-boundary
seq=zeros(5,Nel);
seq(1,1:Nel) = [1:1:Nel];
if(Nel>1)
    for i=0:Nel-1
        % x0
        seq(2,i+1) = mod(i,Nx)*hx;
        % y0
        seq(3,i+1) = (i-mod(i,Nx))/Nx*hy;
        % previous
        if (mod(i+1,Nx)==1)
            seq(4,i+1) = -1;
        else
            seq(4,i+1) = i;
        end
        if i<Nx
            seq(5,i+1) = 1;
        end
        if i>=(Nel-Nx)
            seq(6,i+1) = 1;
        end
        if mod(i+1,Nx)==0
            seq(7,i+1) = 1;
        end
    end
else
    seq(4,1) = -1;
    seq(5,1) = 1;
    seq(6,1) = 1;
    seq(7,1) = 1;
end

% define matrices
U = zeros((N+1)ˆ2,Nel);
A = zeros((N+1)ˆ2,(N+1)ˆ2);
Ak = zeros((N+1)ˆ2,(N+1)ˆ2);
Bk = zeros((N+1)ˆ2,(N+1)ˆ2);
```

```
Ck = zeros((N+1)ˆ2,(N+1)ˆ2);
Aint = zeros(N+1,N+1);
Ufunc = cell(Nel,1);

for i=0:N
    Aint(i+1,i+1) = 2/(2*i+1);
end

% for the norm
if (strcmp(f,’x*y’) | strcmp(f,’(10*cos(10*x)+sin(10*x))*yˆ3’) & mu==1)
    if (strcmp(f,’x*y’))
        exactsol = ’((x-1)+exp(-x))*y’;
    else
        exactsol = ’(sin(10*x))*yˆ3’;
    end
    Lnorm2=0;
    DGnorm2=0;
else
    ’error norm calculation is not possible, exact solution is not known’
    %warndlg(’error norm calculation is not possible, exact solution is not known’);
end

% Construction of Ak
Tk = [[hx/2 0];[0 hy/2]];
for i=0:N
    for j=0:N
        for k=0:N
            for l=0:N
                Iind = i*(N+1)+j+1;
                Jind = k*(N+1)+l+1;
                % Ak
                Ak(Iind,Jind) = mu*det(Tk)*Aint(i+1,k+1)*Aint(j+1,l+1);
                % Bk
                Bk(Iind,Jind) = det(Tk)*dot(Beta,inv(Tk)’*...
                    [Aint(j+1,l+1)*getInt(i,k); Aint(i+1,k+1)*getInt(j,l)]);
                % Ck
                val1 = str2num(maple(’P’,i,-1));
                val2 = str2num(maple(’P’,k,-1));
                Ck(Iind,Jind) = hy/2*val1*val2*Aint(j+1,l+1);
            end
        end
    end
end
Ak = Ak + Bk + Ck;

% loop on the elements
for k=1:Nel
    % get f
    bk = [seq(2,k)+hx/2 seq(3,k)+hy/2]’;
    x0 = seq(2,k);
    x0str = maple(’eval’,x0);
    x0strPlus = maple(’eval’,x0+hx);
    y0 = seq(3,k);
    y0str = maple(’eval’,y0);
    y0strPlus = maple(’eval’,y0+hy);
    invTransX = [’2/(’ maple(’eval’,hx) ’)*(x-’ maple(’eval’,x0) ’)-1’];
    invTransY = [’2/(’ maple(’eval’,hy) ’)*(y-’ maple(’eval’,y0) ’)-1’];
    if (seq(4,k)==-1)
        coeff = gcoeff(counter,:);
    else
        coeff = outflow(:,seq(4,k));
    end
    for i1=0:N
        % First Term
        pi1 = maple(’P’,i1,invTransX);
        func = [’(’ f ’)*(’ pi1 ’)’];
        str = maple(’int’,func,[’x=’ x0str ’..’ x0strPlus]);
        % Second Term
        pi1atmin1 = str2num(maple(’P’,i1,-1));
        for i2=0:N
            % First Term
            pi2 = maple(’P’,i2,invTransY);
            func = [’(’ pi2 ’)*(’ str ’)’];
            str2 = maple(’int’,func,[’y=’ y0str ’..’ y0strPlus]);
            % Second Term
            second = abs(Beta’*[-1 0]’)*hy/2*pi1atmin1*coeff(i2+1)*2/(2*i2+1);
            % Sum
            fk(i1*(N+1)+i2+1) = str2num(str2) + second;
        end
    end

    if (seq(4,k)==-1)
        counter=counter+1;
    end

    % get solution
    U(:,k) = gmres(Ak,fk’,20,10e-15,500);%Ak\fk’;

    % Get grid for evaluation
    x = [x0:h1:x0+hx];
    y = [y0:h2:y0+hy];
    [X(:,:,k) Y(:,:,k)] = meshgrid(x,y);

    % get outflow
    outflow(:,k) = zeros(N+1,1);

    % get solution function in form of string
    % loop over x-basis
    for i=1:N+1
        % loop over y-basis
        for j=1:N+1
            % get func
            if (i==1 & j==1)
                Ufunc(k) = {[’(’ num2str(U(1,k),15) ’*(’ maple(’P’,0,invTransX)...
                            ’)*(’ maple(’P’,0,invTransY) ’))’]};
            else
                Ufunc(k) = {[char(Ufunc(k)) ’+(’ num2str(U((i-1)*(N+1)+j,k),15)...
                            ’*(’ maple(’P’,i-1,invTransX) ’)*(’ maple(’P’,j-1,invTransY) ’))’]};
            end
            % get i-th outflow-coeff for element k
            outflow(i,k) = outflow(i,k) + str2num(maple(’P’,j-1,1))*U((j-1)*(N+1)+i,k);
        end
    end

    % evaluate solution function on grid points for visualization
    s = size(X(:,:,k));
    for i=1:s(1)
        for j=1:s(2)
            x=X(i,j,k);
            y=Y(i,j,k);
            v(i,j,k) = eval(char(Ufunc(k)));
        end
    end

    % calculates norms of this element
    if (strcmp(f,’x*y’) | strcmp(f,’(10*cos(10*x)+sin(10*x))*yˆ3’) & mu==1)
        if (strcmp(f,’x*y’))
            Lnorm2 = Lnorm2 + L2norm(x0,y0,hx,hy,char(Ufunc(k)),exactsol)ˆ2;
            DGnorm2 = DGnorm2 + pureDGnorm(x0,y0,hx,hy,Ufunc,seq(:,k),k,exactsol)ˆ2;
        else
            Lnorm2 = Lnorm2 + L2norm(x0,y0,hx,hy,char(Ufunc(k)),exactsol)ˆ2;
            DGnorm2 = DGnorm2 + pureDGnorm(x0,y0,hx,hy,Ufunc,seq(:,k),k,exactsol)ˆ2;
        end
    end
end

% norm calculation
if (strcmp(f,’x*y’) | strcmp(f,’(10*cos(10*x)+sin(10*x))*yˆ3’) & mu==1)
    Norms(1) = sqrt(Lnorm2);
    Norms(2) = sqrt(DGnorm2);
    Norms(3) = sqrt(DGnorm2+Lnorm2);
else
    Norms = [-1 -1 -1];
end

% visualisation
figure(33)
for k=1:Nel
```

```
    patch(surf2patch(X(:,:,k)’,Y(:,:,k)’,v(:,:,k)’,v(:,:,k)’));
    shading faceted;
    view(3)
end
hold off;
```