CHALMERS
FINITE ELEMENT CENTER

PREPRINT 2000-08

Conservation Properties for the Continuous and Discontinuous Galerkin Methods

Mats G. Larson and A. Jonas Niklasson

Chalmers Finite Element Center
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg Sweden 2001

CHALMERS FINITE ELEMENT CENTER

Preprint 2000-08

Conservation Properties for the Continuous and Discontinuous Galerkin Methods

Mats G. Larson and A. Jonas Niklasson

CHALMERS

Chalmers Finite Element Center
Chalmers University of Technology
SE–412 96 Göteborg Sweden
Göteborg, January 2001

Conservation Properties for the Continuous and Discontinuous Galerkin Methods
Mats G. Larson and A. Jonas Niklasson
NO 2000-08
ISSN 1404–4382

Chalmers Finite Element Center
Chalmers University of Technology
SE–412 96 Göteborg
Sweden
Telephone: +46 (0)31 772 1000
Fax: +46 (0)31 772 3595
www.phi.chalmers.se

Printed in Sweden
Chalmers University of Technology
Göteborg, Sweden 2001

Conservation Properties for the Continuous and Discontinuous Galerkin Methods

Mats G. Larson * † and A. Jonas Niklasson ‡

January 29, 2001

Abstract

In this paper we investigate the relationship between the continuous and the discontinuous Galerkin methods for elliptic problems. In particular, we show that the continuous Galerkin method can be interpreted as the limit of a discontinuous Galerkin method when a stabilization parameter tends to infinity. Based on this observation we derive a method for computing a conservative approximation of the flux on the the boundary of each element for the continuous Galerkin method. The conservative flux is then obtained by actually computing the limit of the natural conservative flux provided by the discontinuous Galerkin method. We prove existence, uniqueness, and optimal order error estimates. Finally, we illustrate our results by a few numerical examples.

1 Introduction

Recently there has been a growing interest in the discontinuous Galerkin (dG) method for elliptic problems. One of the motivations for this interest is the fact that the dG method manufactures a natural flux function, defined on each edge in the triangulation of the domain, which satisfies an elementwise conservation law, a property often desired in applications which the continuous Galerkin (cG) method does not posses.

In this paper we investigate the relationship between the cG and dG method. In particular we show that the cG method may be viewed as the limit of a stabilized dG method when the stabilizing parameter tends to infinity. The stabilization corresponds to penalizing discontinuities. Using this fact a conservative flux for the cG method is naturally

*Corresponding author, Department of Mathematics, Chalmers University of Technology, Göteborg, S-412 96, Sweden, mgl@math.chalmers.se
†Supported by the Swedish Foundation for Strategic Research and the Swedish Council for Engineering Research
‡Divison of Mechanics, Chalmers University of Technology, Göteborg, S-412 96, Sweden, jnik@mec.chalmers.se

obtained by calculating the limit of the natural conservative flux in the dG method as the stabilization parameter tends to infinity. The limit is a correction to the average of the right and left values of the fluxes, which is found by solving a linear symmetric positive definite system of equations. The number of unknowns, in the simplest version, is equal to the number of elements in the triangulation.

The dG method we consider is the classical Nitsche method, see Nitsche [10], on a space of discontinuous piecewise polynomials, which at least contains the continuous piecewise polynomials of degree p and piecewise elementwise constant functions. In fact, the elementwise conservation property emanates from the presence of piecewise constant functions in the testspace and thus enriching the continuous test and trial spaces with piecewise constants produces a minimal conservative dG method with optimal order.

We analyze the limit of the dG method and show existence and uniqueness. Further we derive error estimates of optimal order, which are uniform in the stabilization parameter, using standard techniques.

Conservative fluxes are used, for instance, in a posteriori error estimation to manufacture boundary conditions for local elementwise Neumann problems, see Kelly [8], Ladeveze [9], and Ainsworth and Oden [1], [2]. Another common application is to calculate a certain weighted integral of the flux. This is usually done using a postprocessing procedure which, in the best case give a superconvergent approximation, see for instance [3]. Direct computation of the weighted integrals using the flux proposed here yields exactly the same results as the standard technique.

The conservative structure of the cG method has also been discussed in Hughes et al [6], where a certain nodal conservation property is introduced.

The paper is organized as follows. In Section 2 we present a model problem, the cG method, discuss the local conservation property, and we also give a preliminary derivation of the conservative flux. In Section 3 we analyze the aforementioned limit of the dG method and derive the conservative flux from this perspective. We also show a priori error estimates for the dG solution and the flux which are uniform in the stabilizing parameter. The analytical results are illustrated by numerical examples.

2 The cG method

2.1 Model problem

We consider the following boundary value problem: find $u : \Omega \to \mathbf{R}$ such that

$$\begin{aligned} -\nabla \cdot \sigma(u) &= f \quad \text{in } \Omega, \\ u &= g_D \quad \text{on } \Gamma_D, \\ \sigma_n(u) &= g_N \quad \text{on } \Gamma_N, \end{aligned} \tag{2.1}$$

where Ω denotes a bounded domain in $\mathbf{R}^d$, $d = 1, 2$, or 3, with boundary $\Gamma = \Gamma_D \cup \Gamma_N$, and the normal flux is defined by $\sigma_n(u) = n \cdot \sigma(u)$, where n is the unit outward normal of

Γ and

$$\sigma(u) = A\nabla u, \tag{2.2}$$

with A a uniformly positive definite $d \times d$ matrix with bounded entries $a_{ij} \in C(\Omega)$. Using the notation

$$a_c(v, w) = (\sigma(v), \nabla w), \tag{2.3}$$

$$l_c(v) = (f, v) + (g_N, v)_{\Gamma_N}, \tag{2.4}$$

for all $v, w \in V = \{v \in H^1 : v = g_D \text{ on } \Gamma_D\}$, we may formulate the weak version of (2.1): find $u \in V$ such that

$$a_c(u, v) = l_c(v) \quad \text{for all } v \in V. \tag{2.5}$$

Since A is uniformly positive definite and the entries are bounded there are constants c and C such that

$$c\|\nabla v\|^2 \leq a_c(v, v) \leq C\|\nabla v\|^2 \quad \text{for all } v \in V. \tag{2.6}$$

As is well known (2.5) has a unique solution $u \in V$ for each $f \in H^{-1}$, $g_D \in H^{-1/2}(\Gamma_D)$, and $g_N \in H^{1/2}(\Gamma_N)$, for $\Gamma_D \neq \emptyset$, and if $\Gamma_D = \emptyset$, the solution exists and is unique up to a constant, i.e., $u \in V/\mathbf{R}$ for $f \in H^{-1}$, $g_N \in H^{1/2}(\Gamma)$, and the compatibility condition

$$\int_\Omega f + \int_\Gamma g_N = 0.$$

is satisfied.

2.2 The mesh

To define the numerical methods we introduce a partion $\mathcal{K} = \{K\}$ of Ω called the mesh. For simplicity only we assume that the mesh is quasiuniform with meshsize h, see [4].

The set of edges in the mesh is denoted by $\mathcal{E} = \{E\}$ and we split $\mathcal{E}$ into three disjoint subsets

$$\mathcal{E} = \mathcal{E}_I \cup \mathcal{E}_D \cup \mathcal{E}_N,$$

where $\mathcal{E}_I$ is the set of edges in the interior of Ω, $\mathcal{E}_D$ is the set of edges on the Dirichlet part of the boundary Γ_D, and $\mathcal{E}_N$ is the set of edges in the Neumann part of the boundary Γ_N.

To each edge we associate a fixed unit normal n_E, such that on the boundary Γ, n_E is the outward unit normal. We also use the notation n_K for the outward normal of an element K.

2.3 The cG method

Let $\mathcal{V}_c^p = \mathcal{V}_c^p(\psi)$ denote the space of continuous piecewise polynomials of degree p defined on $\mathcal{K}$, which are equal to ψ on Γ_D,

$$\mathcal{V}_c^p(\psi) = \{v \in C(\Omega) : v|_{\Gamma_D} = \psi, v|_K \in \mathcal{P}_p(K), K \in \mathcal{K}\}, \tag{2.7}$$

where $\mathcal{P}_p(K)$ is the space of polynomials of degree p defined on K. In this note we will be concerned with two cases: $\psi = 0$ for the test space and $\psi = g_D$ for the trial space. We usually write $\mathcal{V}_c^p$ for brevity. The cG method reads: find $U_c \in \mathcal{V}_c^p$ such that

$$a_c(U_c, v) = l_c(v) \quad \text{for all } v \in \mathcal{V}_c^p. \tag{2.8}$$

2.4 The conservation property

Let $\omega \subset \Omega$ be a subdomain of Ω, and χ_ω be the indicator function χ_ω, defined by $\chi_\omega = 1$ on ω and 0 on $\Omega \setminus \omega$. Multiplying (2.1) by χ_ω and integrating by parts yields the conservation law

$$\int_\omega f + \int_{\partial\omega} \sigma_n(u) = 0. \tag{2.9}$$

Note that $\sigma_n(u) = g_N$ on Γ_N. This is the fundamental conservation property which we seek to mimic in the discrete case on an element level, i.e., we seek an approximate flux $\Sigma_n(U_c)$ such that $\Sigma_n(U_c) = g_N$ on Γ_N and

$$\int_K f + \int_{\partial K} \Sigma_{n_K}(U_c) = 0, \tag{2.10}$$

for all elements $K \in \mathcal{K}$.

2.5 A conservative flux for the cG method

Here we shall present a first derivation of the basic conservative flux. A more general version is presented in Section 3. Let v be a piecewise constant function defined on $\mathcal{K}$. We denote the jump at an interior edge $E \in \mathcal{E}_I$ by $[v] = v^+ - v^-$, where $v^\pm(x) = \lim_{t\to 0, t>0} v(x \mp n_E t), x \in E$, and $[v] = v^+$ on edges at the boundary $\mathcal{E}_D \cup \mathcal{E}_N$. Then we can state (2.10) in the form

$$\sum_{E\in\mathcal{E}} (\Sigma_{n_E}(U_c), [v])_E + \sum_{K\in\mathcal{K}} (f, v)_K = 0, \tag{2.11}$$

for all piecewise constant functions v. On each edge $E \in \mathcal{E}_N$, i.e., on the Neumann part of the boundary, we should have

$$\Sigma_{n_E}(U_c) = g_N,$$

since here the normal flux is given. Furthermore, on the remaining edges $\mathcal{E}_I \cup \mathcal{E}_D$, $\Sigma_n(U_c)$ should of course be an approximation of the exact flux $\sigma_n(u)$ of optimal order. For the cG method a natural approximation of $\sigma_n(u)$ is the average

$$\langle \sigma_n(U_c) \rangle, \tag{2.12}$$

where $\langle v \rangle = (v^+ + v^-)/2$, which possesses optimal order of convergence but not the desired local conservation property. We thus write

$$\Sigma_n(U_c) = \langle \sigma_n(U_c) \rangle - \Delta_n(U_c), \tag{2.13}$$

where $\Delta_n = \Delta_n(U_c)$ is some correction making the approximate flux conservative. Inserting this expression into (2.11) we obtain

$$\begin{aligned} \sum_{E \in \mathcal{E}_I \cup \mathcal{E}_D} (\Delta_{n_E}, [v])_E = \sum_{E \in \mathcal{E}_I \cup \mathcal{E}_D} (\langle \sigma_{n_E}(U_c) \rangle, [v])_E \\ + \sum_{E \in \mathcal{E}_N} (g_N, v)_E + \sum_{K \in \mathcal{K}} (f, v)_K. \end{aligned} \tag{2.14}$$

This equation suggests that a natural choice of Δ_{n_E} is

$$\Delta_{n_E} = h^{-1}[V], \tag{2.15}$$

for some piecewise constant function V, to be determined. The scaling with h is motivated by consistency of units. With this choice of V we obtain the problem: find V such that

$$\begin{aligned} \sum_{E \in \mathcal{E}_I \cup \mathcal{E}_D} (h^{-1}[V], [v])_E = \sum_{E \in \mathcal{E}_I \cup \mathcal{E}_D} (\langle \sigma_{n_E}(U_c) \rangle, [v])_E \\ + \sum_{E \in \mathcal{E}_N} (g_N, v)_E + \sum_{K \in \mathcal{K}} (f, v)_K, \end{aligned} \tag{2.16}$$

for all piecewise constant v. This is a linear symmetric system of equations with the number of unknowns equal to the number of elements in the mesh. Note that if $\mathcal{E}_D \neq \emptyset$ then (2.16) is a positive definite system of equations, and thus there exists a unique solution. If $\mathcal{E}_D = \emptyset$ then there exist a solution which is unique up to a constant, and thus the jump $[V]$ is uniquely determined on $\mathcal{E}_I$. See the proof of Theorem 3.1 for details of this argument. Further, setting $v = 1$ on K and 0 on $\Omega \setminus K$, in the righthand side we obtain

$$\begin{aligned} \sum_{E \in \mathcal{E}_I \cup \mathcal{E}_D} (\langle \sigma_{n_E}(U_c) \rangle, [v])_E + \sum_{E \in \mathcal{E}_N} (g_N, v)_E + \sum_{K \in \mathcal{K}} (f, v)_K \\ = (\langle \sigma_{n_K}(U_c) \rangle, 1)_{\partial K \setminus \Gamma_N} + (g_N, 1)_{\partial K \cap \Gamma_N} + (f, 1)_K, \end{aligned} \tag{2.17}$$

i.e., the residual of the average flux approximation when inserted into the element conservation law. Solving (2.16) a conservative flux may be directly computed using the formula

$$\Sigma_{n_E}(U_c) = \begin{cases} \langle \sigma_{n_E}(U_c) \rangle - h^{-1}[V] & E \in \mathcal{E}_I \cup \mathcal{E}_D, \\ g_N \quad E \in \mathcal{E}_N. \end{cases} \tag{2.18}$$

3 The cG method as a limit of the dG method

Starting from the dG method we derive a more general version of the conservative flux approximation by observing that there is a natural conservative flux in the dG method and that the cG method may be viewed as a certain limit of the dG method when a parameter tends to infinity. The corresponding limit of the natural conservative flux in the dG method is a conservative flux for the cG method. This section is devoted to analyzing this limit. In particular, we prove that this limit exists uniquely and we devise a method for calculating the limit. We begin with some preliminaries.

3.1 Discontinuous spaces

Let $\mathcal{W}_d^p$ be a space of discontinuous piecewise polynomials defined on $\mathcal{K}$ such that

$$\mathcal{V}_c^p + \mathcal{V}_d^q \subset \mathcal{W}_d^p, \tag{3.1}$$

where $\mathcal{V}_c^p$ is the space of continuous piecewise polynomials of degree p and $\mathcal{V}_d^q$ is the space of discontinuous piecewise polynomials of degree q. For instance, $\mathcal{W}_d^p = \mathcal{V}_d^p$ or $\mathcal{W}_d^p = \mathcal{V}_c^p + \mathcal{V}_d^0$.

3.2 The dG method

The dG method for (2.1) is defined by: find $U_{d,\beta} \in \mathcal{W}_d^p$ such that

$$a_d(U_{d,\beta}, v) + \beta b_d(U_{d,\beta}, v) = l_d(v) + \beta m_d(v), \tag{3.2}$$

where β is a positive parameter and the bilinear forms are defined by

$$\begin{aligned} a_d(v, w) &= \sum_{K \in \mathcal{K}} (A\nabla v, \nabla w)_K \qquad (3.3)\\ &\quad - \sum_{E \in \mathcal{E}_I \cup \mathcal{E}_D} \Big((\langle \sigma_n(v) \rangle, [w])_E + ([v], \langle \sigma_n(w) \rangle)_E \Big), \\ b_d(v, w) &= \sum_{E \in \mathcal{E}_I \cup \mathcal{E}_D} (h^{-1}[v], [w])_E, \qquad (3.4) \end{aligned}$$

and the linear functionals by

$$\begin{aligned} l_d(w) &= (f, w) + (g_N, w)_{\Gamma_N}, \qquad (3.5)\\ m_d(w) &= (g_D, h^{-1} w)_{\Gamma_D}. \qquad (3.6) \end{aligned}$$

We employed the notation

$$\langle v \rangle = \begin{cases} (v^+ + v^-)/2 & E \in \mathcal{E}_I, \\ v^+ & E \in \mathcal{E}_D, \end{cases} \tag{3.7}$$

and

$$[v] = \begin{cases} v^+ - v^- & E \in \mathcal{E}_I, \\ v^+ & E \in \mathcal{E}_D. \end{cases} \tag{3.8}$$

Here we also used the notation introduced in Subsection 2.2. Furthermore, we assume that g_D can be represented exactly by the continuous space $\mathcal{V}_c^p$.

Remark. Recently, nonsymmetric versions of this classical method has been proposed by Oden, Babuska, and Baumann in [11], and stabilized versions thereof in, for instance, Hughes et al. [7]. Our analysis below extends to the stabilized version and the same elementwise conservative approximation of the flux emanates from this method.

3.3 The local conservation property for the dG method

Introducing the discrete flux

$$\Sigma_{n_E,\beta}(U_{d,\beta}) = \begin{cases} \langle \sigma_{n_E}(U_{d,\beta}) \rangle - \beta h^{-1}[U_{d,\beta}] & E \in \mathcal{E}_I \cup \mathcal{E}_D, \\ g_N & E \in \mathcal{E}_N, \end{cases} \tag{3.9}$$

we obtain the discrete elementwise conservation law

$$\int_K f + \int_{\partial K} \Sigma_{n_K,\beta}(U_{d,\beta}) = 0, \tag{3.10}$$

for each element $K \in \mathcal{K}$. We shall now show that the limit

$$\lim_{\beta \to \infty} \Sigma_{n,\beta}(U_{d,\beta}) \tag{3.11}$$

exists and can be used for computing a conservative flux for the cG method.

3.4 The energy norm

We equip $\mathcal{W}_d^p$ with the following energy norm

$$|||v|||^2 = \|v\|_{\mathcal{K}}^2 + \|\langle \sigma_n(v) \rangle\|_{\mathcal{E}}^2 + \|h^{-1}[v]\|_{\mathcal{E}}^2, \tag{3.12}$$

where

$$\|w\|_{\mathcal{K}}^2 = \sum_{K \in \mathcal{K}} (A \nabla w, \nabla w), \tag{3.13}$$

$$\|w\|_{\mathcal{E}}^2 = \sum_{E \in \mathcal{E}_I \cup \mathcal{E}_D} \|h^{1/2} w\|_E^2. \tag{3.14}$$

Furthermore, the following inverse estimate will be useful

$$\|\langle\sigma_n(v)\rangle\|_{\mathcal{E}}^2 \leq C\|v\|_{\mathcal{K}}^2 \quad \text{for all } v \in \mathcal{W}_d^p, \tag{3.15}$$

with a constant C which depends on the degree of polynomials p but not on the meshsize h. This estimate can be shown by scaling, see [12]. Finally, we need the following interpolation error (or approximation property) estimate

$$|||u - \pi u||| \leq Ch^p\|u\|_{p+1}, \tag{3.16}$$

where $\pi u \in \mathcal{V}_c^p$ is a continuous interpolant of u, and $\|\cdot\|_{p+1}$ denotes the standard Sobolev norm. With $\eta = u - \pi u$ we have

$$|||\eta|||^2 = \|\eta\|_{\mathcal{K}}^2 + \|\langle\sigma_n(\eta)\rangle\|_{\mathcal{E}}^2,$$

since η is continuous and zero on Γ_D. Using the boundedness of A (2.6), we get $\|\eta\|_{\mathcal{K}} \leq C\|\eta\|_1$. For the second term we invoke the trace inequality

$$\|v\|_{\partial K}^2 = C\|v\|_K\left(h^{-1}\|v\|_K + \|v\|_{1,K}\right) \quad \text{for } v \in H^1(K), \tag{3.17}$$

where C is a constant independent of h, trianglewise to obtain

$$\begin{aligned}\|\langle\sigma_n(\eta)\rangle\|_{\mathcal{E}}^2 &\leq C\sum_{K\in\mathcal{K}} h\|\nabla\eta\|_K\left(h^{-1}\|\nabla\eta\|_K + \|\nabla\eta\|_{1,K}\right)\\ &\leq C\sum_{K\in\mathcal{K}} \|\eta\|_{1,K}\left(\|\eta\|_{1,K} + h\|\eta\|_{2,K}\right)\end{aligned}$$

Now (3.16) follows directly from the standard interpolation error estimate

$$\|\eta\|_{s,K} \leq Ch^{p+1-s}\|u\|_{p+1,K},$$

with $s = 1$ and $s = 2$, see [4]. The trace inequality (3.17) follows by mapping to the unit size reference element $\tilde{K}$, employing the trace inequality

$$\|v\|_{\partial\tilde{K}}^2 = C\|v\|_{\tilde{K}}\|v\|_{1,\tilde{K}} \quad \text{for } v \in H^1(\tilde{K}), \tag{3.18}$$

see [4], and finally transforming back to K.

3.5 Main results

We begin by showing an important equivalence of norms.

Lemma 3.1 *There is a positive constant C, independent of h and β, such that*

$$|||w|||_{\mathcal{W}_d^p/\mathcal{V}_c^p}^2 \leq C\|h^{-1}[w]\|_{\mathcal{E}}^2 \quad \textit{for all } w \in \mathcal{W}_d^p/\mathcal{V}_c^p, \tag{3.19}$$

where $|||\cdot|||_{\mathcal{W}_d^p/\mathcal{V}_c^p}^2$ denotes the quotient norm defined by

$$|||w|||_{\mathcal{W}_d^p/\mathcal{V}_c^p}^2 = inf_{v\in\mathcal{V}_c^p}|||w + v|||. \tag{3.20}$$

Proof. Assume that the result does not hold. Then there is a sequence $\{v_n\}$ such that $|||v_n|||^2_{\mathcal{W}^p_d/\mathcal{V}^p_c} = 1$ and $\|h^{-1}[v_n]\|_{\mathcal{E}} \leq 1/n$, for $n = 1, 2, 3, \ldots$ Since $\mathcal{W}^p_d/\mathcal{V}^p_c$ is finite dimensional, $\{v_n\}$ resides in a compact subset and we may extract a convergent subsequence with limit v_∞. But then it follows that v_∞ is a continuous function which is zero on Γ_d and thus $|||v_\infty|||^2_{\mathcal{W}^p_d/\mathcal{V}^p_c} = 0$, which gives a contradiction. Clearly C is independent of β and it follows from scaling that C is also independent of h. □

We begin by stating a useful coercivity result and corresponding error estimate for the dG method.

Lemma 3.2 *There is a constant $\beta_0 > 0$ such that for all $\beta \geq \beta_0$ the coercivity estimate*

$$c|||v|||^2 + (\beta - \beta_0)\|h^{-1}[v]\|^2_{\mathcal{E}} \leq a_d(v, v) + \beta b_d(v, v) \quad \text{for all } v \in \mathcal{W}, \tag{3.21}$$

where c is a positive constant independent of h and β.

Proof. We first have

$$a_d(v, v) + \beta b_d(v, v) = \|v\|^2_{\mathcal{K}} - 2 \sum_{E \in \mathcal{E}_I \cup \mathcal{E}_D} (\langle \sigma_n(v) \rangle, [v])_E + \beta \|h^{-1}[v]\|^2_{\mathcal{E}}.$$

Next we use the standard inequality $2ab < \epsilon a^2 + \epsilon^{-1} b^2$, for any positive ϵ, followed by the inverse trace inequality (3.15) to get

$$\begin{aligned}
-2 \sum_{E \in \mathcal{E}_I \cup \mathcal{E}_D} (\langle \sigma_n(v) \rangle, [v])_E &\geq -\epsilon \|\langle \sigma_n(v) \rangle\|^2_{\mathcal{E}} - \epsilon^{-1} \|h^{-1}[v]\|^2_{\mathcal{E}} \\
&\geq -\epsilon C \|v\|^2_{\mathcal{K}} - \epsilon^{-1} \|h^{-1}[v]\|^2_{\mathcal{E}}.
\end{aligned}$$

Choosing ϵ such that $1 - \epsilon C = c$ and taking $\beta_0 = c + \epsilon^{-1}$ and, finally, using the inverse inequality (3.15) to bound $\|\langle \sigma_n(n) \rangle\|_{\mathcal{E}}$, we obtain the coercivity estimate (3.21). □

Lemma 3.3 *The following error estimates hold*

$$|||u - U_c||| \leq C h^p \|u\|_{p+1}, \tag{3.22}$$

$$|||u - U_{d,\beta}||| \leq C h^p \|u\|_{p+1}, \tag{3.23}$$

provided $\beta \geq \beta_0$ in the second case. Here C denote positive constants independent of h and β.

Proof. To prove the error estimate for the cG method (3.22) we first use the triangle inequality

$$|||u - U_c||| \leq |||u - \pi u||| + |||\pi u - U_c|||.$$

Next, using the inverse inequality (3.15) followed by the definition of $\|\cdot\|_{\mathcal{K}}$ we get

$$\begin{aligned}|||\pi u - U_c|||^2 &\leq C\|\pi u - U_c\|^2 \\ &= Ca_c(\pi u - U_c, \pi u - U_c) \\ &= Ca_c(\pi u - u, \pi u - U_c) \\ &\leq C|||u - \pi u|||\ |||\pi u - U_c|||,\end{aligned}$$

where we used the definition of the cG method to replace U_c by u. Finally, dividing by $|||\pi u - U_c|||$ and using the interpolation error estimate (3.16) yields the estimate.

For the dG method error estimate (3.23) we again use the triangle inequality

$$|||u - U_{d,\beta}||| \leq |||u - \pi u||| + |||\pi u - U_{d,\beta}|||,$$

where $\pi u \in \mathcal{V}_c^p$ is a continuous interpolant of u. For the second term on the right hand side we start from the coercivity estimate (3.21). Using the fact that $\beta \geq \beta_0$ we obtain

$$\begin{aligned}c|||\pi u - U_{d,\beta}|||^2 &\leq a_d(\pi u - U_{d,\beta}, \pi u - U_{d,\beta}) \\ &\quad + \beta b_d(\pi u - U_{d,\beta}, \pi u - U_{d,\beta}) \\ &= a_d(\pi u - u, \pi u - U_{d,\beta}) \\ &\quad + \beta b_d(\pi u - u, \pi u - U_{d,\beta}) \\ &= a_d(\pi u - u, \pi u - U_{d,\beta}) \\ &\leq |||u - \pi u|||\ |||\pi u - U_{d,\beta}|||,\end{aligned}$$

where we used Galerkin orthogonality (3.2) to replace $U_{d,\beta}$ by u. Finally, we used that $[\pi u - u] = 0$ on $\mathcal{E}_I \cup \mathcal{E}_D$ and thus the term $\beta b_d(\pi u - u, \pi u - U_{d,\beta})$ vanished. As a consequence, the right hand side is independent of β. Finally, dividing by $|||\pi u - U_{d,\beta}|||$ and using the interpolation error estimate (3.16) completes the proof. $\square$

Proposition 3.1 *For all $\beta \geq \beta_0$ we have*

$$|||U_c - U_{d,\beta}||| \leq \frac{C}{\beta - \beta_0} h^p \|u\|_{p+1}, \tag{3.24}$$

where C is positive constants independent of h and β.

Remark. Note that in order to prove the coercivity estimate (3.21) in Lemma 3.2 we could in fact replace $b_d(\cdot,\cdot)$ by the weaker term

$$b_d(v, w) = \sum_{E \in \mathcal{E}_I \cup \mathcal{E}_D} (h^{-1} P_{p-1}[v], P_{p-1}[w])_E, \tag{3.25}$$

where P_{p-1} denotes the obvious weighted L^2 projection onto polynomials of order $p-1$ defined on an edge. For triangular meshes and piecewise constant coefficients we obtain the standard nonconforming methods with functions which are continuous at the Gauss points on each edge. For instance, for $p = 1$ and constant A in (2.2) we obtain the space of piecewise linear functions which are continuous at the midpoints of the edges.

Remark. The existence of an interpolation operator with optimal order approximation such that $b(\pi u, v) = 0$ for all $v \in \mathcal{W}_d^p$ is crucial for our results. If for instance, we consider piecewise linear approximation on quadrilaterals, we obtain $\lim_{\beta\to\infty} U_{d,\beta} = 0$. For related results in the context of nearly incompressible materials we refer to Hansbo and Larson [5].

Proof. We begin with an estimate of $|||U_c - U_{d,\beta}|||$. Starting from the coercivity estimate (3.21) and subtracting $(\beta-\beta_0)\|h^{-1}[U_c - U_{d,\beta}]\|_{\mathcal{E}}^2 = (\beta-\beta_0)b_d(U_c - U_{d,\beta}, U_c - U_{d,\beta})$ on both sides we get

$$\begin{aligned} c|||U_c - U_{d,\beta}|||^2 &\leq a_d(U_c - U_{d,\beta}, U_c - U_{d,\beta}) \\ &\quad + \beta_0 b_d(U_c - U_{d,\beta}, U_c - U_{d,\beta}) \\ &\leq a_d(U_c - U_{d,\beta}, U_c - U_{d,\beta} + v) \\ &\quad + \beta_0 b_d(U_c - U_{d,\beta}, U_c - U_{d,\beta} + v) \\ &\leq |||U_c - U_{d,\beta}|||\;|||U_c - U_{d,\beta}|||_{\mathcal{W}_d^p/\mathcal{V}_c^p}, \end{aligned}$$

where we used (2.8) and (3.2) to insert an arbitrary function v in $\mathcal{V}_c^p$, which is zero on Γ_D. Finally we used the Cauchy Schwarz inequality. Dividing by $|||U_c - U_{d,\beta}|||$, and invoking Lemma 3.1 we obtain

$$\begin{aligned} |||U_c - U_{d,\beta}||| &\leq |||U_c - U_{d,\beta}|||_{\mathcal{W}_d^p/\mathcal{V}_c^p} \\ &\leq \|h^{-1}[U_c - U_{d,\beta}]\|_{\mathcal{E}}. \end{aligned} \tag{3.26}$$

Next, in order to estimate $\|h^{-1}[U_c - U_{d,\beta}]\|_{\mathcal{E}}$, we again start from the coercivity estimate (3.21) with $v = U_c - U_{d,\beta}$. We neglect the positive term $c|||U_c - U_{d,\beta}|||^2$ on the left hand side

$$\begin{aligned} (\beta-\beta_0)\|h^{-1}[U_c - U_{d,\beta}]\|_{\mathcal{E}}^2 &\leq a_d(U_c - U_{d,\beta}, U_c - U_{d,\beta}) \\ &\quad + \beta b_d(U_c - U_{d,\beta}, U_c - U_{d,\beta}) \\ &= a_d(U_c - u, U_c - U_{d,\beta}) \\ &\quad + \beta b_d(U_c - u, U_c - U_{d,\beta}) \\ &= a_d(U_c - u, U_c - U_{d,\beta} + v) \\ &\leq |||u - U_c|||\;|||U_c - U_{d,\beta}|||_{\mathcal{W}_d^p/\mathcal{V}_c^p} \\ &\leq Ch^p\|u\|\;\|h^{-1}[U_c - U_{d,\beta}]\|_{\mathcal{E}}. \end{aligned}$$

Here we used (3.2) to replace $U_{d,\beta}$ by u in the first equality, then in the second equality we observed that $\beta b_d(U_c - u, U_c - U_{d,\beta}) = 0$ since $[U_c - u] = 0$ on $\mathcal{E}_I \cup \mathcal{E}_D$ and we inserted an arbitrary function v in $\mathcal{V}_c^p$ which is zero on Γ_D using (2.8). Finally we used the Cauchy

Schwarz inequality followed by Lemma 3.1 and Lemma 3.3. From this estimate we conclude that

$$\|h^{-1}[U_c - U_{d,\beta}]\|_{\mathcal{E}} \leq C(\beta - \beta_0)^{-1} h^p \|u\|_{p+1}. \tag{3.27}$$

Combining (3.27) with (3.26) proves the proposition. □

We summarize our main results in the following theorem.

Theorem 3.1 *Let $U_{d,\beta}$ be the dG approximation defined in (3.2). Then the following holds*

$$\lim_{\beta\to\infty} U_{d,\beta} = U_c, \tag{3.28}$$

where U_c is the cG approximation defined in (2.8), and the limit

$$\Sigma_{n_E}(U_c) = \lim_{\beta\to\infty} \Sigma_{n_E,\beta}(U_{d,\beta}), \tag{3.29}$$

defines a conservative flux for the cG method. Furthermore, the limit (3.29) can be calculated as follows

$$\lim_{\beta\to\infty} \Sigma_{n_E,\beta}(U_{d,\beta}) = \langle \sigma_{n_E}(U_c)\rangle - h^{-1}[V], \tag{3.30}$$

whith $V \in \mathcal{W}_d^p$ a solution of

$$b_d(V, v) = l_d(v) - a_d(U_c, v) \quad \textit{for all } v \in \mathcal{W}_d^p. \tag{3.31}$$

The equation (3.31) is solvable and the jump $[V]$ is uniquely determined.

Proof. The first statement (3.28) is an immediate consequence of Proposition 3.1. Provided the limit (3.28) exists it is clear that $\Sigma_n(U_c)$ must be an elementwise conservative flux since $\Sigma_{n,\beta}(U_{d,\beta})$ is elementwise conservative for all β.

We now proceed to show that the limit is well defined. Using the definition (3.2) of the dG method we get the equation

$$\beta(b_d(U_{d,\beta}, v) - m_d(v)) = l_d(v) - a_d(U_{d,\beta}, v).$$

For the right hand side we get the following limit

$$\lim_{\beta\to\infty} \Big(l_d(v) - a_d(U_{d,\beta}, v)\Big) = l_d(v) - a_d(U_c, v) \quad \text{for all } v \in \mathcal{W}_d^p,$$

using Proposition 3.1. Next, for the left hand side we have

$$\lim_{\beta\to\infty} \beta(b_d(U_{d,\beta}, v) - m_d(v)) = \lim_{\beta\to\infty} \beta(b_d(U_{d,\beta} - U_c, v)) = b_d(V, v),$$

for some $V \in \mathcal{W}_d^p$. The existence of V follows from the fact that the sequence $\{\beta(U_{d,\beta} - U_c)\}$ is bounded in the finite dimensional space $\mathcal{W}_d^p$, and therefore we may extract a convergent subsequence with limit V. Clearly V will satisfy (3.31) and from, the argument below it follows that $[V]$ is uniquely determined, and thus the limit exists and is well defined.

The linear system of equations (3.31) is solvable, since if $b_d(v, w) = 0$ for all v, then w is continuous and equals zero on Γ_D, and then the right hand side $l_d(w) - a_d(U_c, w) = 0$, and thus a solution exists. Next, assume that V and V' are two solutions of (3.31), then $b_d(V - V', w) = 0$, for all $w \in \mathcal{W}_d^p$ and thus $V - V'$ is continuous and zero on Γ_D, therefore $0 = [V - V'] = [V] - [V']$. Hence, the jump is uniquely determined.

□

We conclude the paper with an error estimate for the conservative flux.

Theorem 3.2 *For all $\beta \geq \beta_0$ the following error estimate holds*

$$\|\Sigma_{n,\beta}(U_{d,\beta}) - \sigma_n(u)\|_{\mathcal{E}} \leq Ch^p \|u\|_{p+1}, \tag{3.32}$$

where C is a constant not dependent on β or h. In particular, the error estimate holds for $\Sigma_n(U_c) = \lim_{\beta\to\infty} \Sigma_{n,\beta}(U_{d,\beta})$.

Proof. This result follows imediately from the error estimate in Lemma 3.3. □

Example. We illustrate our estimates on a simple model problem. Consider the Poisson equation (2.1), with A the two by two identity matrix, on the unit square $\Omega = [0, 1]^2$ with homogeneous Dirichlet conditions $g_D = 0$ on the boundary $\Gamma_D = \Gamma$ and right hand side f such that the solution is $u = \sin(\pi x_1)\sin(\pi x_2)$. We solve this problem on a quasiuniform triangulation $\mathcal{K}$ of Ω using the cG method with polynomials of degree $p = 1, 2, 3$, and calculate the conservative flux according to Theorem 3.1 using polynomials of degree q, with $0 \leq q \leq p - 1$.

In Figure 1 we plot the error in the average flux (circle), and the conservative flux with $q = 0$ (triangle), $q = 1$ (square), and $q = 2$ (diamond), as functions of the meshsize h. We observe that the convergence is in agreement with the prediction of Theorem 3.2.

Next, in Figure 2 we illustrate the dependency of $|||U_{d,\beta} - U_c|||$ on β, given in Proposition 3.1. We plot the energynorm for $p = 1$ (circle), $p = 2$ (triangle), and $p = 3$ (square), and increasing β on a fixed grid with $h \approx 0.20$. Clearly, the results verify the estimate in Proposition 3.1.

Finally, we plot the exact flux together with the averaged and conservative fluxes on the side from $(0, 0)$ to $(1, 0)$ on a mesh with triangles of size 0.25. The method uses quadratic approximation, $p = 2$, and we consider correction of the fluxes of order $q = 0$ and 1. We note that the conservative flux certainly is better than the averaged.

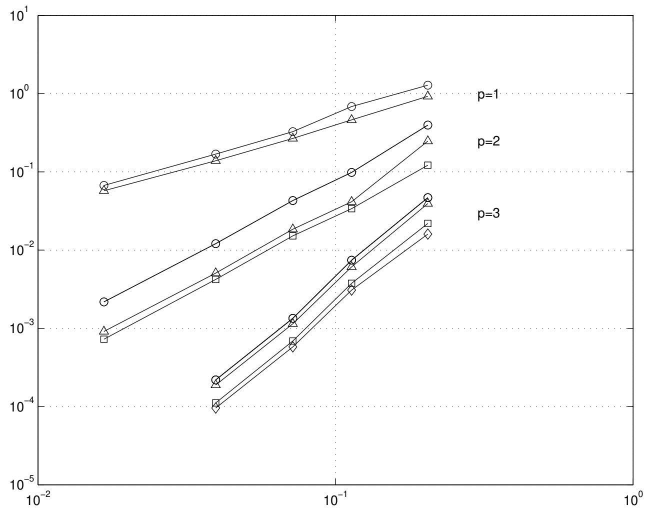


Figure 1: Error in the average (circle) and conservative flux with $q = 0$ (triangle), $q = 1$ (square), and $q = 2$ (diamond), for $p = 1, 2, 3$.

References

[1] M. Ainsworth and J.T. Oden. A procedure for a posteriori error estimation for h-p finite element methods. *Comput. Methods Appl. Mech. Engr.*, 101:73–96, 1992.

[2] M. Ainsworth and J.T. Oden. A unified approach to a posteriori error estimation based on residual methods. *Numer. Math.*, 65:23–50, 1993.

[3] I. Babuška and A. Miller. The post processing approach in the finite element method, part 1: calculation of displacements, stresses and other higher derivatives of the displacements. *Internat. J. Numer. Methods. Engrg*, 34:1085–1109, 1984.

[4] S.C. Brenner and L.R. Scott. *The Mathematical Theory of Finite Element Methods.* Springer-Verlag, Berlin, 1994.

[5] P. Hansbo and M.G. Larson. Discontinuous finite element methods for nearly incompressible elasticity by use of Nitshe's method. Technical report, Chalmers Finite Element Center, 1999. submitted.

[6] T.J.R. Hughes, G. Engel, M.G. Larson, and L. Mazzei. The continuous Galerkin method is conservative. Technical report, Stanford University, 1999. submitted.

[7] T.J.R. Hughes, G. Engel, L. Mazzei, and M.G. Larson. A comparision of discontinuous and continuous Galerkin methods based on error estimates, conservation, robustness, and efficiency. In *Proceedings of International Symposium on Discontinuous Galerkin Methods*, volume 11 of *Lectures in Computational Sciences and Engineering.* Springer, 1999.

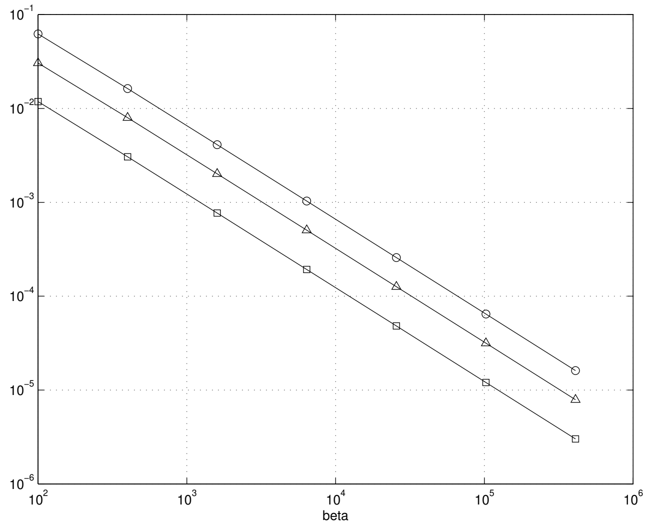


Figure 2: The energynorm $|||U_{d,\beta}-U_c|||$ as a function of β for $p = 1$ (circle), $p = 2$ (triangle), and $p = 3$ (square), on a fixed grid with $h \approx 0.20$.

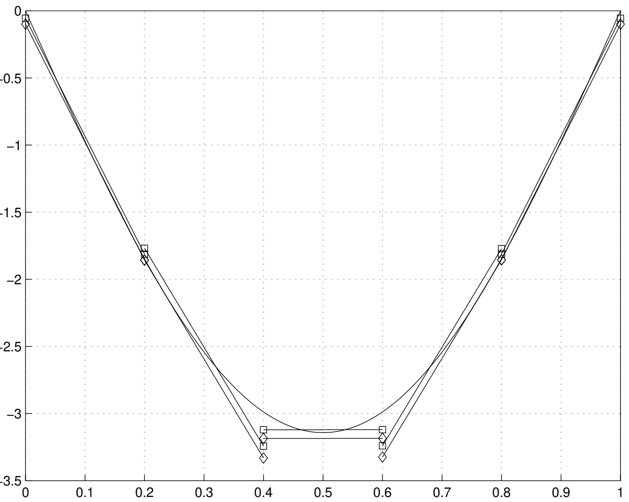
[8] D.W. Kelly. The self equilibration of residuals and complementary error estimates in the finite element method. *Internat. J. Numer. Methods Engrg.*, 20:1491–1506, 1984.

[9] P. Ladeveze and D. Leguillon. Error estimate procedure in the finite element method and applications. *SIAM J. Numer. Anal.*, 20(3):485–509, 1983.

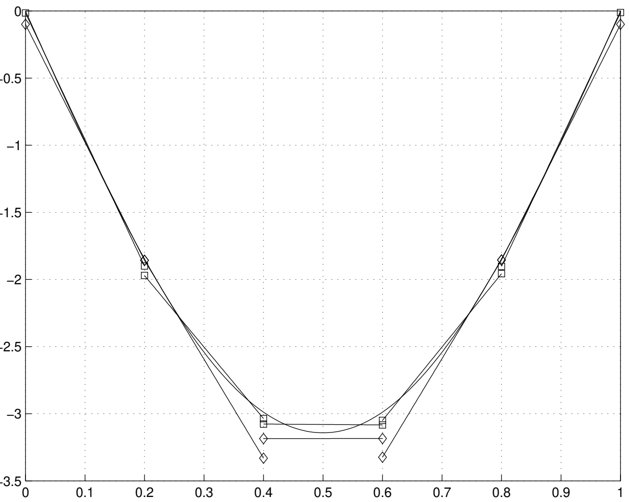
[10] J. Nitsche. Über ein Variationzprinzip zur Lösung von Dirichlet Problemen bei Verwendung von Teilräumen. *Abh. Math. Sem. Univ. Hamburg*, 36(9), 1971.

[11] J.T. Oden, I. Babuska, and C.E. Baumann. A discontinuous hp finite element method for diffusion problems. *J. Comput. Phys.*, 146:491–519, 1998.

[12] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems.* Springer-Verlag, Berlin, 1997.



(a) $q = 0$



(b) $q = 1$

Figure 3: The exact flux (solid), the averaged flux (solid diamond), and the conservative flux (solid square) for quadratic approximation, $p = 2$, and $q = 0, 1$, respectively, plotted on one side of the square

Chalmers Finite Element Center Preprints

2000–01 *Adaptive Finite Element Methods for the Unsteady Maxwell's Equations*
Johan Hoffman

2000–02 *A Multi-Adaptive ODE-Solver*
Anders Logg

2000–03 *Multi-Adaptive Error Control for ODEs*
Anders Logg

2000–04 *Dynamic Computational Subgrid Modeling* (Licentiate Thesis)
Johan Hoffman

2000–05 *Least-Squares Finite Element Methods for Electromagnetic Applications* (Licentiate Thesis)
Rickard Bergström

2000–06 *Discontinuous Galerkin methods for incompressible and nearly incompressible elasticity by Nitsche's method*
Peter Hansbo and Mats G. Larson

2000–07 *A Discountinuous Galerkin Method for the Plate Equation*
Peter Hansbo and Mats G. Larson

2000–08 *Conservation Properties for the Continuous and Discontinuous Galerkin Methods*
Mats G. Larson and A. Jonas Niklasson

These preprints can be obtained from

`www.phi.chalmers.se/preprints`