Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 205 (2005) 401–407
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Short Note

An explicit expression for the penalty parameter of the interior penalty method

Khosro Shahbazi *

Department of Mechanical and Industrial Engineering, University of Toronto, 5 King's College Road, Toronto, Ont., Canada M5S 3G8

Received 23 August 2004; received in revised form 19 November 2004; accepted 19 November 2004

Abstract

We derive an explicit expression for the penalty parameter of the interior penalty method for elliptic problems. The expression yields a coercive bilinear form, and is valid for general meshes comprising of (geometrically nonconforming) simplical elements.

Keywords: Interior penalty method; Penalty parameter; Discontinuous Galerkin methods; High order method

1. Introduction

The interior penalty (IP) method devised in the late 1970s [1] is a type of discontinuous Galerkin method for the spatial discretization of elliptic partial differential equations. The IP method, like other discontinuous Galerkin methods, has advantages over the classical continuous Galerkin method in facilitating hp-adaptivity and yielding block diagonal mass matrices important in time-dependent problems. Moreover, the IP method gives a symmetric, locally conservative, and small-stencil discretization. This last property, in which the degrees of freedom of each element couple only with those of its immediate neighbors, is critical in reducing memory requirements and achieving efficient parallelization in large scale computations.

Despite its early introduction and its advantages, the IP method has not been popular. One drawback to this scheme is that it requires the user to specify a mesh-dependent parameter, known as a penalty parameter. If the value of this parameter is not sufficiently large, the approximate solution is unstable. [1] On the other hand, an arbitrarily large value of the penalty parameter degrades the performance of the iterative

* Tel.: +1 4166040868; fax: +1 4169787753.
E-mail address: shahbazi@mie.utoronto.ca.

[1] Here instability is a direct consequence of non-coercivity of the associated bilinear form. Its characteristic is that small variations in the penalty parameter yield large variations in the field variable, as will be seen below.

0021-9991/$ - see front matter © 2005 Published by Elsevier Inc.
doi:10.1016/j.jcp.2004.11.017

solver of the linear system arising from the IP discretization, as shown in Section 2. In real applications, where highly anisotropic and heterogeneous mesh geometries are used and in adaptive algorithms, where varying approximation orders are also used, it is difficult to know a priori the minimum acceptable value of the penalty parameter. Therefore, the objective of this paper is to derive an explicit expression for the value of the penalty parameter guaranteed to give a stable solution. We consider a domain partitioned into triangular or tetrahedral elements in two or three space dimensions.

Before deriving this expression, we briefly describe the IP method and show the effect of the penalty parameter on the overall efficiency of the scheme.

2. Interior penalty method

We seek the IP formulation of the Poisson equation with Dirichlet boundary conditions:

$$-\Delta u = f \quad \text{in } \Omega, \tag{1a}$$

$$u = g \quad \text{on } \partial\Omega, \tag{1b}$$

where Ω is a polygonal domain of dimension $d = 1, 2$, or 3.

We first introduce some notation. Let K^+ and K^- be two adjacent elements in $\mathscr{T}_h$, a triangulation of Ω; let $\mathbf{x}$ be an arbitrary point of the interior set $e = \partial K^- \cap \partial K^+$, which is assumed to have non-zero dimension $(d-1)$ and is referred to as a face; and let $\mathbf{n}^-$ and $\mathbf{n}^+$ be the corresponding normal vectors at that point. Let u be a smooth function inside each element $K^\pm$ and let us denote by $u^\pm$ the trace of u on e from the interior of $K^\pm$. Then, we define the mean $\{\cdot\}$ and the jump $[\![\cdot]\!]$ at $\mathbf{x} \in e$ as

$$\{u\} := (u^+ + u^-)/2, \quad [\![u]\!] := u^+\mathbf{n}^+ + u^-\mathbf{n}^-.$$

For a point $\mathbf{x}$ on the boundary set $\partial K \cap \partial\Omega$ with normal vector $\mathbf{n}$, we define the trace operators as

$$\{u\} := u, \quad [\![u]\!] := u\mathbf{n}.$$

The mean of a vector-valued function is defined in a similar way. We also denote by Γ_I the union of all interior sets e, and we set $\Gamma = \Gamma_I \cup \partial\Omega$.

We take the discontinuous approximation to the exact solution u, u_h, in the finite element space V_h, where

$$V_h := \{v \in L^2(\Omega) \mid v|_K \in Q_{k_K}(K)\ \forall K \in \mathscr{T}_h\}.$$

Here, $Q_{k_K}(K)$ is the set of polynomials of degree at most k_K on K, $k \geqslant 0$. The approximate solution is then defined by requiring that

$$a(u_h, v_h) = F(v_h) \quad \forall u_h,\ v_h \in V_h,$$

where

$$a(u,v) = \sum_K \int_K \nabla u \cdot \nabla v \,\mathrm{d}\mathbf{x} - \int_\Gamma ([\![v]\!] \cdot \{\nabla u\} + [\![u]\!] \cdot \{\nabla v\}) \,\mathrm{d}s + \int_\Gamma \mu [\![u]\!] \cdot [\![v]\!] \,\mathrm{d}s, \tag{2a}$$

$$F(v) = \int_\Omega fv \,\mathrm{d}\mathbf{x} - \int_{\partial\Omega} g \nabla v \cdot \mathbf{n} \,\mathrm{d}s. \tag{2b}$$

The last term on the r.h.s of (2a) is defined for interior and boundary faces, Γ. This is the penalty term, which is added to enforce the coercivity of the bilinear form $a(u,v)$. We must specify a value for the penalty parameter μ that ensures the coercivity of the bilinear form and, thus, the stability of the approximate solution. In previous work, μ has only been defined to within a multiplicative constant. For example, Arnold [1]

defined $\mu = \gamma_0/l_e$, where $l_e = \mathrm{diam}(e)$, and γ_0 is a large unknown positive constant. In the context of the mixed hp-discontinuous Galerkin finite element method, Schötzau et al. [10] defined $\mu = \eta h^{-1}k^2$, where $h = \min(h_{K^+}, h_{K^-})$, $k = \max(k_{K^+}, k_{K^-})$, and $h_K = \mathrm{diam}(K)$, again leaving η as a large unknown positive constant.

Unfortunately, the above expressions for μ are less than optimal in practice, because a large value of μ has a detrimental effect on the conditioning of the matrix that represents the bilinear form $a(u,v)$. As proved by Castillo [3] for all approximating polynomial degrees, the spectral condition number of this matrix in L_2 norm grows linearly with μ (see Theorem 3.4 in [3]). It is therefore expected that the magnitude of μ will affect the overall efficiency of the iterative solver of the system arising from the IP discretization.

To further investigate this, we conducted the following experiment. Using the nodal high order IP method, we discretized (1a) and (1b) with $g = 0$ and $f = 2\pi^2\sin(\pi x)\sin(\pi y)$, on the square domain $[-1,1]\times[-1,1]$. The corresponding exact solution is $u = \sin(\pi x)\sin(\pi y)$. Our nodal basis was the Lagrange polynomials calculated based on the nodal set of Hesthaven [6] defined on the standard triangle. The domain was partitioned once into 72 structured triangles and once into 72 unstructured (heterogeneous) triangles as shown in Figs. 1(a) and (b), respectively. To solve the resulting linear system, we used the preconditioned conjugate gradient method. The preconditioner was a two-level element-based Schwarz preconditioner. Its local part corresponded to the IP discretization on each element, similar to that of Feng and Karakashian [4], and its global coarse part corresponded to the IP discretization on the same mesh but with the lower approximation order $k = 1$. Our implementation was based on the algorithm oriented mesh database (AOMD) [9] and portable, extensible toolkit for scientific computing (PETSc) [7,8]. We carried out simulations using different values of μ and with the range of approximation orders ($k = 2,\ldots,6$). The initial guess for the conjugate gradient iterations was a vector with random entries confined to the interval $[0,1]$ and the stopping criterion was a relative residual smaller than 10^{-11}. The results are shown in Figs. 1(c) and (d) for the structured and unstructured meshes, respectively. Although the iteration counts are higher for the unstructured mesh, it is clear that the iteration counts in both cases grow almost logarithmically with μ, implying that arbitrarily large values of μ yield unacceptably large iteration counts. It is therefore clear that an explicit expression for the penalty parameter would be useful, so as to guarantee coercivity while minimizing computational expense.

3. Explicit expression for the penalty parameter

Here, we derive an explicit expression for the penalty parameter μ for a d-dimensional simplex. Our derivation is based on the results of Warburton and Hesthaven [11] on trace inverse inequalities. Using orthogonal polynomials, they proved the following inequality for a simplicial element K and $\forall v \in Q_k(K)$:

$$\int_e v^2\,\mathrm{d}s \leqslant \frac{(k+1)(k+d)}{d}\,\frac{\mathscr{A}(e)}{\mathscr{V}(K)}\int_K v^2\,\mathrm{d}\mathbf{x}, \tag{3}$$

where for $d = 3$, $\mathscr{A}$ and $\mathscr{V}$ denote area and volume, respectively, and for $d = 2$, they denote length and area, respectively. For $d = 1$, $\mathscr{A}(\cdot) = 1$ and $\mathscr{V}$ denotes length (see Theorems 2–4 in [11]).

We must find μ such that the bilinear form $a(u,v)$ is coercive, i.e., so that there exists a positive constant c_s such that

$$a(v,v) \geqslant c_s\|v\|_h^2 \quad \forall v \in V_h, \tag{4}$$

where

$$a(v,v) = \sum_K \int_K (\nabla v)^2\,\mathrm{d}\mathbf{x} - 2\int_\Gamma [\![v]\!]\cdot\{\nabla v\}\,\mathrm{d}s + \int_\Gamma \mu[\![v]\!]^2\,\mathrm{d}s \tag{5}$$

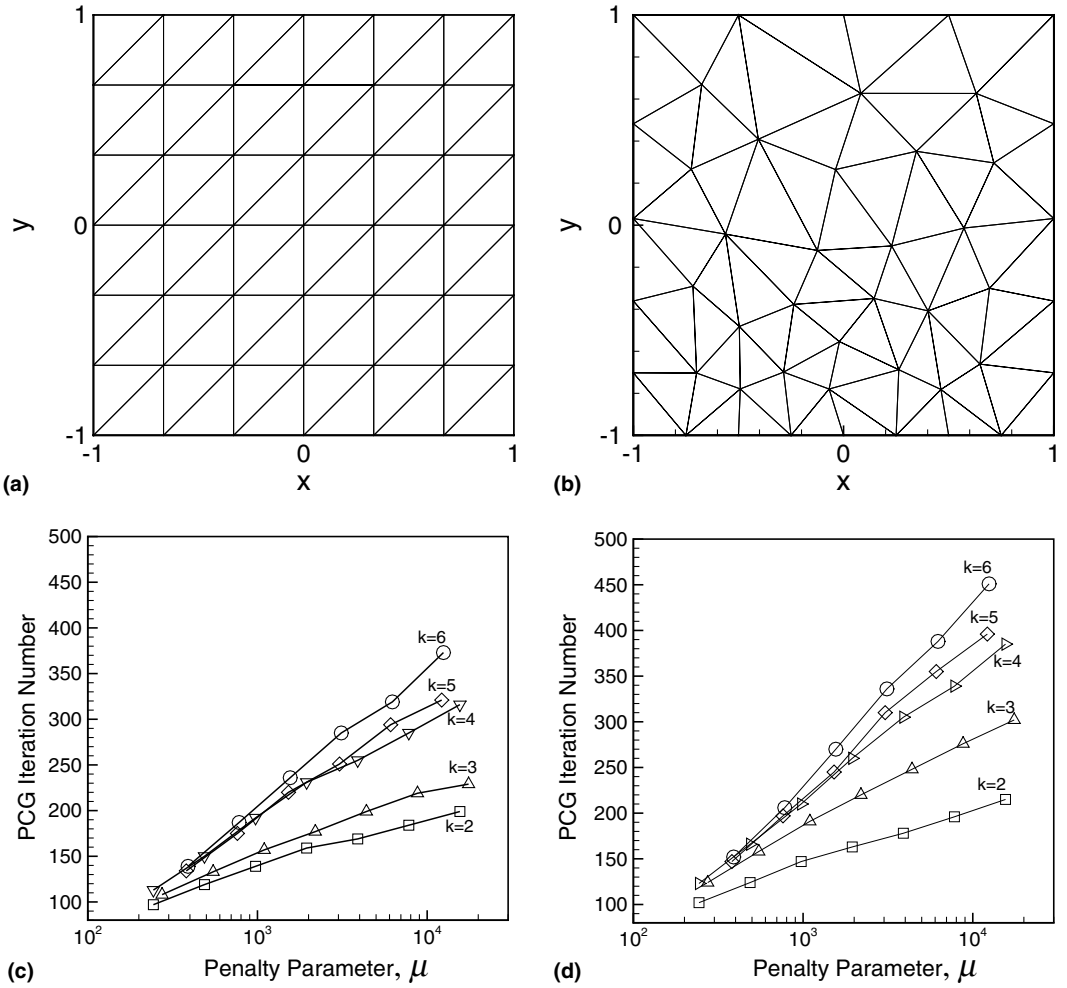

Fig. 1. (a) The computational domain used for tests partitioned into 72 structured triangular elements. (b) The same domain partitioned into 72 unstructured (heterogeneous) triangular elements generated using Gmsh software [5]. (c) and (d) The number of preconditioned conjugate gradient (PCG) iterations needed to solve the Poisson problem vs. the penalty parameter μ using meshes in (a) and (b), respectively. The nodal high order IP method was used, with the range of approximation orders ($k = 2,\ldots,6$).

and

$$\|v\|_h^2 = \sum_K |v|_{1,K}^2 + \int_\Gamma [\![v]\!]^2 \,\mathrm{d}s$$

with the seminorm $|\cdot|_{1,K}$ defined over $H^1(K)$ by

$$|v|_{1,K}^2 = \int_K (\nabla v)^2 \,\mathrm{d}\mathbf{x}.$$

We first find a bound on the negative term on the r.h.s. of (5). Using the arithmetic–geometric mean inequality $ab \leqslant (\epsilon_e/2)a^2 + (1/2\epsilon_e)b^2$ with $\epsilon_e > 0$ yields

$$\int_e [\![v]\!] \cdot \{\nabla v\}\,\mathrm{d}s \leqslant \frac{\epsilon_e}{2}\int_e [\![v]\!]^2\,\mathrm{d}s + \frac{1}{2\epsilon_e}\int_e \{\nabla v\}^2\,\mathrm{d}s \quad \forall e \in \Gamma.$$

Adding the above inequality over all e, noting that on Γ_I, $\{\nabla v\}^2 = (\nabla v^+)^2/4 + (\nabla v^-)^2/4 + (\nabla v^+ \cdot \nabla v^-)/2$ and on $\partial\Omega$, $\{\nabla v\}^2 = (\nabla v)^2$, and using the inequality $a^2 + b^2 + 2ab \leqslant 2a^2 + 2b^2$ yields

$$\int_\Gamma [\![v]\!] \cdot \{\nabla v\}\,\mathrm{d}s \leqslant \sum_{e\in\Gamma}\frac{\epsilon_e}{2}\int_e [\![v]\!]^2\,\mathrm{d}s + \sum_{e\in\Gamma_I}\frac{1}{2\epsilon_e}\int_e \left[\frac{1}{2}(\nabla v^+)^2 + \frac{1}{2}(\nabla v^-)^2\right]\mathrm{d}s + \sum_{e\in\partial\Omega}\frac{1}{2\epsilon_e}\int_e (\nabla v)^2\,\mathrm{d}s.$$

Substituting the above inequality in (5), and then using (3) yields

$$a(v,v) \geqslant \sum_K \sum_{i_e=1}^{n}\left(\frac{c_{e,K}}{c_K} - \frac{c_{e,K}}{\epsilon_e}\right)\int_K (\nabla v)^2\,\mathrm{d}\mathbf{x} + \sum_{e\in\Gamma}\int_e (\mu - \epsilon_e)[\![v]\!]^2\,\mathrm{d}s, \tag{6}$$

where

$$c_{e,K} = \begin{cases} \frac{(k_K+1)(k_K+d)}{d}\,\frac{\mathscr{A}(e)}{\mathscr{V}(K)} & e \in \partial\Omega, \\ \frac{(k_K+1)(k_K+d)}{d}\,\frac{\mathscr{A}(e)/2}{\mathscr{V}(K)} & e \in \Gamma_I, \end{cases}$$

$$c_K = \frac{(k_K+1)(k_K+d)}{d}\,\frac{[\mathscr{A}(\partial K \setminus \partial\Omega)/2 + \mathscr{A}(\partial K \cap \partial\Omega)]}{\mathscr{V}(K)} \quad K \in \mathscr{T}_h. \tag{7}$$

In (6), i_e denotes the local index of face e restricted to the element K, and n denotes the total number of faces of each element K. To yield a positive r.h.s in (6), it is sufficient to choose $\epsilon_e \geqslant c_K$ for $e \in \partial\Omega$, $\epsilon_e \geqslant \max(c_{K^+}, c_{K^-})$ for $e \in \Gamma_I$, and $\mu \geqslant \epsilon_e$. Thus, we choose the local penalty parameter μ_e as

$$\mu_e = c_K \quad \forall e \in \partial\Omega, \tag{8a}$$

$$\mu_e = \max(c_{K^+}, c_{K^-}) \quad \forall e \in \Gamma_I. \tag{8b}$$

Now, if we set

$$c_1 = \min_{K,e}\left(\frac{c_{e,K}}{c_K} - \frac{c_{e,K}}{\epsilon_e}\right) \quad \forall e \in \Gamma,\ K \in \mathscr{T}_h,$$
$$c_2 = \min_e(\mu - \epsilon_e) \quad \forall e \in \Gamma,$$

we obtain

$$a(v,v) \geqslant c_1\sum_K \int_K (\nabla v)^2\,\mathrm{d}\mathbf{x} + c_2\int_\Gamma [\![v]\!]^2\,\mathrm{d}s,$$

Finally, if we choose $c_s = \min(c_1, c_2)$, the coercivity of the bilinear form, (4), results.

Remark 1. The penalty expression in (8a) and (8b) is defined for each face $e \in \Gamma$ and depends on the geometries and approximating polynomial orders within the elements sharing e. Obviously, a global bound for the penalty parameter can be derived as

$$\mu = \max_e(\mu_e). \tag{9}$$

Remark 2. Our estimation in (9) is sharp, which we demonstrate by the following numerical experiment. By once again solving Poisson Eq. (1a) and (1b) using the methodology of Section 2 and the structured triangular mesh in Fig. 1(a), we computed the maximum nodal error versus μ for polynomial approximation orders $k = 1,\ldots,8$ (Fig. 2). It can be seen that the solution is unstable for

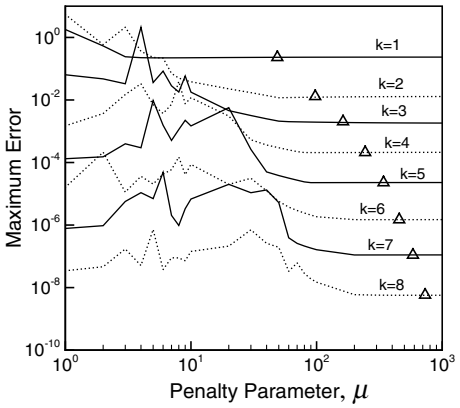

Fig. 2. Maximum nodal error vs. penalty parameter μ for different orders of approximation $k = 1, \ldots, 8$; the triangles represent the values of the penalty parameter calculated using Eq. (9).

$\mu < \mu^*$ (i.e., a small variation in the penalty parameter yields a large variation in the field variable u) and stable for $\mu > \mu^*$ (i.e., variations in μ yields almost no variations in u). The critical values of penalty parameter μ^* are approximately $10, 50, 50, 90, 90, 150, 200$ and 200 for $k = 1, \ldots, 8$, respectively. On the same figure, we also show the value of the penalty parameter computed using Eq. (9). We observe that the estimation based on (9) yields a stable solution and it is roughly three times larger than μ^* at each k. Based on the results of Fig. 1(c), this means that selecting μ according to Eq. (9) guarantees a stable solution at a computational cost within a factor of roughly 1.2 of that of the (unknown) optimal penalty parameter, μ^*, for the higher approximation orders ($k = 6$, 7 and 8), and at least for the mesh used in this study.

Remark 3. For the case of general meshes, when the elements are not face-to-face including those with hanging nodes, an explicit expression for the penalty parameter for a face e_f shared by a collection of adjacent elements $\{K_i | i = 1, \ldots, N\}$ can be defined as

$$\mu_{e_f} = \max(c_{K_i}), \quad i = 1, \ldots, N, \tag{10}$$

where c_{K_i} is computed using (7). Following the above procedure, it is proved that choosing the penalty parameter based on (10) guarantees the coercive bilinear form for general meshes.

Remark 4. Through a similar procedure, an explicit expression for the penalty parameter of the discontinuous Galerkin method of Baker [2] can be derived. The same formulas (8a) and (8b) and (10) are valid in this method, but with a slightly different c_K

$$c_K = \frac{(k_K + 1)(k_K + d)}{d} \frac{\mathscr{A}(\partial K)}{\mathscr{V}(K)} \quad K \in \mathscr{T}_h.$$

Acknowledgements

The author acknowledges fruitful conversations with Prof. C.R. Ethier, University of Toronto and Dr. P.F. Fischer, Argonne National Laboratory. The author also thanks one of the anonymous reviewers for his valuable suggestions. This work was financially supported by National Sciences and Engineering Council of Canada through operating Grant A2191 (C.R. Ethier).

References

[1] D. Arnold, An interior penalty finite element method with discontinuous elements, SIAM J. Numer. Anal. 19 (1982) 742–760.
[2] G.A. Baker, Finite element methods for elliptic equations using nonconforming elements, Math. Comp. 31 (1977) 44–59.
[3] P. Castillo, Performance of discontinuous Galerkin methods for elliptic PDES, SIAM J. Sci. Comput. 24 (2002) 524–547.
[4] X. Feng, O.A. Karakashian, Two-level additive Schwarz methods for a discontinuous Galerkin approximation of second order elliptic problems, SIAM J. Numer. Anal. 30 (2001) 1343–1365.
[5] C. Geuzaine, J.F. Remacle, Gmsh Reference Manual, Edition 1.12, 2003. Available from: <http://www.geuz.org.org/gmsh/>.
[6] J.S. Hesthaven, From electrostatics to almost optimal nodal sets for polynomial interpolation in a simplex, SIAM J. Numer. Anal. 35 (1998) 655–676.
[7] S. Balay, K. Buschelman, W.D. Gropp, D. Kaushik, M. Knepley, L.C. McInnes, B.F. Smith, H. Zhang, PETSc home page, 2001. Available from: <http://www.mcs.anl.gov/petsc>.
[8] S. Balay, K. Buschelman, V. Eijkhout, W.D. Gropp, D. Kaushik, M.G. Knepley, L.C. McInnes, B.F. Smith, H. Zhang, PETSc Users Manual, ANL-95/11 – Revision 2.1.5, Argonne National Laboratory, 2004.
[9] J. Remacle, AOMD home page, 2002. Available from: <http://www.scorec.rpi.edu/AOMD/>.
[10] D. Schötzau, C. Schwab, A. Toselli, Mixed hp-DGFEM incompressible flows, SIAM J. Numer. Anal. 40 (2003) 2171–2194.
[11] T. Warburton, J.S. Hesthaven, On the constants in hp-finite element trace inverse inequality, Comput. Meth. Appl. Mech. Eng. 192 (2003) 2765–2773.