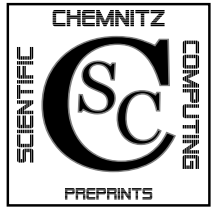
TECHNISCHE UNIVERSITÄT CHEMNITZ

T. Eibner J. M. Melenk

p-FEM quadrature error analysis on tetrahedra

CSC/06-01



Chemnitz Scientific Computing Preprints

Chemnitz Scientific Computing Preprints – ISSN 1864-0087

Impressum:

Chemnitz Scientific Computing Preprints — ISSN 1864-0087
(1995–2005: Preprintreihe des Chemnitzer SFB393)

Herausgeber:
Professuren für
Numerische und Angewandte Mathematik
an der Fakultät für Mathematik
der Technischen Universität Chemnitz

Postanschrift:
TU Chemnitz, Fakultät für Mathematik
09107 Chemnitz
Sitz:
Reichenhainer Str. 41, 09126 Chemnitz

`http://www.tu-chemnitz.de/mathematik/csc/`

Other titles in the SFB393 series:

05-01 C. Pester. A residual a posteriori error estimator for the eigenvalue problem for the Laplace-Beltrami operator. January 2005.

05-02 J. Badía, P. Benner, R. Mayo, E. Quintana-Ortí, G. Quintana-Ortí, J. Saak. Parallel Order Reduction via Balanced Truncation for Optimal Cooling of Steel Profiles. February 2005.

05-03 C. Pester. CoCoS – Computation of Corner Singularities. April 2005.

05-04 A. Meyer, P. Nestler. Mindlin-Reissner-Platte: Einige Elemente, Fehlerschätzer und Ergebnisse. April 2005.

05-05 P. Benner, J. Saak. Linear-Quadratic Regulator Design for Optimal Cooling of Steel Profiles. April 2005.

05-06 A. Meyer. A New Efficient Preconditioner for Crack Growth Problems. April 2005.

05-07 A. Meyer, P. Steinhorst. Überlegungen zur Parameterwahl im Bramble-Pasciak-CG für gemischte FEM. April 2005.

05-08 T. Eibner, J. M. Melenk. Fast algorithms for setting up the stiffness matrix in hp-FEM: a comparison. June 2005.

05-09 A. Meyer, P. Nestler. Mindlin-Reissner-Platte: Vergleich der Fehlerindikatoren in Bezug auf die Netzsteuerung Teil I. June 2005.

05-10 A. Meyer, P. Nestler. Mindlin-Reissner-Platte: Vergleich der Fehlerindikatoren in Bezug auf die Netzsteuerung Teil II. July 2005.

05-11 A. Meyer, R. Unger. Subspace-cg-techniques for clinch-problems. September 2005.

05-12 P. Ciarlet, Jr, B. Jung, S. Kaddouri, S. Labrunie, J. Zou. The Fourier Singular Complement Method for the Poisson Problem. Part III: Implementation Issues. October 2005.

05-13 T. Eibner, J. M. Melenk. Multilevel preconditioning for the boundary concentrated *hp*-FEM. December 2005.

05-14 M. Jung, A. M. Matsokin, S. V. Nepomnyaschikh, Yu. A. Tkachov. Multilevel preconditioning operators on locally modified grids. December 2005.

05-15 S. Barrachina, P. Benner, E. S. Quintana-Ortí. Solving Large-Scale Generalized Algebraic Bernoulli Equations via the Matrix Sign Function. December 2005.

05-16 B. Heinrich, B. Jung. Nitsche- and Fourier-finite-element method for the Poisson equation in axisymmetric domains with re-entrant edges. December 2005.

05-17 M. Randrianarivony, G. Brunnett. C^0-paving of closed meshes with quadrilateral patches. December 2005.

05-18 M. Randrianarivony, G. Brunnett. Quadrilateral removal and 2-ear theorems. December 2005.

05-19 P. Benner, E. S. Quintana-Ortí, G. Quintana-Ortí. Solving linear-quadratic optimal control problems on parallel computers. December 2005.

05-20 P. Benner, M. Pester. Parallel Library in Control: Model Reduction (PLiCMR) - Usage on CLiC. December 2005.

The complete list of current and former preprints is available via
`http://www.tu-chemnitz.de/sfb393/preprints.html`.

TECHNISCHE UNIVERSITÄT CHEMNITZ

Chemnitz Scientific Computing Preprints

T. Eibner **J. M. Melenk**

p-FEM quadrature error analysis on tetrahedra

CSC/06-01

Abstract

In this paper we consider the p-FEM for elliptic boundary value problems on tetrahedral meshes where the entries of the stiffness matrix are evaluated by numerical quadrature. Such a quadrature can be done by mapping the tetrahedron to a hexahedron via the Duffy transformation.

We show that for tensor product Gauss-Lobatto-Jacobi quadrature formulas with $q+1 = p+1$ points in each direction and shape functions that are adapted to the quadrature formula, one again has discrete stability for the fully discrete p-FEM.

The present error analysis complements the work Eibner and Melenk [2006] for the p-FEM on triangles/tetrahedra where it is shown that by adapting the shape functions to the quadrature formula, the stiffness matrix can be set up in optimal complexity.

CSC/06-01 **ISSN 1864-0087** **October 2006**

Contents

1 Introduction **1**

2 Quadrature Error Analysis **2**

2.0.1 Notation 2

2.1 Gauss-Lobatto-Jacobi quadrature 3

2.1.1 Gauss-Lobatto-Jacobi quadrature in 1D 3

2.1.2 Gauss-Lobatto-Jacobi quadrature on $\widehat{K}$ 3

2.2 Discrete Stability 4

2.3 Convergence Analysis of Fully Discrete p-FEM 4

3 Proof of Theorem 2.5 **5**

4 Numerical example **7**

T. Eibner
TU Chemnitz
Fakultät für Mathematik
D-09107 Chemnitz
`http://www.tu-chemnitz.de/~eti`

J.M. Melenk
TU Wien
Institut für Analysis und Scientific Computing, E101
A-1040 Wien
Wiedner Hauptstrasse 8-10
`http://www.math.tuwien.ac.at/~melenk`

References

C. Bernardi and Y. Maday. *Approximations spectrales de problèmes aux limites elliptiques*. Mathématiques & Applications. Springer Verlag, 1992.

C. Canuto and A. Quarteroni. Approximation results for orthogonal polynomials in Sobolev spaces. *Math. Comput.*, 38(257):67–86, 1982.

T. Eibner and J.M. Melenk. Fast algorithms for setting up the stiffness matrix in *hp*-fem: a comparison. In Elias A. Lipitakis, editor, *Computer Mathematics and its Applications-Advances and Developments (1994-2005)*, pages 575–596. LEA Publisher, 2006.

G.E. Karniadakis and S.J. Sherwin. *Spectral/hp Element Methods for CFD*. Oxford University Press, 1999.

Y. Maday and E.M. Rønquist. Optimal error analysis of spectral methods with emphasis on non-constant coefficients and deformed geometries. *Comput. Meth. Appl. Mech. Engrg.*, 80:91–115, 1990.

J.M. Melenk and C. Schwab. *hp* FEM for reaction-diffusion equations I: Robust exponentional convergence. *SIAM J. Numer. Anal.*, 35:1520–1557, 1998.

C. Schwab. *p- and hp-Finite Element Methods*. Oxford University Press, 1998.

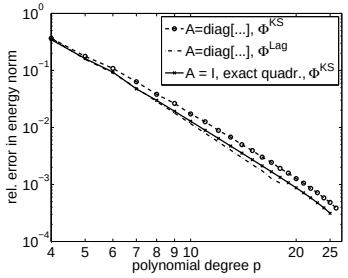


Figure 1: rel. energy norm error

where $r^2 = x_1^2 + x_2^2 + x_3^2$. We base the p-FEM on a single element on two different sets of shape functions: Φ^{KS} is the set of shape functions proposed by Karniadakis & Sherwin Karniadakis and Sherwin [1999] and spans $\mathcal{P}_p(\widehat{K}) \cap H_0^1(\widehat{K})$; the set Φ^{Lag} is, roughly, speaking, the set of Langrange interpolation points in the quadrature points (on $\mathcal{Q}$); it spans a space that contains $\mathcal{P}_p(\widehat{K}) \cap H_0^1(\widehat{K})$ and we refer to Eibner and Melenk [2006] for details. In both cases the stiffness matrix is set up using the minimal quadrature, i.e., $q = p$. Fig. 1 shows the relative energy norm error $(\frac{E_{exact} - a^q(u_N, u_N)}{E_{exact}})^{1/2}$ for both cases, where $E_{exact} = \int_\Omega \nabla u \cdot A \nabla \mathrm{d}\Omega$. To illustrate that the optimal rate of convergence is not affected by the quadrature, we include in Fig. 1 a calculation (based on Φ^{KS}) that corresponds to (15) with $A = I$; in this case the linear system of equations can be set up without quadrature errors. We observe indeed that the rate of convergence is the same as in the case of quadrature.

We close by pointing out that the shape functions in Φ^{Lag} are adapted to the quadrature rule. While the number of functions in Φ^{Lag} is (asymptotically for large p) 6 times that of Φ^{KS}, setting up the stiffness matrix is not slower than setting up the stiffness matrix based on Φ^{KS}. We refer to Eibner and Melenk [2006] for a detailed study.

1 Introduction

In the p-FEM and the closely related spectral method, the solution of an elliptic boundary value problems is approximated by piecewise (mapped) polynomials of degree p on a fixed mesh $\mathcal{T}$. In practice, the entries of the p-FEM stiffness matrix cannot be evaluated exactly due to variable coefficients and/or non-affine element maps and one has to resort to numerical quadrature to obtain a fully discrete method. Computationally, choosing shape functions that are related to the quadrature formula employed can significantly improve the computational complexity. For example, for tensor product elements (i.e., quadrilaterals, hexahedra) choosing tensor product Gauss-Lobatto quadrature with $q + 1 = p + 1$ points in each spatial direction and taking as shape functions the Langrange interpolation polynomials (of degree p) in the Gauss-Lobatto points effectively leads to a spectral method. The quadrature error analysis for the p-FEM/spectral method is available even for this case of minimal quadrature (see, e.g., Maday and Ronquist [1990], Melenk and Schwab [1998] and reference there). Key to the error analysis is a one-dimensional discrete stability result for the Gauss-Lobatto quadrature due to Canuto and Quarteroni [1982] (corresponding to $\alpha = 0$ in Lemma 2.3 below) that can readily be extended to quadrilaterals/hexahedra by tensor product arguments.

In the present paper, we show an analog of the error analysis of the above minimal quadrature for the p-FEM on tetrahedral meshes (the easier case of triangles can be treated completely analogously). Quadrature on a tetrahedron can be done by a mapping to a hexahedron via the Duffy transformation D of (3). We show in Thm. 2.5 that for tensor product Gauss-Lobatto-Jacobi quadrature formulas with $q+1 = p+1$ points in each direction, one again has discrete stability for the fully discrete p-FEM. A complete quadrature error analysis (Thm. 2.7, Cor. 2.9) then follows from Strang's lemma and shows that the convergence rates of the Galerkin p-FEM (where all integrals are evaluated exactly) is retained by the fully discrete p-FEM. The present error analysis complements the work Eibner and Melenk [2006] for the p-FEM on triangles/tetrahedra where it is shown that by adapting the shape functions to the quadrature formula, the stiffness matrix can be set up in optimal complexity. However, we mention that the approximation spaces employed in Eibner and Melenk [2006] are no longer the classical spaces $S^{p,1}(\mathcal{T})$ of piecewise polynomials but the spaces $S^{p,1}(\mathcal{T})$ augmented by bubble shape functions for each element, which makes the static condensation more expensive.

To fix ideas, we consider

$$-\nabla \cdot (A(x)\nabla u) = f \qquad \text{on } \Omega \subset \mathbb{R}^3, \qquad u|_{\partial\Omega} = 0, \tag{1}$$

where $A \in C(\overline{\Omega}, \mathbb{R}^{3\times 3})$ is pointwise symmetric positive definite. We require A

and f to be analytic on $\overline{\Omega}$ and the standard ellipticity condition

$$0 < \lambda_{min} \leq A(x) \leq \lambda_{max}, \qquad \forall x \in \Omega.$$

2 Quadrature Error Analysis

2.0.1 Notation

The reference tetrahedron $\widehat{K}$ and the reference cube $\mathcal{Q}$ are defined as

$$\widehat{K} = \{(x,y,z) \mid -1 < x,y,z \ \wedge \ x+y+z < -1\}, \qquad \mathcal{Q} := (-1,1)^3. \quad (2)$$

The Duffy transformation $D : \mathcal{Q} \to \widehat{K}$ is given by

$$D(\eta_1,\eta_2,\eta_3) := \left(\frac{(1+\eta_1)(1-\eta_2)(1-\eta_3)}{4} - 1, \frac{(1+\eta_2)(1-\eta_3)}{2} - 1, \eta_3 \right). \quad (3)$$

Lemma 2.1. *The Duffy transformation is a bijection between the (open) cube $\mathcal{Q}$ and the (open) tetrahedron $\widehat{K}$. Additionally,*

$$D'(\eta_1,\eta_2,\eta_3) := \left[\frac{\partial \xi_i}{\partial \eta_j}\right]_{i,j=1}^{3} = \begin{bmatrix} \frac{1}{4}(1-\eta_2)(1-\eta_3) & 0 & 0 \\ -\frac{1}{4}(1+\eta_1)(1-\eta_3) & \frac{1}{2}(1-\eta_3) & 0 \\ -\frac{1}{4}(1+\eta_1)(1-\eta_2) & -\frac{1}{2}(1+\eta_2) & 1 \end{bmatrix}^{\top},$$

$$(D'(\eta_1,\eta_2,\eta_3))^{-1} = \frac{1}{(1-\eta_2)(1-\eta_3)} \begin{bmatrix} 4 & 2(1+\eta_1) & 2(1+\eta_1) \\ 0 & 2(1-\eta_2) & 1-\eta_2^2 \\ 0 & 0 & (1-\eta_2)(1-\eta_3) \end{bmatrix},$$

$$\det D' = \left(\frac{1-\eta_2}{2}\right)\left(\frac{1-\eta_3}{2}\right)^2. \quad (4)$$

Proof. See, for example, Karniadakis and Sherwin [1999]. □

We employ standard notation by writing $\mathcal{P}_p(\widehat{K})$ for the space of polynomials of degree p on $\widehat{K}$, and by denoting $\mathcal{Q}_p(\mathcal{Q})$ the tensor-product space of polynomials of degree p in each variable, Schwab [1998]; additionally we set

$$\widetilde{\mathcal{Q}}_p := \{u \in \mathcal{Q}_p(\mathcal{Q}) \,|\, \partial_1 u = \partial_2 u = \partial_3 u = 0 \text{ on } \eta_3 = 1 \text{ and } \partial_1 u = 0 \text{ on } \eta_2 = 1\}.$$

Remark 2.2. *The Duffy transformation D maps the face $\eta_3 = 1$ to the point $(-1,-1,1)$ and the face $\eta_2 = 1$ to a line. An important property of $\widetilde{\mathcal{Q}}_p$ is that $u \in \mathcal{P}_p(\widehat{K})$ implies $u \circ D \in \widetilde{\mathcal{Q}}_p$.* ∎

Proof of Theorem 2.5. We will only show (8) as (9) follows easily from Lemma 2.4. Let u be such that $\widehat{u} := u \circ D \in \widetilde{\mathcal{Q}}_p$. In view of the positivity of the quadrature weights and Lemma 3.1 we get for $\widetilde{E} := \operatorname{diag}((D')^{-1}(D')^{-\top})$

$$\begin{aligned} \mathrm{GLJ}_{\widehat{K},q}(\nabla u \cdot A \nabla u) &\geq \lambda_{min}\, \mathrm{GLJ}_{\widehat{K},q}(|\nabla u|^2) \\ &= \lambda_{min}\, \mathrm{GLJ}_{\mathcal{Q},q}(\nabla \widehat{u} \cdot (D')^{-1}(D')^{-\top} \nabla \widehat{u}) \geq \frac{\lambda_{min}}{3468}\, \mathrm{GLJ}_{\mathcal{Q},q}(\nabla \widehat{u} \cdot \widetilde{E} \nabla \widehat{u}). \end{aligned}$$

A calculation reveals $\widetilde{E} = \left(E^{(1)}\right)^2 + \left(E^{(2)}\right)^2$ if we introduce

$$\begin{aligned} E^{(1)} &:= \operatorname{diag}\left\{ \frac{\sqrt{8}(1+\eta_1)}{(1-\eta_2)(1-\eta_3)}, \frac{1+\eta_2}{1-\eta_3}, 1 \right\}, \\ E^{(2)} &:= \operatorname{diag}\left\{ \frac{4}{(1-\eta_2)(1-\eta_3)}, \frac{2}{1-\eta_3}, 0 \right\} \end{aligned}$$

The assumption $\widehat{u} \in \widetilde{\mathcal{Q}}_p$ implies that the components of $E^{(1)}\nabla\widehat{u}$ and $E^{(2)}\nabla\widehat{u}$ are in $\mathcal{Q}_p(\mathcal{Q})$; hence, from Lemma 2.4

$$\begin{aligned} \mathrm{GLJ}_{\mathcal{Q},q}(\nabla \widehat{u} \cdot \widetilde{E} \nabla \widehat{u}) &= \mathrm{GLJ}_{\mathcal{Q},q}(|E^{(1)}\nabla\widehat{u}|^2) + \mathrm{GLJ}_{\mathcal{Q},q}(|E^{(2)}\nabla\widehat{u}|^2) \\ &\geq \int_{\mathcal{Q}} |E^{(1)}\nabla\widehat{u}|^2 |\det D'| \mathrm{d}\Omega + \int_{\mathcal{Q}} |E^{(2)}\nabla\widehat{u}|^2 |\det D'| \mathrm{d}\Omega \\ &= \int_{\mathcal{Q}} (\nabla\widehat{u})^{\top} \widetilde{E} \nabla\widehat{u} \, |\det D'| \mathrm{d}\Omega \\ &\geq \frac{1}{3}\int_{\mathcal{Q}} (\nabla\widehat{u})^{\top}(D')^{-1}(D')^{-\top}\nabla\widehat{u} \, |\det D'| \mathrm{d}\Omega = \frac{1}{3}\int_{\widehat{K}} |\nabla u|^2 \mathrm{d}\Omega, \end{aligned}$$

where we also appealed to Lemma 3.1. Collecting our findings, we arrive at

$$\mathrm{GLJ}_{\widehat{K},q}(\nabla u \cdot A \nabla u) \geq \frac{\lambda_{min}}{3468}\frac{1}{3}\|\nabla u\|^2_{L^2(\widehat{K})} \geq \frac{\lambda_{min}}{10404\lambda_{max}} \int_{\widehat{K}} \nabla u \cdot A \nabla u \mathrm{d}\Omega.$$

4 Numerical example

Cor. 2.9 states that the fully discrete p-FEM converges at the same rate as a Galerkin p-FEM where all integrals are evaluated exactly. We illustrate this behavior for the following example:

$$-\nabla \cdot (A \nabla u) = 1 \quad \text{on } \Omega := \widehat{K} \quad \text{and} \quad u = 0 \quad \text{on } \partial\Omega, \quad (15)$$

$$A(x_1,x_2,x_3) := \operatorname{diag}\left[\frac{1}{r^2+1}, \exp\left(r^2\right), \cos\left(\frac{1}{r^2+1}\right) \right], \quad (16)$$

Lemma 3.1. *Let* $E(\eta) := (D'^{-1}D'^{-\top})(\eta)$ *and denote by* $\operatorname{diag} E(\eta) \in \mathbb{R}^{3\times 3}$ *the diagonal of* $E(\eta)$. *Then*

$$\frac{1}{3468}\operatorname{diag} E(\eta) \le E(\eta) \le 3\operatorname{diag} E(\eta) \qquad \forall \eta \in \mathcal{Q}. \tag{12}$$

Proof. One easily shows for any invertible matrix $G \in \mathbb{R}^{n\times n}$

$$B \le A \quad \Longleftrightarrow \quad G^\top B G \le G^\top A G. \tag{13}$$

In order to prove (12), we define the diagonal matrix

$$B(\eta) := \operatorname{diag}\left[(1-\eta_2)(1-\eta_3), (1-\eta_3), 1\right]$$

and in view of (13) we are led to showing

$$\frac{1}{3468}(B^\top(\operatorname{diag} E)B)(\eta) \le (B^\top E B)(\eta) \le 3(B^\top(\operatorname{diag} E)B)(\eta) \quad \forall \eta \in \mathcal{Q}. \tag{14}$$

Explicitly computing

$$(B^\top E B)(\eta) = \begin{pmatrix} 8(1+\eta_1)^2+16 & (1+\eta_1)\{4+2(1+\eta_2)\} & 2(1+\eta_1) \\ \text{sym.} & 4+(1+\eta_2)^2 & (1+\eta_2) \\ \text{sym.} & \text{sym.} & 1 \end{pmatrix}$$

and applying the three estimates

$$\begin{aligned} &2(1+\eta_1)\{4+2(1+\eta_2)\}v_1v_2 \le 8(1+\eta_1)^2v_1^2 + [4+(1+\eta_2)^2]v_2^2, \\ &4(1+\eta_1)v_1v_3 \le 4(1+\eta_1)^2v_1^2 + v_3^2, \qquad 2(1+\eta_2)v_2v_3 \le (1+\eta_2)^2v_2^2 + v_3^2 \end{aligned}$$

for all $\eta \in \mathcal{Q}, v_1, v_2, v_3 \in \mathbb{R}$, we conclude for any vector $v = (v_1, v_2, v_3)^\top \in \mathbb{R}^3$

$$v^\top (B^\top E B)(\eta) v \le v^\top \operatorname{diag}\left[20(1+\eta_1)^2+16, 8+3(1+\eta_2)^2, 3\right] v.$$

In view of $(B^\top(\operatorname{diag} E)B)(\eta) = \operatorname{diag}\left[8(1+\eta_1)^2+16, 4+(1+\eta_2)^2, 1\right]$ we arrive at $(B^\top E B)(\eta) \le 3(B^\top(\operatorname{diag} E)B)(\eta)$. In order to prove the lower bound of (14) we observe that $(B^\top E B)(\eta)$ is symmetric positive definite for all $\eta \in \mathcal{Q}$; denoting by $0 < \lambda_1 \le \lambda_2 \le \lambda_3$ the three eigenvalues of $(B^\top E B)(\eta)$, we conclude from the Gershgorin circle theorem $0 < \lambda_1 \le \lambda_2 \le \lambda_3 \le 68$ for all $\eta \in \mathcal{Q}$. Moreover, a direct calculation shows $\det(B^\top E B)(\eta) = 64$. Thus, $\lambda_1 \ge \det(B^\top E B)/\lambda_2^2 \ge 4/289$ for all $\eta \in \mathcal{Q}$. Hence for all $\eta \in \mathcal{Q}$

$$\begin{aligned} (B^\top E B)(\eta) &\ge \frac{4}{289} I \ge \frac{4}{289}\operatorname{diag}\left[\frac{8(1+\eta_1)^2+16}{48}, \frac{4+(1+\eta_2)^2}{8}, 1\right] \\ &\ge \frac{1}{3468}(B^\top(\operatorname{diag} E)B)(\eta). \end{aligned}$$

□

2.1 Gauss-Lobatto-Jacobi quadrature

2.1.1 Gauss-Lobatto-Jacobi quadrature in 1D

For $\alpha > -1$, $n \in \mathbb{N}$, the Gauss-Lobatto-Jacobi quadrature formula is given by

$$\mathrm{GLJ}_{(\alpha,n)}(f) := \sum_{i=0}^{n} \omega_i^{(\alpha,n)} f(x_i^{(\alpha,n)}) \approx \int_{-1}^{1} (1-x)^\alpha f(x)\mathrm{d}x; \tag{5}$$

(see, e.g., [Karniadakis and Sherwin, 1999, App. B]): the quadrature nodes $x_i^{(\alpha,n)}$, $i = 0, \dots, n$, are the zeros of the polynomial $x \mapsto (1-x^2)P_n^{(\alpha+1,1)}(x)$, where $P_n^{(\alpha,\beta)}$ denotes the Jacobi polynomial of degree n with respect to the weight function $(1-x)^\alpha(1+x)^\beta$. The quadrature weights $\omega_i^{(\alpha,n)}$, $i = 0, \dots, n$, are *positive* and explicit formulas can be found, for example, in [Karniadakis and Sherwin, 1999, App. B]. We have:

Lemma 2.3. *Let* $\mathcal{P}_n$ *be the space of polynomials of degree* n. *Then for* $\alpha > -1$*:*

1. *For all* $f \in \mathcal{P}_{2n-1}$ *there holds* $\mathrm{GLJ}_{(\alpha,n)}(f) = \int_{-1}^{1} f(x)(1-x)^\alpha dx$.
2. *For all* $f \in \mathcal{P}_n$ *there holds*

$$\int_{-1}^{1} f^2(x)(1-x)^\alpha \mathrm{d}x \le \mathrm{GLJ}_{\alpha,n}(f^2) \le \left(2 + \frac{\alpha+1}{n}\right)\int_{-1}^{1} f^2(x)(1-x)^\alpha \mathrm{d}x.$$

Proof. The first assertion is well-known. The second assertion follows by the same arguments as in the case $\alpha = 0$, which can be found, for example, in Canuto and Quarteroni [1982] or [Bernardi and Maday, 1992, Cor. 1.13]. □

2.1.2 Gauss-Lobatto-Jacobi quadrature on $\widehat{K}$

Using the change of variables formula $\int_{\widehat{K}} g\mathrm{d}x = \int_{\mathcal{Q}} (g \circ D)|\det D'|\mathrm{d}x$, we can introduce a quadrature formulas such that

$$\mathrm{GLJ}_{\mathcal{Q},n}(f) \approx \int_{\mathcal{Q}} f(\eta)|\det D'(\eta)|\mathrm{d}\eta, \qquad \mathrm{GLJ}_{\widehat{K},n}(g) \approx \int_{\widehat{K}} g(\xi)\,\mathrm{d}\xi$$

by setting

$$\begin{aligned} \mathrm{GLJ}_{\mathcal{Q},n}(f) &:= 1/8 \sum_{i_1,i_2,i_3=0}^{n} \omega_{i_1}^{(0,n)}\omega_{i_2}^{(1,n)}\omega_{i_3}^{(2,n)} f\left(x_{i_1}^{(0,n)}, x_{i_2}^{(1,n)}, x_{i_3}^{(2,n)}\right), & (6) \\ \mathrm{GLJ}_{\widehat{K},n}(g) &:= \mathrm{GLJ}_{\mathcal{Q},n}(g \circ D). & (7) \end{aligned}$$

Using standard tensor product arguments one can deduce from the properties of the quadrature rules $\mathrm{GLJ}_{\alpha,n}$ and the formula (4) the following result:

Lemma 2.4. *Let $1 \le p \le q$ and let $\widehat{u} \in \mathcal{Q}_p(\mathcal{Q})$, $\widehat{v} \in \mathcal{Q}_{2q-1}(\mathcal{Q})$. Set $u := \widehat{u} \circ D^{-1}$, $v := \widehat{v} \circ D^{-1}$. Then the equalities $\mathrm{GLJ}_{\mathcal{Q},q}(\widehat{v}) = \int_{\mathcal{Q}} \widehat{v} |\det D'| \mathrm{d}\Omega$ and $\mathrm{GLJ}_{\widehat{K},q}(v) = \int_{\widehat{K}} v \mathrm{d}\Omega$ are true and, for $\underline{C} := (2+1/p)(2+2/p)(2+3/p) \le 60$,*

$$\begin{aligned} \int_{\mathcal{Q}} |\widehat{u}|^2 |\det D'| d\Omega \le \quad & \mathrm{GLJ}_{\mathcal{Q},q}(\widehat{u}^2) \quad \le \underline{C} \int_{\mathcal{Q}} |\widehat{u}|^2 |\det D'| d\Omega, \\ \|u\|^2_{L^2(\widehat{K})} \le \quad & \mathrm{GLJ}_{\widehat{K},q}(u^2) \quad \le \underline{C} \|u\|^2_{L^2(\widehat{K})}. \end{aligned}$$

2.2 Discrete Stability

The following discrete stability result is the heart of the quadrature error analysis; its proof is deferred to Section 3.

Theorem 2.5. *Let $A \in C(\overline{\widehat{K}}, \mathbb{R}^{3\times 3})$ be pointwise symmetric positive definite, $c \in C(\overline{\widehat{K}})$. Assume the existence of λ_{min}, λ_{max}, $c_{min} > 0$ with*

$$\lambda_{min} \le A(x) \le \lambda_{max}, \qquad c_{min} \le c(x) \forall x \in \widehat{K}.$$

Then for $q \ge p$ there holds for all $u \in \{u \,|\, u \circ D \in \widetilde{\mathcal{Q}}_p\}$

$$\mathrm{GLJ}_{\widehat{K},q}(\nabla u \cdot A \nabla u) \;\ge\; \frac{\lambda_{min}}{10404} \|\nabla u\|^2_{L^2(\widehat{K})} \ge \frac{\lambda_{min}}{10404 \lambda_{max}} \int_{\widehat{K},q} \nabla u \cdot A \nabla u \mathrm{d}\Omega, \quad (8)$$

$$\mathrm{GLJ}_{\widehat{K},q}(cu^2) \;\ge\; c_{min} \|u\|^2_{L^2(\widehat{K})}. \quad (9)$$

2.3 Convergence Analysis of Fully Discrete p-FEM

For the model problem (1) and given mesh $\mathcal{T}$ consisting of (curvilinear) tetrahedra with element maps $F_K : \widehat{K} \to K$, we define the discrete bilinear form a^q and right-hand side F^q by

$$\begin{aligned} a^q(u,v) \;&:=\; \sum_{K\in\mathcal{T}} \mathrm{GLJ}_{\hat{K},q}\left(((\nabla u \cdot A \nabla v)|_K \circ F_K) |\det F_K'| \right), \\ F^q(u) \;&:=\; \sum_{K\in\mathcal{T}} \mathrm{GLJ}_{\hat{K},q}\left(((f u)|_K \circ F_K) |\det F_K'| \right). \end{aligned}$$

We let $S_0^{p,1}(\mathcal{T}) := \{u \in H^1_0(\Omega) \,|\, u|_K \circ F_K \in \mathcal{P}_p(\widehat{K}) \quad \forall K \in \mathcal{T}\}$ and consider finite dimensional spaces V_N satisfying

$$S_0^{p,1}(\mathcal{T}) \subset V_N \subset \widetilde{S}_0^{p,1}(\mathcal{T}) := \{u \in H^1_0(\Omega) \,|\, u|_K \circ F_K \circ D \in \widetilde{Q}_p \quad \forall K \in \mathcal{T}\}. \quad (10)$$

Remark 2.6. *By Remark 2.2, choosing $V_N = S_0^{p,1}(\mathcal{T})$ is admissible. Taking V_N larger than $S_0^{p,1}(\mathcal{T})$ permits adapting the shape functions to the quadrature points and permits efficient ways to generate the stiffness matrix, Eibner and Melenk [2006].* ∎

The fully discrete problem is then:

$$\text{Find } u_N \in V_N \text{ s.t.} \qquad a^q(u_N, v) = F^q(v) \qquad \forall v \in V_N. \quad (11)$$

The discrete stability result Theorem 2.5 for a single element is readily extended to meshes with several elements and existence and uniqueness of solutions to (11) follows. An application of Strang's Lemma then gives error estimates:

Theorem 2.7. *Let the mesh $\mathcal{T}$ be fixed and the element maps F_K be analytic on $\overline{\widehat{K}}$. Assume (10) and $q \ge p$. Let u solve (1) and u_N solve (11). Then there exist C, $b > 0$ depending only on Ω, the analytic data A, f of (1), and the analytic element maps F_K such that*

$$\|u - u_N\|_{H^1(\Omega)} \le C \left(\inf_{v \in S_0^p(\mathcal{T})} \|u - v\|_{H^1(\Omega)} + C r^3 e^{-b(2q+p-r)} \right)$$

for arbitrary $1 \le r \le \min\{p, 2(q-1)-p\}$.

Proof. The proof follows along the lines of [Melenk and Schwab, 1998, Secs. 4.2, 4.3]: Thm. 2.5 enables us to use a Strang lemma, and the resulting consistency terms can be made exponentially small by the analyticity of A, f, and the F_K. □

Remark 2.8. *It is worth stressing that analyticity of $\partial\Omega$ is not required in Thm. 2.7—only analyticity of the element maps is necessary. Hence, also piecewise analytic geometries are covered by Thm. 2.7. The requirement that A, f be analytic can be relaxed to the condition that $A|_K$, $f|_K$ be analytic on $\overline{K}$ for all elements.* ∎

We note that choosing $r = \lfloor p/2 \rfloor$ in Thm. 2.7 implies that the *rate of convergence* of the fully discrete p-FEM is typically the same as the Galerkin p-FEM in which all quadratures are performed exactly:

Corollary 2.9. *Assume the hypotheses of Theorem 2.7. Then:*

1. *If $\inf_{v\in S_0^{p,1}(\mathcal{T})} \|u - v\|_{H^1(\Omega)} = O(p^{-\alpha})$, then $\|u - u_N\|_{H^1(\Omega)} = O(p^{-\alpha})$.*
2. *If $\inf_{v\in S_0^{p,1}(\mathcal{T})} \|u - v\|_{H^1(\Omega)} = O(e^{-bp})$ for some $b > 0$, then there exists $b' > 0$ such that $\|u - u_N\|_{H^1(\Omega)} = O(e^{-b'p})$.*

3 Proof of Theorem 2.5

The heart of the proof of Theorem 2.5 consists in the assertion that for the Duffy transformation D, the matrix $(D')^{-1}(D')^{-\top}$ is equivalent to its diagonal. To that end, we recall for square matrices A, $B \in \mathbb{R}^{n\times n}$ the standard notation $A \le B$ which expresses $v^\top A v \le v^\top B v$ for all $v \in \mathbb{R}^n$. We have: