Preconditioned Chebyshev collocation methods and triangular finite elements

M.O. Deville
Unité de Mécanique Appliquée, Université Catholique de Louvain, B-1348 Louvain-La-Neuve, Belgium

E.H. Mund
Service de Métrologie Nucléaire, Université Libre de Bruxelles, B-1050 Brussels and Unité de Thermodynamique, Université Catholique de Louvain, B-1348 Louvain-La-Neuve, Belgium

V. Van Kemenade
Unité de Mécanique Appliquée, Université Catholique de Louvain, B-1348 Louvain-La-Neuve, Belgium

Abstract

This paper analyzes triangular finite elements for the preconditioning of Chebyshev collocation solutions of elliptic boundary value problems. Results are given for scalar model problems and for both Stokes and Navier–Stokes equations.

1. Introduction

Spectral collocation methods applied to elliptic problems lead to condition numbers of $O(N^4)$, where $N$ is the polynomial degree in one space dimension. This condition number occurs for Jacobi polynomials in general, and for Chebyshev polynomials, in particular, which will be the basic ingredient of the present paper. The numericists facing this difficulty resort to preconditioning techniques. In the last decade, finite elements have emerged as a primer tool for that purpose.

The algorithmic approach defines the collocation grid by the tensor product of a one-dimensional set of discretization points based on the Gauss–Lobatto–Chebyshev quadrature rule. A high order Lagrangian interpolant using Chebyshev polynomials defines the approximation of the dependent variable. The projection method rests upon a weighted residual technique with Dirac functions as projectors. The system of collocated equations is preconditioned by the finite element (FE) Galerkin method applied to the weak formulation of the problem. The FE basis is built upon tensor products of 1-D piecewise interpolants. The vertices of the finite element mesh coincide with the Chebyshev grid. An iterative procedure is set up through a Richardson method where the right-hand side represents the residual of the partial differential equation to be solved. One may also view the iterative process as a deferred correction method with a highly accurate computation of the residual.

Up to now, bilinear elements have proved to be very powerful on elliptic problems [1–3] as well as on hyperbolic or advective–diffusive type equations [4, 5]. The efficient performance of the FE preconditioning is due to the clustered eigenspectrum around the origin for the iteration operator which governs the convergence rate of the Richardson method. This is in contrast with the finite difference preconditioning which is twice as slow to reach the same level of accuracy.

Based on this background, FE preconditioning was applied to Stokes and Navier–Stokes equations [6, 7]. The two-dimensional 9-node Lagrangian element involving biquadratics for velocities and bilinear interpolants for pressures (Q2-Q1 element) was found to be excellent with impressive
convergence properties. A series of test problems was solved showing the feasibility of the technique. More recently, the method was extended to curved geometries by use of the transfinite interpolation technique [8]. The FE preconditioner employs isoparametric representations to approximate the geometry while the spectral part of the calculation still retains exponential convergence properties.

The purpose of this paper is to investigate FE bases which are not obtained through tensor products. Typically, we have triangular elements in mind. They are widely used especially in some industrial codes and constitute a cornerstone of FE mesh generation. Their extension to three-dimensional problems leads to tetrahedra. The bandwidth of the global algebraic solver is reduced and the computational complexity induced by the solution decreased. Therefore, we will be interested in the number of iterations needed to reach (for example) machine accuracy on smooth problems.

The paper is organized as follows. Section 2 treats elliptic model problems and analyzes eigenspectrum computations. Section 3 deals with Stokes and Navier–Stokes equations where P2–P1 elements are investigated. Finally, the last section draws some conclusions.

2. An elliptic model problem

Let us consider the following two-dimensional elliptic model problem with homogeneous Dirichlet boundary conditions:

\[ Lu = -\Delta u(x, y) + \lambda u(x, y) = f(x, y), \quad (x, y) \in D, \]
\[ u(x, y) = 0, \quad (x, y) \in \partial D, \quad (1) \]

where \( D \) is the open unit square \((0, 1) \times (0, 1)\), \( \Delta \) the Laplacian operator, \( \lambda \) a non-negative constant and \( f(x, y) \) a prescribed function.

The Chebyshev collocation method uses a high-order Lagrangian interpolant based on Chebyshev polynomials of the first kind. The collocation grid \( G \) is the tensor product of 1-D Gauss–Lobatto–Chebyshev quadrature points

\[ z_j = \cos \frac{\pi j}{N}, \quad j \in [0, N], \quad z = x \text{ or } y. \]

These abscissae are the roots of the equation

\[ (1 - z^2)T_N'(z) = 0, \]

where \( T_N' \) is the derivative of the Chebyshev polynomial of degree \( N \). An affine transformation is used to map the domain \( D \) onto the reference square \([-1, 1]^2\) where the Chebyshev polynomials are defined and the dependent variable \( u \) is expanded as a double series of polynomials of degree \( N_x \) and \( N_y \) in each spatial direction,

\[ u_N(x, y) = \sum_{i=0}^{N_x} \sum_{j=0}^{N_y} u_{ij} h_i(x) h_j(y). \]

The one-dimensional Lagrangian interpolants are given by

\[ h_i(z_k) = \frac{(-1)^{i+1}(1 - z^2)T_N'(z_k)}{\tilde{c}_i N_i(N_i - 1)}, \quad \tilde{c}_i = \tilde{c}_N = 2, \quad \tilde{c}_0 = 1, \quad \forall i \in [1, N_x - 1], \]

and satisfy the cardinality relation

\[ h_i(z_k) = \delta_{ik}. \]

with \( \delta_{ik} \) denoting the Kronecker symbol. The discrete equations are obtained by a weighted residual method where the residual comes from the insertion of Eq. (4) in (1) with Dirac functions as test functions on the interior nodes of \( G \):

\[ L_{\varepsilon}u_N = -\Delta u_N(x_i, y_j) + \lambda u_N(x_i, y_j) = f_N(x_i, y_j), \quad \forall (x_i, y_j) \in G \cap \Omega. \]
Introducing matrix notation we denote the set of Eqs. (7) in the sequel by

$$L_c \bar{u} = \bar{f}. \quad (8)$$

Analysis of the eigenspectrum of $L_c$ shows that its condition number grows as $N^4$, $N + 1$ being the number of degrees of freedom in one direction. In order to circumvent round-off errors, Eq. (8) is solved iteratively, using a preconditioning technique, written as

$$L_{ap} \bar{u}^{k+1} = L_{ap} \bar{u}^k - \alpha_k (L_c \bar{u}^k - \bar{f}), \quad k = 0, 1, \ldots, \quad (9)$$

where the ‘preconditioner’ $L_{ap}$ should be as close as possible to the collocation matrix $L_c$. The coefficient $\alpha_k$ is a relaxation parameter needed to ensure convergence of the Richardson method and superscript $k$, an iteration index. It is well known from linear algebra considerations that the iterative process Eq. (9) converges if the iteration operator

$$A = I - \alpha_k L_{ap}^{-1} L_c \quad (10)$$

has a spectral radius $\rho(A) < 1$. When the eigenvalues of $L_{ap}^{-1} L_c$ are real, an optimal value of the relaxation factor $\alpha_k$ is given by

$$\alpha_{opt} = \frac{2}{\lambda_m + \lambda_M}, \quad (11)$$

where $\lambda_m$ and $\lambda_M$ are the lower and upper bounds of the spectrum of $L_{ap}^{-1} L_c$.

DeviUe and Mund have shown that for a very large class of elliptic problems, the eigenvalues of $L_{ap}^{-1} L_c$ are located inside the unit circle (allowing for a relaxation factor $\alpha_k$ equal to 1) when the preconditioning algorithm uses low order FE. The most efficient preconditioner (in terms of numerical work) is the bilinear element $Q_1$ whose stencil on the reference square $(-1, 1) \times (-1, 1)$ is represented in Fig. 1(a) (cf. [1, 2]). Other elements might be used like the triangular element $P_1$, a combination of the elemental functions $\{1, x, y\}$, whose stencil is represented in Fig. 1(b). Let us recall that the vertices of the FE triangulation have to coincide with the nodes of the Gauss–Lobatto–Chebyshev grid. The loss in accuracy is partly compensated by a diminution of one unit in the half-bandwidth length of the stiffness matrix.

Figure 2 shows the eigenspectrum of the iterative operator Eq. (10) ($\alpha_k = 1$) associated to the preconditioning of the elliptic problem

$$Lu = -\nabla((1 + x^2 y^2)\nabla u(x, y)) + u(x, y), \quad (12)$$

with Dirichlet boundary conditions, on the unit square. The left part of the figure corresponds to the preconditioning with $Q_1$ elements whereas the right part of the figure relates to $P_1$ preconditioning. The spectral radius is larger in the latter case, leading to slower convergence. Figure 3 displays (with the same convention) the eigenspectrum when the preconditioner, instead of being the FE approximation of Eq. (12), is the FE approximation of

$$-\Delta u(x, y) + u(x, y). \quad (13)$$

Quite remarkably the spectrum is once again inside the unit circle.

---

Fig. 1. Reference FE stencils: (a) quadrilateral element $Q_1$, (b) triangular element $P_1$. 

---
Finally, Fig. 4 shows the convergence history for an analytical model problem corresponding to the Helmholtz operator $-\Delta u + u$. Slower convergence for the P1 preconditioning is clearly illustrated by the results.

3. Navier–Stokes equations and finite element preconditioning

Let us look at the steady Navier–Stokes equations written in stress formulation:

$$\rho (\mathbf{u} \cdot \nabla \mathbf{u}) = \nabla \cdot \sigma + \rho f,$$

$$\sigma = -pI + \mu (\nabla \mathbf{u} + (\nabla \mathbf{u})^T),$$

$$\nabla \cdot \mathbf{u} = 0,$$

where $\mathbf{v}$ is the velocity field, $p$ the pressure, $\mu$ the dynamic viscosity, $f$ the body force term, $\sigma$ the stress tensor, $I$ the unit tensor and $\rho$ the density. The superscript $T$ indicates the transpose. Equation (14) is the momentum equation. Equation (15) defines the constitutive relationship of a viscous Newtonian fluid. Equation (16) enforces the incompressibility constraint.

The boundary conditions may be split into two parts. On $\partial D_c$, essential boundary conditions are applied:

$$\mathbf{u}(\mathbf{r}) = \mathbf{g}_1(\mathbf{r}) \quad \forall \mathbf{r} \in \partial D_c.$$

Fig. 2. Eigenspectrum of the iteration operator Eq. (10): Q1 elements on the left and P1 on the right.

Fig. 3. Eigenspectrum of (10). The preconditioner is the FE approximation of Eq. (13). Same convention as in Fig. 2.
In Eq. (17), \( \vec{r} \) denotes the position vector. On \( \partial D_\alpha \), natural boundary conditions apply:

\[
\tau(\vec{r}) = \sigma \cdot n = g_2(\vec{r}), \quad \forall \vec{r} \in \partial D_\alpha.
\]  (18)

In (18), \( \tau \) is the stress vector acting on \( \partial D_\alpha \), where \( n \) is the unit outward normal vector.

A Newton's linearization technique is carried out on Eqs. (14) and (16). Denoting by \( \delta v \) and \( \delta \sigma \) the variation of \( v \) and \( \sigma \), respectively, the linearized Navier-Stokes equations are

\[
\rho \left[ (v^n \cdot \nabla)\delta v + (\delta v \cdot \nabla)v^n \right] = \text{div}(\sigma^n + \delta \sigma) + \rho f^n - \rho (v^n \cdot \nabla)v^n,
\]  (19)

\[
\text{div} \delta v = -\text{div} v^n,
\]  (20)

\[
v^{n+1} = v^n + \delta v, \quad \sigma^{n+1} = \sigma^n + \delta \sigma,
\]  (21)

where the superscript is an iteration counter of Newton's scheme. Eqs. (19)–(21) are subject to the following boundary conditions:

\[
\delta v = 0 \quad \forall \vec{r} \in \partial D_\infty,
\]  (22)

and

\[
\delta \tau = \delta \sigma \cdot n = -\tau^n + g_2(\vec{r}) \quad \forall \vec{r} \in \partial D_\alpha.
\]  (23)

Expanding each dependent variable \( v, p \) in a double Chebyshev series as in Eq. (4) and inserting them in the weighted residual projection method applied to Eqs. (14)–(16), we obtain the collocation equations which we shall write formally as

\[
Lc x = b,
\]  (24)

with

\[
x = (v, p)^T.
\]

The right-hand side \( b \) corresponds to the body forces and the imposed functions through the boundary conditions. The details to construct (24) are given in [7].

The collocation equations (24) are non-linear and ill-conditioned. In order to solve them, we use finite element preconditioning inside a Richardson iteration scheme. The algorithm is now
where $L_{FE}$ is a standard Galerkin finite element method applied to the Navier–Stokes equations (19)–(21). This is in contrast with the numerical procedure used formerly by Demaret and Deville [7] where two nested iterations were performed. The previous Newton (outer) iteration and the Richardson (inner) iterations have been merged in a single non-linear Richardson process as recommended by Boyd [9]. The convergence rate of (25) depends on the iteration operator

$$A = I - \alpha_n L_{FE} L_c,$$

which is no longer linear. However, if the continuous problem reduces to the Stokes equations, an eigenspectrum of (26) might be evaluated.

In Demaret and Deville [6, 7], the so-called Q2–Q1 finite element was extensively used. Here we concentrate on triangular elements and each quadrangle whose vertices coincide with the collocation grid $\mathcal{S}_h$ is subdivided in two triangles where P2–P1 finite element approximations are applied to the velocity and pressure fields, respectively.

A full solution of the Navier–Stokes equations (14)–(16) may be written in the general form:

$$u = -\cos\left(\frac{\pi}{2} x\right) \sin\left(\frac{\pi}{2} y\right), \quad v = \sin\left(\frac{\pi}{2} x\right) \cos\left(\frac{\pi}{2} y\right), \quad p = -\pi \sin\left(\frac{\pi}{2} x\right) \sin\left(\frac{\pi}{2} y\right),$$

$$f_x = -\frac{\pi^2}{\rho} \cos\left(\frac{\pi}{2} x\right) \sin\left(\frac{\pi}{2} y\right) - a \frac{\pi}{4} \sin(\pi x), \quad f_y = -a \frac{\pi}{4} \sin(\pi y).$$

where the constant $a$ is set equal to 1 for a Navier–Stokes problem and equal to 0 for the Stokes equation.

Figure 5 compares the convergence history with respect to the iteration number for the Stokes case when either Q2–Q1 quadrilaterals or P2–P1 triangles constitute the FE approach. The results were obtained with the POLYFLOW package [10]. One observes that to reach machine accuracy, triangular FE need about four times more iterations than quadrilaterals. In Fig. 6, we report the convergence history for a Navier–Stokes problem characterized by a Reynolds number equal to 100. Again, the triangular elements are less efficient than the quadrilaterals and need about twice as many iterations to achieve full accuracy. In all previous calculations, $N$ was set equal to 12 on a single domain. For the

![Fig. 5. Convergence analysis of the relative errors on a Stokes problem.](image-url)
Stokes problem, the relaxation parameter was set equal to $2/3$ (as in the Q2-Q1 case) while for the Navier–Stokes test, it was adjusted in a range between 0.6 and 0.3.

4. Discussion and conclusions

In two-dimensional problems, the dismal performance of triangular finite elements in terms of the number of iterations with respect to the conventional quadrilaterals is not balanced by a significant diminution of computing time per iteration. The reduction in bandwidth, one subdiagonal per half bandwidth, is only marginal. An extension of the method to three-dimensional problems with tetrahedrons replacing (expensive) triquadratic-trilinear parallelepipeds, might be more rewarding in terms of CPU computing cost. This remains to be demonstrated.

References