29

A Multigrid Iterated Penalty Method for Mixed Elements

Ridgway Scott and Shangyou Zhang

1 Introduction

We consider the following stationary Stokes problem: Find functions $\mathbf{u}$ (the fluid velocity) and p (the pressure) on a bounded polygonal/polyhedral domain $\Omega \subset \mathbf{R}^d$ ($d = 2, 3$), such that

$$\begin{aligned} -\Delta\mathbf{u} + \mathrm{grad}p &= \mathbf{f} && \text{in } \Omega, \\ \mathrm{div}\mathbf{u} &= 0 && \text{in } \Omega, \\ \mathbf{u} &= 0 && \text{on } \partial\Omega, \end{aligned} \tag{1}$$

where $\mathbf{f}$ is the body force. In the variation form of velocity-pressure formulation of the Stokes equations, the velocity and pressure are in the Sobolev spaces $H_0^1(\Omega)^d$ and $L_0^2(\Omega)$, respectively. The mixed element approximation spaces can be chosen to be the corresponding subspaces. A natural pairing would be continuous piecewise-polynomials of degree $(k+1)$ and discontinuous piecewise-polynomials of degree k for the velocity and pressure, respectively. Such mixed element solutions satisfy the incompressibility condition, i.e. pointwise divergence free. Scott and Vogelius [SV85] showed that the Babuška-Brezzi inequality holds for such $\mathcal{P}_{k+1}$-P_k triangular mixed-elements in 2D if the polynomial degree k is 3 or higher and if the meshes are singular-vertex free. This result is partially extended to 3D in [Zha94]. It is shown that, when defined on tetrahedral meshes of a macro-element type, the above $\mathcal{P}_{k+1}$-P_k elements are stable if the polynomial degree for velocity is 3 or higher.

The mixed elements approximation to (1) in weak formulation is: Find $[\mathbf{u}_j, p_j] \in V_j$, such that

$$a(\mathbf{u}_j, \mathbf{v}) + b(\mathbf{v}, p_j) + b(\mathbf{u}_j, q) = (\mathbf{f}, \mathbf{v}) \quad \forall [\mathbf{v}, q] \in V_j, \tag{2}$$

where $a(\mathbf{u}, \mathbf{v}) := (\nabla\mathbf{u}, \nabla\mathbf{v})$ and $b(\mathbf{v}, p) := -(\mathrm{div}\mathbf{v}, p)$. Here $\{V_j\}$ are multilevel $\mathcal{P}_{k+1}$-P_k mixed finite element spaces:

$$V_j = (\mathcal{P}^0_{k+1,\mathcal{T}_j})^d \times P^0_{k,\mathcal{T}_j} \subset (H_0^1(\Omega))^d \times L_0^2(\Omega),$$

Ninth International Conference on Domain Decomposition Methods
Editor Petter E. Bjørstad, Magne S. Espedal and David E. Keyes

defined on nested or nonnested triangular/tetrahedral grids $\{\mathcal{T}_j\}$ covering Ω, where $\mathcal{P}^0_{k,\mathcal{T}_j} = \mathcal{P}_{k,\mathcal{T}_j} \cap H^1_0(\Omega)$, $P^0_{k,\mathcal{T}_j} = P_{k,\mathcal{T}_j} \cap L^2_0(\Omega)$, $\mathcal{P}_{k,\mathcal{T}_j}$ and $P_{k,\mathcal{T}_j}$ are the piecewise continuous and discontinuous polynomials of degree k on the mesh $\mathcal{T}_j$, respectively.

We note that for $\mathcal{P}_{k+1}$-P_k elements, the divergence of the discrete velocity space is precisely the discrete pressure space, i.e., $\{\mathrm{div}\mathbf{u} \,|\, \mathbf{u} \in (\mathcal{P}^0_{k+1,\mathcal{T}_j})^d\} = P^0_{k,\mathcal{T}_j}$. Therefore, we can avoid introducing the discrete space $P^0_{k,\mathcal{T}_j}$ in computation completely in the following iterated penalty method (3), which is a special form of the augmented Lagrangian method [FG83] . Let $r \geq 1$. The iterated penalty method [BS94] defines $\mathbf{u}^n \in (\mathcal{P}^0_{k+1,\mathcal{T}_j})^d$ by

$$\begin{aligned} a(\mathbf{u}^n, \mathbf{v}) + r(\mathrm{div}\mathbf{u}^n, \mathrm{div}\mathbf{v}) &= (\mathbf{f}, \mathbf{v}) + (\mathrm{div}\mathbf{v}, \mathrm{div}\mathbf{w}^n) \qquad \forall \mathbf{v} \in (\mathcal{P}^0_{k+1,\mathcal{T}_j})^d, \\ \mathbf{w}^{n+1} &= \mathbf{w}^n + r\mathbf{u}^n \end{aligned} \tag{3}$$

sequentially given $\mathbf{w}^0 \in (\mathcal{P}^0_{k+1,\mathcal{T}_j})^d$ (which is usually $\mathbf{0}$). $p^n = \mathrm{div}\mathbf{w}^n \in P^0_{k,\mathcal{T}_j}$. The key point of the iterated penalty method is that the system of equations represented by the first equation in (3) for $\mathbf{u}^n$ will be symmetric and positive definite. Each time when we solve the linear system (3), we apply the multigrid method to get an inner iteration. Another essential point here is that we only do one or two multigrid iterations for (3), which needs not to be solved accurately as the solution $\mathbf{u}^n$ is not final. We show that the overall convergence rate of the combined iterated penalty – multigrid method is independent of the size of the discrete system, and that the overall computation cost for solving the discrete Stokes equations up to the order of approximation is proportional to the system size if the penalty parameter r is chosen not too big. Therefore the method is optimal in the order of computation. The analysis is confirmed by our numerical computation.

2 Analysis

We first analyze the the iterated penalty method. We will define a multigrid iterated penalty method. Then we will apply the general theory developed in [BP87], [BPX91] and [SZ92] to obtain a convergence analysis of the multigrid iterated penalty method.

Theorem 1. (Convergence of the iterated penalty method) *For $\mathbf{u}^n$ defined in (3), the following error reduction relation holds:*

$$|||\mathbf{u}_j - \mathbf{u}^n|||_{1,j} \leq \frac{C}{r} |||\mathbf{u}_j - \mathbf{u}^{n-1}|||_{1,j}$$

where C is independent of j.

PROOF. Let the errors be denoted as $\mathbf{e}^n = \mathbf{u}^n - \mathbf{u}_j$ and $\epsilon^n = p^n - p_j = \mathrm{div}\mathbf{w}^n - p_j$. We will use a short notation

$$\mathbf{U}_j := (\mathcal{P}^0_{k+1,\mathcal{T}_j})^d.$$

Subtracting (2) from (3), it follows

$$a(\mathbf{e}^n, \mathbf{v}) + r(\mathrm{div}\mathbf{e}^n, \mathrm{div}\mathbf{v}) = b(\mathbf{v}, \epsilon^n) \quad \forall \mathbf{v} \in \mathbf{U}_j, \tag{4}$$

$$\epsilon^{n+1} = \epsilon^n + r\mathrm{div}\mathbf{u}^n = \epsilon^n + r\mathrm{div}\mathbf{e}^n. \tag{5}$$

Hence, combining the above two equations we can get

$$\begin{aligned} a(\mathbf{e}^{n+1},\mathbf{e}^{n+1})+r(\mathrm{div}\mathbf{e}^{n+1},\mathrm{div}\mathbf{e}^{n+1}) &= b(\mathbf{e}^{n+1},\epsilon^{n+1}) \qquad (6)\\ &= b(\mathbf{e}^{n+1},\epsilon^{n})-r(\mathrm{div}\mathbf{e}^{n+1},\mathrm{div}\mathbf{e}^{n})\\ &= a(\mathbf{e}^{n+1},\mathbf{e}^{n})\le|\mathbf{e}^{n+1}|_{H^1}|\mathbf{e}^{n}|_{H^1}. \qquad (7)\end{aligned}$$

By (4), $a(\mathbf{e}^n,\mathbf{v})=0$ for all divergence-free function $\mathbf{v}\in\mathbf{U}_j$. Therefore, for our special choices of $\mathbf{U}_j$, we have

$$(\nabla\times\mathbf{e}^n,\nabla\times\mathbf{e}^n)=0 \quad \text{and} \quad a(\mathbf{e}^n,\mathbf{e}^n)=(\mathrm{div}\mathbf{e}^n,\mathrm{div}\mathbf{e}^n).$$

The theorem is proven with $C=1$. ∎

By Theorem 1, we can see that $|||\mathbf{u}_j-\mathbf{u}^n|||_{1,j}=O(r^{-n})$. One may like to choose a very large r to get a fast convergence. However, large n will cause a bad conditioning for the linear system in (3), which will in turn increase the work of the multigrid method when solving the linear system.

In the multigrid method, we have two steps, the fine-level smoothing and the coarse-level correction. The fine-level smoothing is usually the Richardson iteration, where we can introduce an L^2-equivalent discrete inner-product, $(\cdot,\cdot)_j$, on $\mathbf{U}_j$ defined by the diagonal entries of the L^2 inner-product for the nodal basis. We note the equivalence constant depends on the polynomial degree, $k+1$, but is independent of the level number j. We start by defining a family of symmetric positive definite operators, $A_j:\mathbf{U}_j\to\mathbf{U}_j$,

$$(A_j\mathbf{u},\mathbf{v})_j=a_r(\mathbf{u},\mathbf{v}):=a(\mathbf{u},\mathbf{v})+r(\mathrm{div}\mathbf{u},\mathrm{div}\mathbf{v}) \quad \forall\mathbf{u},\mathbf{v}\in\mathbf{U}_j$$

Let $\rho(A_j)$ be the spectral radius of A_j. Then it is well known by the inverse inequality that

$$\rho(A_j)=Ch_j^{-2}r, \tag{8}$$

where the constant C depends on k but not on j. We assume the following regularity for the solution of the continuous version of the first equation (3):

$$\|\mathbf{u}_g\|_{H^{1+\alpha}(\Omega)^d}\le C\|\mathbf{g}\|_{H^{-1+\alpha}(\Omega)^d} \qquad \forall\mathbf{g}\in H^{-1+\alpha}(\Omega)^d, \tag{9}$$

where $\alpha>0$ and $\mathbf{u}_g$ is defined by $a_r(\mathbf{u}_g,\mathbf{v})=(\mathbf{g},\mathbf{v}) \quad \forall\mathbf{v}\in H_0^1(\Omega)^d$. In (9) the constant C is independent of the penalty parameter r. By the standard finite element theory, we have the following estimate in the energy norm

$$|||\mathbf{u}_g-\mathrm{P}_j\mathbf{u}_g|||_{1,j}\le Crh^s\|\mathbf{u}_g\|_{H^{1+s}(\Omega)^d}, \quad s=\min\{k+1,\alpha\}, \tag{10}$$

where $\mathrm{P}_j\mathbf{u}_g$ denotes the finite element solution for $\mathbf{u}_g$. Here the triple-bar norms are defined by $|||\mathbf{v}|||_{s,j}^2:=(A_j^s\mathbf{v},\mathbf{v})_j, \quad 0\le s\le 2$.

Definition 1 (One level j W-cycle symmetric nested/nonnested multigrid iteration)**.**
1. For $j=1$, the problem (3) or the residual problem (12) below is solved exactly.
2. For $j>1$, $\mathbf{w}_{2m+1}$ will be generated from the initial guess, $\mathbf{w}_0$ (which is either 0

or a previous solution on the same level, the s below is great than 1), as follows.
2-a. m presmoothings are performed to generate $\mathbf{w}_m$:

$$(\mathbf{w}_l - \mathbf{w}_{l-1}, \mathbf{v})_j = \lambda_j^{-1}(\mathbf{F}(\mathbf{v}) - a_r(\mathbf{w}_{l-1}, \mathbf{v})) \quad \forall \mathbf{v} \in \mathbf{U}_j, \quad l = 1, 2, \ldots, m, \tag{11}$$

where $\rho(A_j)/\lambda_k \le \omega$ for some fixed ω satisfying $0 < \omega < 2$ and $\mathbf{F}$ is either the functional defined in the right-hand side of (3) or the $\tilde{\mathbf{F}}$ in (12).
2-b. $\mathbf{w}_m$ is corrected by $\mathrm{I}_j q$ to generate w_{m+1}: Let $\bar{q}$ solve the following coarse–level residual problem,

$$a_r(\bar{\mathbf{q}}, \mathbf{v}) = \mathbf{F}(\mathrm{I}_j \mathbf{v}) - a_r(\mathbf{w}_m, \mathrm{I}_j \mathbf{v}) =: \tilde{\mathbf{F}}(\mathbf{v}) \quad \forall \mathbf{v} \in \mathbf{U}_{j-1}. \tag{12}$$

Let $\mathbf{q} \in \mathbf{U}_{j-1}$ be the approximation of $\bar{\mathbf{q}}$ obtained by applying s ($s = 2$) iterations of the $(j-1)$st level multigrid scheme to (12) starting with initial guess zero. Then $\mathbf{w}_{m+1} = \mathbf{w}_m + \mathrm{I}_j q$. Here, $\mathrm{I}_j : \mathbf{U}_{j-1} \to \mathbf{U}_j$ is the usual Lagrange interpolation operator if $\mathbf{U}_{j-1} \not\subset \mathbf{U}_j$, or just an identity operator if $\mathbf{U}_{j-1} \subset \mathbf{U}_j$.
2-c. m postsmoothings of the form (11) are performed to generate w_{2m+1} from w_{m+1}.
▮

Definition 2 (A multigrid iterated penalty method)**.** Let $\tilde{\mathbf{w}}^0 = 0$. $\tilde{\mathbf{u}}^n$ is defined by doing l jth level W-cycle symmetric nested/nonnested multigrid iteration for the problem defined by the first equation in (3) where $\mathbf{w}^n$ being replaced by $\tilde{\mathbf{w}}^n$, i.e., for the equation

$$a(\bar{\mathbf{u}}^n, \mathbf{v}) + r(\mathrm{div}\bar{\mathbf{u}}^n, \mathrm{div}\mathbf{v}) = (\mathbf{f}, \mathbf{v}) + (\mathrm{div}\mathbf{v}, \mathrm{div}\tilde{\mathbf{w}}^n) \quad \forall \mathbf{v} \in \mathbf{U}_j. \tag{13}$$

Here the initial guess for the multigrid iteration is $\tilde{\mathbf{u}}^{n-1}$ and $\tilde{\mathbf{u}}^{-1} = \mathbf{0}$. $\tilde{\mathbf{w}}^n$ is defined by $\tilde{\mathbf{w}}^n = \tilde{\mathbf{w}}^{n-1} + r\mathrm{div}\tilde{\mathbf{u}}^n$. ▮

By the assumptions (8–9) and (10) it is standard to verify the "regularity and approximation" assumption introduced by Bramble *et al.* (see (3.2) in [BP87], also [BPX91] and [SZ92]):

$$a_r(\mathbf{u} - \mathrm{P}_{j-1}, \mathbf{u}) \le C_\beta^2 \left(\frac{|||A_j \mathbf{u}|||_{0,j}^2}{\lambda_j} \right)^\beta a_r(\mathbf{u}, \mathbf{u})^{1-\beta} \quad \forall \mathbf{u} \in \mathbf{U}_j$$

with $C_\beta^2 = Cr^{1+\alpha/2}$ and $\beta = \alpha/2$ (α is defined in (9)). Therefore, we can get the following theorem by the theory of [BP87].

Theorem 2. *The error reduction factor in the $|||\cdot|||_{1,j}$ norm for one jth level W-cycle symmetric nested/nonnested multigrid iteration is bounded by*

$$\gamma := \left(\frac{Cr^{1/\alpha+1/2}}{Cr^{1/\alpha+1/2} + m} \right)^\alpha < 1,$$

where α is introduced in (9). ▮

Theorem 3. *The error reduction factor in the $|||\cdot|||_{1,j}$ norm for the multigrid iterated penalty method defined in Definition 2 can be estimated by*

$$|||\tilde{\mathbf{u}}^{n+1} - \mathbf{u}_j|||_{1,j} \le \left(\frac{C}{r} + \gamma^l\right) |||\tilde{\mathbf{u}}^n - \mathbf{u}_j|||_{1,j},$$

where γ is defined in Theorem 2.

PROOF. Let the errors be denoted by $\tilde{\mathbf{e}}^n = \tilde{\mathbf{u}}^n - \mathbf{u}_j$. Let $\bar{\mathbf{u}}^n$ be the exact solution of the first equation in (3) with $\mathbf{w}^n$ there being replaced by $\tilde{\mathbf{w}}^n$. Repeat (4 – 6) in the proof for Theorem 1, it follows also that

$$|||\bar{\mathbf{u}}^{n+1} - \mathbf{u}_j|||_{1,j} \le \frac{C}{r}|||\tilde{\mathbf{u}}^n - \mathbf{u}_j|||_{1,j}. \tag{14}$$

By (14) and Theorems , we get that

$$\begin{aligned}
|||\tilde{\mathbf{e}}^{n+1}|||_{1,j} &\le |||\tilde{\mathbf{u}}^{n+1} - \bar{\mathbf{u}}^{n+1}|||_{1,j} + |||\bar{\mathbf{u}}^{n+1} - \mathbf{u}_j|||_{1,j} \\
&\le \gamma^s|||\tilde{\mathbf{u}}^n - \bar{\mathbf{u}}^{n+1}|||_{1,j} + |||\bar{\mathbf{u}}^{n+1} - \mathbf{u}_j|||_{1,j} \\
&\le \gamma^s|||\tilde{\mathbf{e}}^n|||_{1,j} + (1+\gamma^s)|||\bar{\mathbf{u}}^{n+1} - \mathbf{u}_j|||_{1,j}. \\
&\le \gamma^s|||\tilde{\mathbf{e}}^n|||_{1,j} + \frac{C(1+\gamma^s)}{r}|||\tilde{\mathbf{e}}^n|||_{1,j}. \blacksquare
\end{aligned}$$

According to Theorem 3, we have a constant convergence-rate for the multigrid iterated penalty method. By it, with a full multigrid iteration, one gets in a standard way the optimal order of computation for the algorithm (cf. [SZ92]), i.e., the work to solve (2) up to the finite element truncation error is proportional to the number of unknown in the linear system (2).

3 Numerical Test

Table 1 Number of iteration for the iterated penalty method ($l = 1$, $k = 3$, $m = 10$, $r = 2$)

Grid level	2	3	4	5	6
Iteration number	20	9	9	8	8

In our numerical test, we let $\Omega = (0,1) \times (0,1)$. We let $\mathbf{u} = \nabla \times g$ and $p = \Delta g$ be the exact solutions for (1), where $g = 100(x_1 - x_1^2)^2(x_2 - x_2^2)^2$. The first level grid consists of two triangles, and higher level grids are defined by refining the previous level triangles into 4 subtriangles. In our computation, we use the V-cycle, symmetric multigrid iteration with $m = 10$ as the inner iteration. The iteration number of the iterated penalty method is listed in the Table 1 for degree 4 polynomial approximation

Table 2 Number of iteration for the iterated penalty method ($l = 1$, $k = 6$, $m = 10$, $j = 3$)

Penalty parameter r	1/2	1	3	5	10	100
Iteration number	24	10	9	9	12	37

of $\mathbf{u}$ on different grids, where the penalty parameter $r = 2$. The stopping criterion is 2×10^{-7} for $a(\mathbf{e}^n, \mathbf{e}^n)$.

Next we use degree 7 polynomials to approximate the velocity $\mathbf{u}$. In this case, the mixed finite element solutions are exact, the same as $\nabla \times g$. We numerically test the dependence of the iteration number of the iterated penalty method on the penalty parameter r. We can see from Table 2 that for r between 1 and 10, the number of iteration seems to be the least. But with bigger r, we have to increase the smoothing parameter m or use W-cycle multigrid methods to get a more accurate multigrid solution for each penalty problem. This is indicated by Theorem 3. For example, if we let $m = 50$, then the iteration number would become 5 if $r = 100$. But the overall work is about 3 times as much as that for the case $m = 10$ and $r = 5$.

REFERENCES

[BP87] Bramble J. H. and Pasciak J. E. (1987) New convergence estimates for multigrid algorithms. *Math. Comp.* 49: 311–329.

[BPX91] Bramble J. H., Pasciak J. E., and Xu J. (1991) The analysis of multigrid algorithms with nonnested spaces or non-inherited quadratic forms. *Math. Comp.* 56: 1–34.

[BS94] Brenner S. C. and Scott L. R. (1994) *The Mathematical Theory of Finite Element Methods*. Springer-Verlag, New York.

[FG83] Fortin M. and Glowinski R. (1983) *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-value Problems*. North Holland, Amsterdam.

[SV85] Scott L. R. and Vogelius M. (1985) Norm estimates for a maximal right inverse of the divergence operator in spaces of piecewise polynomials. *Math. Modelling Numer. Anal.* 19: 111–143.

[SZ92] Scott L. R. and Zhang S. (1992) Higher dimensional non-nested multigrid methods. *Math. Comp.* 58: 457–466.

[Zha94] Zhang S. (1994) A new family of stable mixed finite elements for the 3D Stokes equations. *Center for the Mathematics of Waves, University of Delaware* Tech. Report 94-18: 457–466.