[31] Christopher C. Paige and Michael A. Saunders. Solutions of sparse linear systems of linear equations. *SIAM J. Numer. Anal.*, 12:617–629, 1975.

[32] Einar M. Rønquist. A domain decomposition solver for the incompressible Navier-Stokes equations. 1994. Workshop on Spectral Element Methods, North Carolina State University.

[33] Torgeir Rusten and Ragnar Winther. A preconditioned iterative method for saddle point problems. *SIAM J. Matrix Anal. Appl.*, 13:887–904, 1992.

[34] Marcus Sarkis. Two-level Schwarz methods for nonconforming finite elements and discontinuous coefficients. In N. Duane Melson, Thomas A. Manteuffel, and Steve F. McCormick, editors, *Proceedings of the Sixth Copper Mountain Conference on Multigrid Methods, Volume 2*, number 3224, pages 543–566, Hampton VA, 1993. NASA.

[35] Marcus Sarkis. *Schwarz preconditioners for elliptic problems with discontinuous coefficients using conforming and non-conforming elements.* PhD thesis, Courant Institute of the Mathematical Sciences, New York University, 1994.

[36] Morten D. Skogen. *Schwarz Methods and Parallelism.* PhD thesis, Department of Informatics, University of Bergen, Norway, February 1992.

[37] Gerard L.G. Sleijpen, Henk A. Van der Vorst, and Diederik R. Fokkema. BiCGstab(l) and other hybrid Bi-CG methods. Technical report, University of Utrecht, February 1994.

[38] Barry F. Smith. *Domain decomposition algorithms for the partial differential equations of linear elasticity.* PhD thesis, Courant Institute of the Mathematical Sciences, New York University, 1990.

[39] Barry F. Smith. A parallel implementation of an iterative substructuring algorithm for problems in three dimensions. *SIAM J. Sci. Comput.*, 14(2):406–423, 1993.

[40] Barry F. Smith. Extensible PDE solvers package users manual. Technical Report ANL 94/40, Argonne National Laboratory, September 1994.

[41] David Sylvester and Andrew Wathen. Fast iterative solutions of stabilised Stokes systems Part II: Using general block preconditioners. *SIAM J. Numer. Anal.*, 31(5):1352–1367, 1994.

[42] Rüdiger Verfürth. Error estimates for a mixed finite element approximation of the Stokes equations. *R.A.I.R.O. Num. Anal.*, 18(2):175–182, 1984.

[43] Andrew Wathen and David Sylvester. Fast iterative solution of stabilised Stokes systems Part I: Using simple diagonal preconditioners. *SIAM J. Numer. Anal.*, 30(3):630–649, 1993.

[44] Jinchao Xu. *Theory of Multilevel Methods.* PhD thesis, Cornell University, May 1989.

[45] Xuejun Zhang. Multilevel Schwarz methods. *Numer. Math.*, 63(4):521–539, 1992.

[9] James H. Bramble and Joseph E. Pasciak. A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems. *Math. Comp.*, 50:1–17, 1988.

[10] James H. Bramble, Joseph E. Pasciak, and Apostol E. Vasilev. Analysis of the inexact Uzawa algorithm for saddle point problems. Technical report, Cornell University, 1994.

[11] Susanne C. Brenner. A nonconforming mixed multigrid method for the pure displacement problem in planar linear elasticity. *SIAM J. Numer. Anal.*, 30(1):116–135, 1993.

[12] Susanne C. Brenner. A nonconforming mixed multigrid method for the pure traction problem in planar linear elasticity. *Math. Comp.*, 63(208):435–460, October 1994.

[13] Susanne C. Brenner and L. Ridgway Scott. *The Mathematical Theory of Finite Element Methods.* Springer-Verlag, 1994.

[14] Franco Brezzi. On the existence, uniqueness, and approximation of saddle-point problems arising from Lagrange multipliers. *R.A.I.R.O. Anal. Numer.*, 8:129–151, 1974.

[15] Franco Brezzi and Michel Fortin. *Mixed and Hybrid Finite Element Methods.* Springer-Verlag, 1991.

[16] Tony F. Chan and Tarek P. Mathew. Domain decomposition algorithms. *Acta Numerica*, 1994.

[17] Maksymilian Dryja, Marcus Sarkis, and Olof B. Widlund. Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions. Technical Report 662, Department of Computer Science, Courant Institute, March 1994.

[18] Maksymilian Dryja, Barry F. Smith, and Olof B. Widlund. Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions. *SIAM J. Numer. Anal.*, 31(6), December 1994.

[19] Maksymilian Dryja and Olof B. Widlund. Schwarz methods of Neumann-Neumann type for three-dimensional elliptic finite element problems. Technical Report 626, Department of Computer Science, Courant Institute, March 1993. To appear in Comm. Pure Appl. Math.

[20] Howard Elman. Multigrid and Krylov subspace methods for the discrete Stokes equations. Technical Report UMIA CS-TR-94-76, University of Maryland, June 1994.

[21] Howard Elman and Gene Golub. Inexact and preconditioned Uzawa algorithms for saddle point problems. *SIAM J. Numer. Anal.*, 31(6), December 1994.

[22] Vivette Girault and Pierre-Arnaud Raviart. *Finite Element Methods for Navier-Stokes Equations.* Springer-Verlag, 1986.

[23] Michael Griebel. *Multilevelmethoden als Iterationsverfahren über Erzeugendensystemen.* Teubner Skripten zur Numerik, Teubner-Verlag, Stuttgart, 1994.

[24] Michael Griebel and Peter Oswald. On the abstract theory of additive and multiplicative Schwarz algorithms. *Numer. Math.* To appear.

[25] William D. Gropp and Barry F. Smith. Scalable, extensible, and portable numerical libraries. In *Proceedings of Scalable Parallel Libraries Conference*, pages 87–93. IEEE, 1994.

[26] Wolfgang Hackbusch. *Iterative Lösung großer schwachbesetzter Gleichungssysteme.* Teubner-Verlag, Stuttgart, 1991.

[27] Thomas J.R. Hughes. *The Finite Element Method.* Prentice-Hall, 1987.

[28] Michael Jung. Konvergenzfaktoren von Mehrgitterverfahren für Probleme der ebenen, linearen Elastizitätstheorie. *ZAMM*, 67:165–173, 1987.

[29] David G. Luenberger. The conjugate residual method for constrained minimization problems. *SIAM J. Numer. Anal.*, 7(3):390–398, 1970.

[30] Peter Oswald. *Multilevel Finite Element Approximation. Theory and Applications.* Teubner Skripten zur Numerik, Teubner Verlag, Stuttgart, 1994.

TABLE 6

PCR-method with two-level multigrid with a standard V-cycle defining $\hat{A}$ and $\hat{C} = C$ on a 80×40 grid.

ν	0.3	0.4	0.49	0.499	0.4999	0.49999	0.499999	0.5
Iter	26	29	33	33	33	33	33	33

TABLE 7

PCR-method with an exact solver as $\hat{A}$ and a one-level symmetric multiplicative overlapping Schwarz method with minimal overlap as $\hat{C}$ on a 80×40 grid.

ν	0.3	0.4	0.49	0.499	0.4999	0.49999	0.499999	0.5
Iter	21	21	23	25	25	25	25	25

TABLE 8

PCR-method with two-level multigrid with a standard V-cycle as $\hat{A}$ and a one-level symmetric multiplicative overlapping Schwarz method with minimal overlap as $\hat{C}$ on a 80×40 grid.

ν	0.3	0.4	0.49	0.499	0.4999	0.49999	0.499999	0.5
Iter	26	29	33	33	33	33	33	33

sents an efficient and robust iterative solver for saddle point problems with a penalty term. For a comparison of the convergence rates and the efficiency of multgrid and Krylov subspace methods for saddle point problems; see Elman [20].

Acknowledgements. I would like to thank Olof Widlund for all the time and guidance he has been devoting to me and for all the helpful discussions. I am also indepted to Barry Smith for his excellent advice on the implementation of the algorithms in PETSc.

REFERENCES

[1] Kenneth J. Arrow, Leonid Hurwicz, and Hirofumi Uzawa. *Studies in Linear and Nonlinear Programming*. Stanford University Press, 1958.

[2] Steven F. Ashby, Thomas A. Manteuffel, and Paul E. Saylor. A taxonomy for conjugate gradient methods. *SIAM J.Numer.Anal.*, 27:1542–1568, 1990.

[3] Ivo Babuška and Manil Suri. Locking effects in the finite element approximation of elasticity problems. *Numer. Math.*, 62(4):439–463, 1992.

[4] Ivo Babuška. The finite element method with Lagrange multipliers. *Numer. Math.*, 20:179–192, 1973.

[5] Randolph E. Bank, Bruno D. Welfert, and Harry Yserentant. A class of iterative methods for solving saddle point problems. *Numer. Math.*, 56:645–666, 1990.

[6] Dietrich Braess. A multigrid method for the membrane problem. *Comp. Mech.*, 3:321–329, 1988.

[7] Dietrich Braess. *Finite Elemente*. Springer-Verlag, 1992.

[8] Dietrich Braess and C. Blömer. A multigrid method for a parameter dependent problem in solid mechanics. *Numer. Math.*, 57:747–761, 1990.

TABLE 1

PCR-method with exact solvers as preconditioners for A and C, $\nu = 0.3$.

Grid	20×10	40×20	60×30	80×40	100×50	120×60	140×70
Iter	17	19	19	21	21	21	21

TABLE 2

PCR-method with a two-level multigrid preconditioner with a standard V-cycle defining $\hat{A}$, and $\hat{C} = C$, and $\nu = 0.3$.

Grid	20×10	40×20	60×30	80×40	100×50	120×60	140×70
Iter	20	23	24	26	26	26	26

TABLE 3

PCR-method with a two-level multigrid preconditioner with a standard V-cycle defining $\hat{A}$ and a one-level symmetric multiplicative overlapping Schwarz method with minimal overlap defining $\hat{C}$, and $\nu = 0.3$.

Grid	20×10	40×20	60×30	80×40	100×50	120×60	140×70
Iter	20	23	24	26	26	26	26

TABLE 4

PCR-method with a two-level multigrid preconditioner with a standard V-cycle defining $\hat{A}$, a diagonal preconditioner $\hat{C} = diag(C)$, and $\nu = 0.3$.

Grid	20×10	40×20	60×30	80×40	100×50	120×60	140×70
Iter	46	53	56	58	58	58	60

TABLE 5

PCR-method with exact solvers as preconditioners for A and C on a 80×40 grid.

ν	0.3	0.4	0.49	0.499	0.4999	0.49999	0.499999	0.5
Iter	21	21	23	25	25	25	25	25

has a condition number independent of h, a simple diagonal preconditioning is not the best choice as far as the number of iterations is concerned; cf. Table 4. The cost can be reduced at a little extra expense. Interpreting diagonal preconditioning as a one-level additive non-overlapping Schwarz method with 1×1 nodes per subdomain; see e. g. Hackbusch [26], p. 343, it is natural to try to improve the convergence by introducing some overlap or by using a multiplicative scheme. Our experiments show that already the minimal overlap of one node combined with a one-level symmetric multiplicative Schwarz method yields results matching those obtained by the exact solver for C; see Tables 2 and 3 (resp. Tables 6 and 8). It would of course be more efficient to use more than two levels in the multigrid preconditioner but in this paper we primarily wish to analyze the parameter dependence of the preconditioned methods.

The experiments show that the preconditioned conjugate residual method repre-

Here, we have used that $M := \hat{\mathcal{B}}_\alpha^{-1/2}\mathcal{A}\hat{\mathcal{B}}_\alpha^{-1/2}$ is normal. Analogously, we obtain by using that $\lambda_{min}^{abs}(M) = \rho(M^{-1})$,

$$\lambda_{min}^{abs}(\hat{\mathcal{B}}_\alpha^{-1/2}\mathcal{A}\hat{\mathcal{B}}_\alpha^{-1/2}) \geq \lambda_{min}^{abs}(\mathcal{B}_\alpha^{-1/2}\mathcal{A}\mathcal{B}_\alpha^{-1/2}) \min\{a_0^2, c_0^2\}. \tag{47}$$

Since $\hat{\mathcal{B}}_\alpha^{-1}\mathcal{A}$ and $\hat{\mathcal{B}}_\alpha^{-1/2}\mathcal{A}\hat{\mathcal{B}}_\alpha^{-1/2}$ have the same spectrum, Theorem 4 follows.

□

From Theorem 3 and 4 follows immediately

COROLLARY 1.

$$\kappa(\hat{\mathcal{B}}_\alpha^{-1}\mathcal{A}) \leq \frac{\max\{a_1^2, c_1^2\}}{\min\{a_0^2, c_0^2\}} \frac{\alpha_{max}/2 + \sqrt{m_1^2\beta_1^2 + (\alpha_{max})^2}}{-\alpha_{min}/2 + \sqrt{m_0^2\beta_0^2 + (\alpha_{min}/2)^2}}. \tag{48}$$

Hence, we have now derived an estimate for the condition number which is independent of the discretization and the penalty parameter and for which we can guarantee that the convergence rate of the Krylov space method considered will not deteriorate when t and h decrease. Corollary 1, shows that the condition number estimate of $\hat{\mathcal{B}}_\alpha^{-1}\mathcal{A}$ is completely defined by the quality of the preconditioners $\hat{A}, \hat{C}$, the condition number of $BA^{-1}B^t$ and the choice of the scaling parameter α.

5. Numerical examples. We apply our preconditioner to the problem of planar, linear elasticity; see Section 2. For simplicity, we work with the formulation given in Remark 2. All the results shown are for mixed boundary conditions and the region $[-1,1]\times[-1,1]$. In our experience, the number of iterations for homogeneous Dirichlet boundary conditions are always 4-5 fewer than for the mixed case. As there is no difference between the two cases with regard to the critical parameters, we present just the worst case. We note that our model is mathematically equivalent to the full elasticity problem only in the case of homogeneous Dirichlet conditions. The numerical results confirm that the number of iterations is bounded independently of the critical parameters h and t.

All computations were carried out on a Sun workstation with 256 Mbyte memory using the numerical software package PETSc developed by William Gropp and Barry Smith at the Argonne National Laboratory; see Gropp and Smith [25] or Smith [40]. The initial guess is 0, and the stopping criterion is $\|r_k\|_2/\|r_0\|_2 < 10^{-5}$, where r_k is the k-th residual. The Krylov space method used is always the PCR-method. In all runs we have chosen $\alpha = 1$. We could not detect a difference in the convergence rate for other choices of α, e.g. $\alpha = t$.

To see how the PCR-method behaves under the best of circumstances, we first conducted some experiments using exact solvers, i.e. $\hat{A} = A$ and $\hat{C} = C$. The method works as predicted for both limit cases, $h \to 0$ and $t \to 0$, see Tables 1 and 5.

In another series of experiments, we applied different preconditioners for A and C. We present results with a two-level multigrid preconditioner with a V-cycle including one pre- and one post-smoothing symmetric Gauß-Seidel step defining $\hat{A}$, and a one-level symmetric multiplicative overlapping Schwarz method with minimal overlap as $\hat{C}$; see Tables 2, 3, 6, 7, 8. We note, that although C is uniformly well conditioned, i.e.

By construction, see (26), we have

$$\frac{p^t(BA^{-1}B^t)p}{p^tM_pp} \le \beta_1^2 \tag{39}$$

Using (22) and $C=\hat{C}$, we get

$$\frac{p^t(BA^{-1}B^t)p}{p^t\hat{C}p} = \frac{q^t\hat{C}^{-1/2}BA^{-1}B^t\hat{C}^{-1/2}q}{q^tq} \le \beta_1^2m_1^2. \tag{40}$$

Applying (21), results in

$$\frac{q^t(\hat{C}^{-1/2}B\hat{A}^{-1/2})(\hat{A}^{-1/2}B^t\hat{C}^{-1/2})q}{q^tq} \le \beta_1^2m_1^2a_0^2. \tag{41}$$

The definition of $\tilde{B}$ gives

$$\frac{q^t\tilde{B}\tilde{B}^tq}{q^tq} \le \beta_1^2m_1^2a_0^2 = \beta_1^2m_1^2. \tag{42}$$

Here, $a_0=1$, since we are using $\hat{A}=A$. Analogously, using (21),(22), and (24), we obtain

$$\frac{q^t\tilde{B}\tilde{B}^tq}{q^tq} \ge \beta_0^2m_0^2a_1^2 = \beta_0^2m_0^2. \tag{43}$$

The theorem follows from Lemma 2.

□

REMARK 5. *In many applications we have $C=M_p$. In this case the bound in Theorem 3 simplifies to*

$$\frac{\alpha_{max}/2+\sqrt{\beta_1^2+(\alpha_{max})^2}}{-\alpha_{min}/2+\sqrt{\beta_0^2+(\alpha_{min}/2)^2}}. \tag{44}$$

We next give an upper bound of the condition number when a general preconditioner $\hat{\mathcal{B}}_\alpha^{-1}$ is used.

THEOREM 4.

$$\kappa(\hat{\mathcal{B}}_\alpha^{-1}\mathcal{A}) \le \frac{\max\{a_1^2,c_1^2\}}{\min\{a_0^2,c_0^2\}}\kappa(\mathcal{B}_\alpha^{-1}\mathcal{A}). \tag{45}$$

Proof: We consider

$$\begin{aligned}
\lambda_{max}^{abs}(\hat{\mathcal{B}}_\alpha^{-1/2}\mathcal{A}\hat{\mathcal{B}}_\alpha^{-1/2}) &= \rho(\hat{\mathcal{B}}_\alpha^{-1/2}\mathcal{A}\hat{\mathcal{B}}_\alpha^{-1/2})\\
&= \|\hat{\mathcal{B}}_\alpha^{-1/2}\mathcal{A}\hat{\mathcal{B}}_\alpha^{-1/2}\|_2\\
&= \sup_{x\neq 0}\left(\frac{x^t\hat{\mathcal{B}}_\alpha^{-1/2}\mathcal{A}\hat{\mathcal{B}}_\alpha^{-1/2}x}{x^tx}\right)\\
&\le \sup_{x\neq 0}\left(\frac{x^t\mathcal{A}x}{x^t\mathcal{B}_\alpha x}\right)\sup_{x\neq 0}\left(\frac{x^t\mathcal{B}_\alpha x}{x^t\hat{\mathcal{B}}_\alpha x}\right)\\
&\le \lambda_{max}^{abs}(\mathcal{B}_\alpha^{-1/2}\mathcal{A}\mathcal{B}_\alpha^{-1/2})\max\{a_1^2,c_1^2\}.
\end{aligned} \tag{46}$$

Solving for λ, we obtain

$$\lambda = \frac{1 - t^2/\alpha^2}{2} \pm \sqrt{\frac{\mu}{\alpha^2} + \left(\frac{1 + t^2/\alpha^2}{2}\right)^2}. \tag{32}$$

Since $t^2 \leq \alpha^2$, we get, by comparing the distances of the two values of λ to the origin,

$$\begin{aligned} \lambda^{abs}_{max} &= \tfrac{1-t^2/\alpha^2}{2} + \sqrt{\tfrac{\mu_{max}}{\alpha^2} + \left(\tfrac{1+t^2/\alpha^2}{2}\right)^2}, \\ \lambda^{abs}_{min} &= \left| \tfrac{1-t^2/\alpha^2}{2} - \sqrt{\tfrac{\mu_{min}}{\alpha^2} + \left(\tfrac{1+t^2/\alpha^2}{2}\right)^2} \right|, \\ &= \tfrac{t^2/\alpha^2-1}{2} + \sqrt{\tfrac{\mu_{min}}{\alpha^2} + \left(\tfrac{1+t^2/\alpha^2}{2}\right)^2}. \end{aligned} \tag{33}$$

□

The next lemma provides a condition number estimate for the case of the exact preconditioner.

LEMMA 2.

$$\kappa(\mathcal{B}_\alpha^{-1}\mathcal{A}) \leq \frac{\alpha_{max}/2 + \sqrt{\mu_{max} + (\alpha_{max})^2}}{-\alpha_{min}/2 + \sqrt{\mu_{min} + (\alpha_{min}/2)^2}}. \tag{34}$$

Proof: Since the common factor $\frac{1}{\alpha(t)}$ of λ^{abs}_{max} and λ^{abs}_{min} cancels, it is sufficient to provide an upper bound for

$$\frac{\alpha - t^2/\alpha}{2} + \sqrt{\mu_{max} + \left(\frac{\alpha + t^2/\alpha}{2}\right)^2} \tag{35}$$

and a lower bound for

$$\frac{t^2/\alpha - \alpha}{2} + \sqrt{\mu_{min} + \left(\frac{\alpha + t^2/\alpha}{2}\right)^2}. \tag{36}$$

From the assumptions that $0 \leq t \leq \alpha$ and that $\alpha(t) > 0$ is continuous on $[0, 1]$, we immediately find the upper bound

$$\frac{\alpha_{max}}{2} + \sqrt{\mu_{max} + (\alpha_{max})^2}. \tag{37}$$

Knowing that $\alpha(t) > 0$ and the fact that $\alpha(t)$ has a minimum on $[0, 1]$ gives the lower bound.

□

We are now able to prove

THEOREM 3.

$$\kappa(\mathcal{B}_\alpha^{-1}\mathcal{A}) \leq \frac{\alpha_{max}/2 + \sqrt{m_1^2\beta_1^2 + (\alpha_{max})^2}}{-\alpha_{min}/2 + \sqrt{m_0^2\beta_0^2 + (\alpha_{min}/2)^2}}. \tag{38}$$

Proof: Obviously, we only have to provide an upper bound for μ_{max} and a lower bound for μ_{min}, i.e. upper and lower bounds for the Rayleigh quotient $p^t\tilde{B}\tilde{B}^tp/p^tp$.

Although the (pressure) mass matrix is uniformly well conditioned, i.e. has a condition number independent of h, we can decrease the number of iterations at a small expense by replacing a diagonal preconditioner by a one-level overlapping Schwarz method with minimal overlap on small subdomains; see Section 5.

We will next establish some bounds for $BA^{-1}B^t$. These bounds depend directly on that of the inf-sup condition and on the boundedness of B. We then consider the case of the special preconditioner $\mathcal{B}_\alpha$ and use these results to obtain a condition number estimate for the general case $\hat{\mathcal{B}}_\alpha$.

From the inf-sup condition, we obtain

$$\beta_0^2\, p^t M_p p \le p^t B A^{-1} B^t p \quad \forall p \in M; \tag{24}$$

see Babuška [4], Brezzi [14], Brezzi and Fortin [15], and Sylvester and Wathen [43]. Since B is continuous, we have

$$\exists \beta_1 > 0, \text{ such that } u^t B^t p \le \beta_1^2\, (p^t M_p p)^{1/2} (u^t A u)^{1/2} \quad \forall u \in X, \forall p \in M. \tag{25}$$

Substituting $u = A^{-1}B^t p$ and cancelling a common factor, we get

$$p^t B A^{-1} B^t p \le {\beta_1}^2\, p^t M_p p \quad \forall p \in M, \tag{26}$$

and the following inequality holds with positive constants β_0, β_1

$$\beta_0^2\, p^t M_p p \le p^t B A^{-1} B^t p \le \beta_1^2\, p^t M_p p \quad \forall p \in M. \tag{27}$$

These constants are independent of the penalty parameter t.

We next consider the case of the special preconditioner $\mathcal{B}_\alpha$.

Lemma 1.

$$\begin{array}{rcl} \lambda_{max}^{abs}(\mathcal{B}_\alpha^{-1}\mathcal{A}) & = & \frac{1}{\alpha}\left(\frac{\alpha - t^2/\alpha}{2} + \sqrt{\mu_{max} + \left(\frac{\alpha + t^2/\alpha}{2}\right)^2}\right) \\ \lambda_{min}^{abs}(\mathcal{B}_\alpha^{-1}\mathcal{A}) & = & \frac{1}{\alpha}\left(\frac{t^2/\alpha - \alpha}{2} + \sqrt{\mu_{min} + \left(\frac{\alpha + t^2/\alpha}{2}\right)^2}\right) \end{array} \tag{28}$$

Proof: We consider the preconditioned system $\mathcal{B}_\alpha^{-1/2}\mathcal{A}\mathcal{B}_\alpha^{-1/2}$ and the related eigenvalue problem

$$\begin{pmatrix} I & \frac{1}{\alpha}\tilde{B}^t \\ \frac{1}{\alpha}\tilde{B} & -\frac{t^2}{\alpha^2} I \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \lambda \begin{pmatrix} u \\ p \end{pmatrix} \tag{29}$$

From this equation, it follows that

$$\tilde{B}\tilde{B}^t p = (\lambda - 1)(t^2 + \alpha^2 \lambda) p. \tag{30}$$

Denoting the eigenvalues of $\tilde{B}\tilde{B}^t$ by μ, we get

$$(\lambda - 1)(t^2 + \alpha^2\lambda) = \mu. \tag{31}$$

4. The preconditioner. To construct the preconditioner, we work with the matrix representation of the saddle point problem,

$$\mathcal{A} := \begin{pmatrix} A & B^t \\ B & -t^2 C \end{pmatrix}, \tag{18}$$

and assume that the preconditioner has the form

$$\hat{\mathcal{B}}_\alpha := \begin{pmatrix} \hat{A} & O \\ O & \alpha^2 \hat{C} \end{pmatrix}. \tag{19}$$

Here $\hat{A}$ is X-elliptic, $\hat{C}$ M-elliptic and $\alpha > 0$ a parameter at our disposal. Thus, $\hat{\mathcal{B}}_\alpha$ is positive definite.

We will require that $\alpha = \alpha(t)$ is a continuous function on $[0,1]$ and that $0 \le t \le \alpha$. We denote the case of $\hat{A} = A$ and $\hat{C} = C$ by $\mathcal{B}_\alpha$. In view of Theorem 2, it is our goal to give an estimate of the condition number $\kappa(\hat{\mathcal{B}}_\alpha^{-1}\mathcal{A})$. A simple computation shows that

$$\begin{aligned} \hat{\mathcal{B}}_\alpha^{-1/2}\mathcal{A}\hat{\mathcal{B}}_\alpha^{-1/2} &= \begin{pmatrix} \hat{A}^{-1/2}A\hat{A}^{-1/2} & \frac{1}{\alpha}\hat{A}^{-1/2}B^t\hat{C}^{-1/2} \\ \hat{C}^{-1/2}B\hat{A}^{-1/2} & -\frac{t^2}{\alpha^2}\hat{C}^{-1/2}C\hat{C}^{-1/2} \end{pmatrix} \\ &=: \begin{pmatrix} \tilde{A} & \frac{1}{\alpha}\tilde{B}^t \\ \frac{1}{\alpha}\tilde{B} & -\frac{t^2}{\alpha^2}\tilde{C} \end{pmatrix}. \end{aligned} \tag{20}$$

Finally let μ denote an eigenvalue of $\tilde{B}\tilde{B}^t$ and α_{max} and α_{min} the maximum and minimum value of $\alpha(t)$ on $[0,1]$.

We make the following assumptions on $\mathcal{A}$ and $\hat{\mathcal{B}}_\alpha$:
The matrix $\hat{A}$ is a good preconditioner for A, i. e.

$$\exists a_0, a_1 > 0 \quad a_0^2 u^t \hat{A} u \le u^t A u \le a_1^2 u^t \hat{A} u \quad \forall u \in X. \tag{21}$$

The constants a_0, a_1 should preferably be independent of the discretization parameters but there are also interesting cases with a polylogarithmic dependence on H/h; see [17,34,35]. (The parameter H represents the diameter of a subdomain in a domain decomposition method.) Multigrid and domain decomposition methods are examples of preconditioners that meet these requirements. They have also been successfully implemented on parallel machines; see e.g. for details [16,17,18,19,24,23,30,36,39,44,45], and the references therein.

We also require that $\hat{C}$ is a good preconditioner for the pressure mass matrix M_p, i.e.

$$\exists m_0, m_1 > 0 \quad m_0^2 p^t \hat{C} p \le p^t M_p p \le m_1^2 p^t \hat{C} p \quad \forall p \in M \tag{22}$$

and we finally assume that C is spectrally equivalent to $\hat{C}$, i.e.

$$\exists c_0, c_1 > 0 \quad c_0^2 p^t \hat{C} p \le p^t C p \le c_1^2 p^t \hat{C} p \quad \forall p \in M. \tag{23}$$

A good choice for $\hat{C}$ is a one-level overlapping Schwarz method. This family of methods also includes the popular algorithms that use diagonal and block preconditioning.

The PCR-method is an algorithm to solve $\mathcal{A}x = b$ with a symmetric indefinite matrix $\mathcal{A}$ and a positive definite preconditioner $\hat{\mathcal{B}}$. We will give a stable version that is based on a three term recurrence; see Hackbusch [26], p. 270.

ALGORITHM 1.

$$
\begin{aligned}
&\textit{Initialization}: && \\
&& r_0 &:= b - \mathcal{A}x_0, \\
&& p_{-1} &:= 0, \\
&& p_0 &:= \hat{\mathcal{A}}^{-1} r_0, \\
&\textit{Iteration}: && \\
&& \lambda &:= \frac{r_m^t \hat{\mathcal{A}}^{-1} \mathcal{A} p_m}{p_m^t \mathcal{A} \hat{\mathcal{B}}^{-1} \mathcal{A} p_m} \\
&& x_{m+1} &:= x_m + \lambda p_m, \\
&& r_{m+1} &:= r_m - \lambda \mathcal{A} p_m, \\
&& \alpha_0 &:= \frac{p_m^t \mathcal{A} \hat{\mathcal{B}}^{-1} \mathcal{A} \hat{\mathcal{B}}^{-1} \mathcal{A} p_m}{p_m^t \mathcal{A} \hat{\mathcal{B}}^{-1} \mathcal{A} p_m}, \\
&& \alpha_1 &:= \frac{p_m^t \mathcal{A} \hat{\mathcal{B}}^{-1} \mathcal{A} \hat{\mathcal{B}}^{-1} \mathcal{A} p_{m-1}}{p_{m-1}^t \mathcal{A} \hat{\mathcal{B}}^{-1} \mathcal{A} p_{m-1}}, \\
&& p_{m+1} &:= \hat{\mathcal{B}}^{-1} \mathcal{A} p_m - \alpha_0 p_m - \alpha_1 p_{m-1}.
\end{aligned} \tag{15}
$$

REMARK 3. *This algorithm needs two matrix/vector products with the original matrix and one with the preconditioner, but it is easy to derive a version with just one matrix/vector product with each of* $\mathcal{A}$ *and* $\hat{\mathcal{B}}^{-1}$ *by introducing an additional recursion for* $a^m := \mathcal{A}p^m$. *We have used this less expensive version in our numerical experiments.*

REMARK 4. *Attention should be paid to scaling: If we assume that* $\hat{\mathcal{A}}$ *is an optimal positive definite preconditioner but choose* $\hat{\mathcal{B}} := c\hat{\mathcal{A}}$, $c \in \mathbf{R}^+$, *then* λ *will grow in proportion to* $\frac{1}{c^{m+1}}$. *This can easily be seen by induction. This phenomenon seems to be well known but not discussed in the literature. The easiest way of fixing this problem is to normalize* p_{m+1} *in every iteration; we have done so in our implementation.*

We introduce the following notation

$$
\kappa(\hat{\mathcal{B}}^{-1}\mathcal{A}) := \frac{\lambda_{max}^{abs}}{\lambda_{min}^{abs}} := \frac{\max\{|\lambda| : \lambda \in \sigma(\hat{\mathcal{B}}^{-1}\mathcal{A})\}}{\min\{|\lambda| : \lambda \in \sigma(\hat{\mathcal{B}}^{-1}\mathcal{A})\}}, \tag{16}
$$

where $\sigma(\hat{\mathcal{B}}^{-1}\mathcal{A})$ denotes the spectrum of $\hat{\mathcal{B}}^{-1}\mathcal{A}$. The next theorem can be found in Hackbusch [26], p. 270. It gives an upper bound for the convergence rate of the PCR-method.

THEOREM 2. *Let the regular matrix* $\mathcal{A}$ *be symmetric and* $\hat{\mathcal{B}}$ *be positive definite. Then the m-th iterate* x_m *of Algorithm 1 satisfies*

$$
||\hat{\mathcal{B}}^{-1/2}\mathcal{A}(x^m - x^*)||_2 \le \frac{2c^\mu}{1 + c^{2\mu}} ||\hat{\mathcal{B}}^{-1/2}(\mathcal{A}x^0 - b)||_2 \tag{17}
$$

where $c := \frac{\kappa - 1}{\kappa + 1}$, $\kappa := \kappa(\hat{\mathcal{B}}^{-1}\mathcal{A})$ *and* $\frac{m}{2} - 1 < \mu \le \frac{m}{2}$ $\forall \mu \in \mathbf{Z}$.

The convergence rate of the PCR-method is thus determined by the condition number of the preconditioned system.

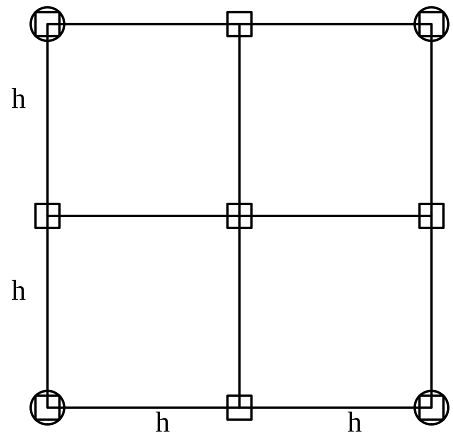

FIG. 1. *The Taylor-Hood Element:* $\square$ *denotes the points of the displacement discretization,* $\circ$ *the ones of the Lagrangian parameter.*

REMARK 2. *If we have homogeneous Dirichlet conditions on the whole boundary it is possible to use an equivalent simpler formulation; see Brezzi and Fortin [15], p.201,*

$$\begin{array}{lllll} \mu(\nabla u, \nabla v)_0 & + & (\mathrm{div} v, p)_0 & = & < f, v > & \forall v \in X := H^1_\Gamma(\Omega), \\ (\mathrm{div} u, q)_0 & - & \frac{1}{\lambda+\mu}(p, q)_0 & = & 0 & \forall q \in M := L_2(\Omega). \end{array} \tag{13}$$

We know from Korn's inequality that the bilinear form $(\varepsilon(u), \varepsilon(v))_0$ is X- elliptic. Thus (12) is very similar to the Stokes problem and we can use finite elements developed for that problem. In this paper, we discretize by the Taylor-Hood element; see Fig. 1. For the displacements u, we use piecewise linear polynomials on quadrilaterals on a grid with the meshsize h and for the Lagrange parameter p piecewise linear polynomials on quadrilaterals with mesh size $2h$. The corresponding finite element spaces are

$$\begin{array}{lll} X^h & := & \{v_h \in (\mathcal{C}(\Omega))^d \cap H^1_\Gamma(\Omega) : v_{h|T} \in \mathcal{P}_1,\ T \in \tau_h\}, \\ M^h & := & \{q_h \in \mathcal{C}(\Omega) \cap L_2(\Omega) : q_{h|T} \in \mathcal{P}_1,\ T \in \tau_{2h}\}. \end{array} \tag{14}$$

For a proof that the inf-sup condition holds in this case; see Verfürth [42], Girault and Raviart [22], or Brezzi and Fortin [15]. All the conditions for Theorem 1 are satisfied and the finite element method converges independently of the Lamé constants (resp. the Poisson ratio). It is of course no restriction, to assure that $\frac{1}{\lambda} =: t^2$ is in $[0, 1]$ since we are mainly interested in the nearly incompressible case.

3. Iterative methods for indefinite problems. The preconditioned conjugate gradient method (PCG) has gained great popularity for positive definite problems. A natural generalization for symmetric indefinite problems is the preconditioned conjugate residual method (PCR); see Ashby et al. [2], Hackbusch[26], Luenberger [29], and Paige and Saunders [31]. There are also other methods that could be used for indefinite problems, e.g. Bi-CGSTAB, CGS, BiCGstab(l); see Sleijpen, Van der Vorst, and Fokkema [37]. We only describe the PCR-method and give a convergence estimate that is determined by the condition number of the preconditioned linear system.

REMARK 1. *The conditions (i) and (ii) are known as the Babuška-Brezzi condition for the saddle point problem (3).*

This result is also valid for finite element spaces. We then require, additionally, that the constants ϑ, β_0 are independent of h. The continuity assumptions turn into uniform boundedness with respect to h; see Brezzi and Fortin [15], Braess [7].

2.2. The equations of linear elasticity. An example of a saddle point problem with a penalty term arises from the displacement formulation of the equations of linear elasticity. In the rest of this paper u, v (resp. p, q) will always denote vector valued (resp. scalar) functions. We define

$$
\begin{array}{ll}
\nabla p := \left(\frac{\partial p}{\partial x_i}\right)_{i\in\{1,\ldots,d\}}, & \operatorname{div} v := \sum_{i=1}^{d} \frac{\partial v_i}{\partial x_i}, \\
\nabla v := (\nabla v_i)_{i\in\{1,\ldots,d\}}, & \varepsilon(v) := \frac{1}{2}\left(\nabla v + (\nabla v)^t\right), \\
(\varepsilon(v),\varepsilon(u))_0 := \int_\Omega \varepsilon(v):\varepsilon(u)dx, & (v,u)_0 := \int_\Omega vu dx, \\
< f, v > := \int_\Omega f v dx + \int_{\Gamma_1} g_1 v dx, & \Gamma_1 := \partial\Omega \setminus \Gamma_0,
\end{array} \tag{7}
$$

with the inner product for matrices defined by

$$
\sigma : \tau \ := \ \sum_{i=1}^{d}\sum_{j=1}^{d} \sigma_{ij}\tau_{ij} \tag{8}
$$

and Γ_0 denoting the part of the boundary where a homogeneous Dirichlet condition is imposed. The following Sobolev spaces are used

$$
H^1_\Gamma(\Omega) := \left(H^1_{\Gamma_0}(\Omega)\right)^d, d = 2,3. \tag{9}
$$

The variational formulation for the equations of elasticity is then

$$
2\,\mu\,(\varepsilon(u),\varepsilon(v))_0 + \lambda\,(\operatorname{div} u, \operatorname{div} v)_0 = < f, v > \quad \forall v \in X := H^1_\Gamma(\Omega). \tag{10}
$$

Instead of using the Lamé constants λ and μ, we can also work with Young's elasticity modulus E and the Poisson ratio ν. These constants are related to each other by the following equations

$$
\begin{array}{lllllll}
\lambda & = & \frac{E\nu}{(1+\nu)(1-2\nu)} & , & \nu & = & \frac{\lambda}{2(\lambda+\mu)}, \\
\mu & = & \frac{E}{2(1+\nu)} & , & E & = & \frac{\mu(3\lambda+2\mu)}{\lambda+\mu}.
\end{array} \tag{11}
$$

Some materials, e.g. rubber, are nearly incompressible, i.e. small changes in the density of the material lead to a rapid growth of the energy. Almost incompressible materials are characterized by a Poisson ratio near $\frac{1}{2}$. In terms of the Lamé constants, this means that λ tends to infinity. This leads to the phenomenon of locking if a low order finite element model is used in a pure displacement setting; see Section 1. To obtain a finite element model that converges independently of locking as $h \to 0$, we reformulate the pure displacement model as a saddle point problem with a penalty term; see Brezzi and Fortin [15], and Braess [7]. We introduce a new variable $p := \lambda \operatorname{div} u$ and obtain from (10)

$$
\begin{array}{lllll}
2\,\mu\,(\varepsilon(u),\varepsilon(v))_0 & + & (\operatorname{div} v, p)_0 & = & < f, v > & \forall v \in X := H^1_\Gamma(\Omega), \\
(\operatorname{div} u, q)_0 & - & \frac{1}{\lambda}(p,q)_0 & = & 0 & \forall q \in M := L_2(\Omega).
\end{array} \tag{12}
$$

2. Saddlepoint problems with a penalty term. In this section, we give different formulations of the saddle point problem. We first define it by bilinear forms in a Hilbert space setting and then give an equivalent matrix/operator representation. Finally we describe a mixed formulation of the displacement method of linear elasticity. The abstract theory shows that this model does not suffer from locking.

2.1. Abstract theory. Let X and M be two Hilbert spaces and let

$$a(\cdot,\cdot): X \times X \to \mathbf{R},\ b(\cdot,\cdot): X \times M \to \mathbf{R},\ c(\cdot,\cdot): M \times M \to \mathbf{R}, \tag{2}$$

be three continuous bilinear forms. We assume that $a(\cdot,\cdot)$ is X- elliptic and that $c(\cdot,\cdot)$ is M-positive definite. Furthermore let $f \in X'$, $g \in M'$, where X', M' are the dual spaces of X, M, and let $< \cdot, \cdot >$ denote the dual pairing for X, X' (resp. M, M'). Consider the following problem:
Find $(u,p) \in X \times M$ with

$$\begin{array}{lcrcll} a(u,v) & + & b(v,p) & = & < f, v > & \forall v \in X \\ b(u,q) & - & t^2 c(p,q) & = & < g, q > & \forall q \in M \quad t \in [0,1]. \end{array} \tag{3}$$

It is often useful to reformulate the problem in operator form. For a specific discretization, the operators will be represented by matrices. It is well known that there exist linear operators

$$\begin{array}{rcccll} A: & X & \to & X' & < Au, v >= a(u,v) & \forall v \in X, \\ B: & X & \to & M' & < Bu, q >= b(u,q) & \forall q \in M, \\ B': & M & \to & X' & < B'p, v >= b(v,p) & \forall v \in X, \\ C: & M & \to & M' & < Mp, q >= c(p,q) & \forall q \in M. \end{array} \tag{4}$$

Thus (3) is equivalent to

$$\begin{array}{lcrcl} Au & + & B'p & = & f \\ Bu & - & t^2 Cp & = & g. \end{array} \tag{5}$$

To obtain a well-posed problem, the bilinear form $b(\cdot,\cdot)$ must satisfy the inf-sup condition; see Babuška [4], Brezzi [14], and Brezzi and Fortin [15]. The following theorem can be found in Braess [7], p.129.

THEOREM 1. *The saddle point problem (3) defines an isomorphism $L : X \times M \to X' \times M'$ if the following conditions are fulfilled:*

$$\begin{array}{l} (i)\ \textit{The continuous bilinear form } a(\cdot,\cdot) \textit{ is } X\textit{-elliptic, i.e.} \\ \quad \exists \vartheta > 0,\ \textit{such that}\ a(v,v) \geq \vartheta \|v\|_X^2\ \forall v \in X. \\ (ii)\ \textit{The continuous bilinear form } b(\cdot,\cdot) \textit{ fulfills the inf-sup condition, i.e.} \\ \quad \exists \beta_0 > 0,\ \textit{such that}\ \inf_{q \in M} \sup_{v \in X} \frac{b(v,q)}{\|v\|_X \|q\|_M} \geq {\beta_0}^2 \\ (iii)\ \textit{The continuous bilinear form } c(\cdot,\cdot) \textit{ is } M\textit{- positive semi-definite, i.e.} \\ \quad c(q,q) \geq 0\ \forall q \in M. \end{array} \tag{6}$$

Under these assumptions, the operator L^{-1} is uniformly bounded for $t \in [0,1]$.

is better suited for multigrid methods than reduced/selected integration; the convergence rate of the iterative method for the linear system arising from the latter model still deteriorates.

In this paper we focus on the construction of an iterative method for certain saddle point problems with a penalty term. We describe a preconditioned iterative solver for

$$\begin{pmatrix} A & B^t \\ B & -t^2 C \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} \tag{1}$$

that converges independently of the penalty and discretization parameters.

Several techniques have been developed to solve such problems iteratively. As far as we know, the parameter dependence of the convergence rate never seems to have been considered systematically in these studies. The oldest algorithm is known as the Uzawa algorithm; see Arrow, Hurwicz, and Uzawa [1]. It is mainly a gradient algorithm applied to the Schur complement $t^2C + BA^{-1}B^t$ of the indefinite linear system. When this algorithm is used, we have to solve a linear system of the form $A\delta = d$. This can be quite expensive since A normally is not well conditioned. Therefore many authors have introduced an inner iteration for A^{-1}, see Bank, Welfert, and Yserentant [5], Bramble, Pasciak, and Vassilev [10], Elman and Golub [21], and Rønquist [32]. To avoid inner and outer iterations and to provide a much simpler approach, some authors have in recent years tried to precondition the whole indefinite system and to use a conjugate residual algorithm as an accelerator; see e.g. Rusten and Winther [33], and Sylvester and Wathen [43,41]. In the first of these papers, a preconditioner for a saddle point problem without a penalty term is analyzed; in the others a problem arising from stabilized and unstabilized Stokes flow is considered. Only the stabilized case results in a saddle point problem with a penalty term. In contrast to problems arising in elasticity, where the penalty parameter arises from the material or the geometry, the penalty term in the stabilized Stokes case can and should be chosen to stabilize an otherwise unstable discretization and to ensure fast convergence of the iterative method. The main goal in Sylvester and Wathen [43,41] is to provide a good criterion for choosing this parameter in the context of preconditioning. A third possibility is to transform the indefinite problem into a positive definite one by introducing a new inner product. Then the conjugate gradient method can be applied; see Bramble and Pasciak [9].

In this paper, we show that the condition number of the preconditioned system is bounded independently of the finite element discretization and the penalty parameter. Therefore the Krylov space method used to solve the preconditioned system converges at a rate which is independent of the critical parameters; see Section 3.

The outline of the remainder of the paper is as follows. In Section 2, we describe the abstract saddle point problem with a penalty term and the finite element theory thereof. An example for the equations of linear elasticity is given. In Section 3, we consider the preconditioned conjugate residual method (PCR) as an example of a Krylov space method for indefinite linear systems. In Section 4, we analyze the proposed preconditioner and give our condition number estimate. In Section 5, we discuss numerical results for the problem of linear elasticity.

AN OPTIMAL PRECONDITIONER FOR A CLASS OF SADDLE POINT PROBLEMS WITH A PENALTY TERM

AXEL KLAWONN *

Abstract. Iterative methods are considered for a class of saddle point problems with a penalty term arising from finite element discretizations of certain elliptic problems. An optimal preconditioner which is independent of the discretization and the penalty parameter is constructed. This approach is then used to design an iterative method with a convergence rate independent of the Lamé parameters occuring in the equations of linear elasticity.

Key words. mixed finite elements, saddle point problems, penalty term, nearly incompressible materials, elasticity, preconditioned conjugate residual method, domain decomposition, multigrid

AMS(MOS) subject classifications. 65F10, 65N22, 65N30, 65N55, 73V05

1. Introduction. In recent years, modern iterative methods, e.g. domain decomposition and multigrid methods, have been applied to parameter dependent problems arising in solid mechanics; see Braess [6], Braess and Blömer [8], Jung [28], and Smith [38]. If the direct approach of (low order) conforming finite elements is used in a pure displacement setting, the phenomenon of locking leads to problems. Locking occurs when a parameter, e.g. the Poisson ratio of a material, approaches a limit. The convergence rate of the iterative method and that of the finite element model deteriorates severely when the limit is approached, e.g. when the Poisson ratio tends to 1/2 in the problem of linear elasticity; see Braess [6] and Jung [28]; note that one has to make a distinction between the convergence rate of the finite element model and the convergence rate of the iterative method. This deterioration of the convergence rates can be explained by interpreting locking as a problem of ill conditioning; see Braess [7], pp. 253-254. For a detailed discussion of the locking phenomenon in the finite element model; see Babuška and Suri [3].

There are different approaches to overcome the problem of locking in the finite element model; nonconforming finite element methods, reduced/selected integration and a reformulation in terms of a saddle point problem with a penalty term. Most of them can be analyzed as saddle point problems with a penalty term; see Braess [7], Brenner [11,12], Brenner and Scott [13], Brezzi and Fortin [15], and Hughes [27]. For all of these approaches it can be proven, that the finite element solution converges uniformly with regard to the penalty parameter but there is still a difference between these methods as far as the iterative solution of the resulting linear systems is concerned. Thus it was observed in Braess and Blömer [8] that the mixed formulation

* Westfälische Wilhelms-Universität, Institut für Numerische und instrumentelle Mathematik, Einsteinstraße 62, 48149 Münster, Germany. Electronic mail address: klawonn@math.uni-muenster.de.
This work was supported by a scholarship of the Hochschulsonderprogramm II/AUFE des Deutschen Akademischen Austauschdienstes (DAAD) and carried out while visiting Prof. Olof Widlund at the Courant Institute in New York during the academic year 1993/94 and Dr. Barry Smith at Argonne National Laboratory, Argonne, IL. Dieser Artikel ist Teil einer geplanten Dissertation an der Mathematisch-Naturwissenschaftlichen Fakultät der Westfälischen Wilhelms-Universität.