

## A Multigrid Version of a Simple Finite Element Method for the Stokes Problem

By Juhani Pitkäranta and Tuomo Saarinen

**Abstract.** We consider a finite element method for the Stokes problem on a rectangular domain based on piecewise bilinear velocities and piecewise constant pressures on a uniform rectangular grid. It is shown that by a simple stabilization strategy the method can be implemented in a convergent multigrid procedure.

**Introduction.** One of the most natural ways of discretizing the Stokes equations on a two-dimensional rectangular domain is to apply the finite element techniques with continuous piecewise bilinear velocities and piecewise constant pressures on a rectangular grid. This method is easily extended to more general quadrilateral meshes and it has proved to be quite effective in practice. The convergence analysis of the method was first carried out in [10] on a rectangular domain. In [12] the estimates of [10] are improved and extended to more general quadrilateral meshes.

So far, the above method has been used mainly in connection with direct band solvers, usually combined with penalty/perturbation techniques to eliminate the pressure [9]. In this paper we consider an iterative variant of the method based on multigrid techniques. Such an algorithm seems attractive, especially since the underlying finite-difference equations are relatively simple.

Multigrid methods for the Stokes problem and for more general elliptic systems were considered previously by Hackbusch [8] and Verfürth [14]. In [14] it is shown that a convergent multigrid algorithm for the Stokes problem can be constructed by appropriately scaling the variables, provided that the underlying finite element scheme satisfies the Babuška-Brezzi stability condition [2], [5]. In our case, however, the finite element method is not stable in this sense, and we have not been able to show that the straightforward application of the algorithm proposed in [14] yields a convergent process. Therefore, we suggest first modifying the finite element method in such a way that it becomes stable in the ordinary sense.

The usual way of stabilizing an unstable mixed method is to add more velocity (or primary) variables. A classical example of this is the quadratic/linear velocity-pressure element of Crouzeix and Raviart [6], where added “bubble” functions act as stabilizers in the velocity space. Increasing the dimension of the velocity space would be an alternative in our case. However, we propose another method which seems to yield a simpler algorithm. It is based on adding an extra stabilizing term into the

---

Received May 1, 1984.

1980 *Mathematics Subject Classification.* Primary 65H30.

©1985 American Mathematical Society  
0025-5718/85 \$1.00 + \$.25 per page

finite element functional while keeping the velocity and pressure spaces unchanged. The added term introduces an extra consistency error, the magnitude of which depends on the smoothness of the pressure. At worst, it is of the same order of magnitude as the discretization error in the original scheme. The main point in the modification is, however, that it increases the computational work of the multigrid algorithm by only a negligible amount.

For simplicity, we consider only a model problem on a two-dimensional rectangular domain. The ideas could be easily extended to more general variants of the method. In particular, the same kind of stabilization is possible on the quadrilateral meshes considered in [12] or in the three-dimensional version of the method analyzed in [11].

The plan of the paper is as follows. In Section 2 we introduce the finite element method and establish some basic stability and error estimates. In Section 3 we consider an iterative two-grid algorithm based on the finite element method of Section 2. The analysis of a more general multilevel algorithm is omitted, as it is a straightforward extension of the two-grid case, cf. [4], [13], [14]. In Section 4 we report on some numerical experiments on a three-level algorithm.

**2. The Finite Element Method.** Let  $\Omega \in \mathbf{R}^2$  be a rectangular domain,  $\Omega = \{(x_1, x_2): 0 < x_i < a_i, i = 1, 2\}$ . We consider the Stokes differential equations for an incompressible fluid with viscosity equal to unity:

$$(2.1) \quad \begin{aligned} -\Delta u + \nabla p &= f && \text{in } \Omega, \\ \operatorname{div} u &= 0 && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega, \end{aligned}$$

$$\int_{\Omega} p \, dx = 0.$$

Here  $u = (u_1, u_2)$  is the velocity field,  $p$  denotes the pressure and  $f \in [L_2(\Omega)]^2$  is the external force field. For simplicity, we assume homogeneous Dirichlet boundary conditions and normalize the pressure to have a zero mean value.

Finite element methods for solving the system (2.1) are usually formulated as: Given the subspaces  $V_h \subset [H_0^1(\Omega)]^2 \equiv V$  and  $Q_h \subset L_2(\Omega) \equiv Q$ , find  $(u_h, p_h) \in V_h \times Q_h$  such that  $\int_{\Omega} p_h \, dx = 0$  and

$$(2.2a) \quad (\nabla u_h, \nabla v) - (p_h, \operatorname{div} v) = (f, v), \quad v \in V_h,$$

$$(2.2b) \quad (\operatorname{div} u_h, q) = 0, \quad q \in Q_h,$$

where  $(\cdot, \cdot)$  denotes the inner product of  $L_2(\Omega)$  or  $[L_2(\Omega)]^2$ . Here we consider a particular case of algorithm (2.2) where the subspaces are chosen as follows: Let  $\mathcal{C}_h$  be a uniform rectangular partitioning of  $\Omega$ ,

$$(2.3) \quad \mathcal{C}_h = \left\{ K_{ij} = \{(x_1, x_2): ih_1 < x_1 < (i+1)h_1, jh_2 < x_2 < (j+1)h_2\}, \right. \\ \left. i = 0, \dots, m_1 - 1, j = 0, \dots, m_2 - 1 \right\},$$

where  $h_i = a_i/m_i$ , and we set  $h \equiv h_1$ . Then, let  $V_h$  consist of continuous vector functions which vanish on  $\partial\Omega$  and are bilinear on each  $K \in \mathcal{C}_h$ , and let  $Q_h$  consist of functions which are constant on each  $K \in \mathcal{C}_h$ .

When used in connection with direct solvers, the above algorithm performs well, as is known from both practice [9] and theory [10], [12], [15]. The multigrid implementation of the method, however, seems nontrivial since the method is not stable in the classical sense of Babuška [2] and Brezzi [5] (see Lemma 2.1 below). We therefore suggest modifying the scheme so that instead of (2.2) we seek  $(u_h, p_h) \in V_h \times Q_h$  which satisfies (2.2a) and

$$(2.2b') \quad \kappa(p_h, \pi_h q) + (\operatorname{div} u_h, q) = 0, \quad q \in Q_h,$$

where  $\kappa \geq 0$  is a parameter and  $\pi_h$  is an  $L_2$ -projection into a subspace of  $Q_h$  as defined below. The scheme (2.2a, b') performs well whenever  $\kappa$  is nonnegative and bounded (see the remark following Theorem 1 below), but it has the desirable multigrid characteristics only when  $\kappa = \mathcal{O}(1)$ . Therefore, we mainly assume that  $\kappa = 1$  below.

Let us now define the projection  $\pi_h$  in (2.2b'). To this end, assume that  $m_1$  and  $m_2$  in (2.3) are even numbers, and denote by  $\mathcal{C}_{2h}$  a coarser rectangular subdivision of  $\Omega$  defined by

$$\mathcal{C}_{2h} = \left\{ T_{\nu\mu}, \nu = 0, \dots, \frac{m_1}{2} - 1, \mu = 0, \dots, \frac{m_2}{2} - 1 \right\},$$

where

$$T_{\nu\mu} = \{(x_1, x_2) : 2\nu h_1 < x_1 < 2(\nu + 1)h_1, 2\mu h_2 < x_2 < 2(\mu + 1)h_2\}.$$

For each  $T_{\nu\mu} \in \mathcal{C}_{2h}$ , define a function  $\xi_{\nu\mu} \in Q_h$  as follows:

$$(2.4) \quad \xi_{\nu\mu}(x) = \begin{cases} 0 & \text{if } x \notin T_{\nu\mu}, \\ (-1)^{i+j} & \text{if } x \in K_{ij} \cap T_{\nu\mu}, K_{ij} \in \mathcal{C}_h. \end{cases}$$

Then let  $\pi_h$  be the  $L_2$ -projection into the subspace of  $Q_h$  spanned by the functions  $\xi_{\nu\mu}$ .

In the remaining part of this section we analyze the convergence of the modified finite element method (2.2a, b'). We denote below by  $\|\cdot\|_s$  and  $|\cdot|_s$ , respectively, the norm and the seminorm of the Sobolev space  $H^s(\Omega)$  or  $[H^s(\Omega)]^2$ ,  $s \geq 0$ . By  $C$ , we denote a generic constant which may take different values on different usages. The constant may depend on  $\Omega$  and on the ratio  $h_1/h_2$ , but not on any other parameters unless indicated explicitly. By  $P_k(K)$ , we denote the space of polynomials on  $K$  of degree  $\leq k$ .

To shorten the notation, let us introduce the bilinear form

$$(2.5) \quad \mathcal{B}((u, p); (v, q)) = (\nabla u, \nabla v) - (p, \operatorname{div} v) - (\operatorname{div} u, q)$$

defined on  $V \times Q$ . Then the variational formulation of problem (2.1) is: Find  $(u, p) \in V \times Q$  such that  $(p, 1) = 0$  and

$$(2.6) \quad \mathcal{B}((u, p); (v, q)) = g(v, q), \quad (v, q) \in V \times Q,$$

where  $g(v, q) = (f, v)$ . Similarly, (2.2a, b') may be written as

$$(2.7) \quad \mathcal{B}_h((u_h, p_h); (v, q)) = g(v, q), \quad (v, q) \in V_h \times Q_h,$$

where

$$(2.8) \quad \mathcal{B}_h((u, p); (v, q)) = \mathcal{B}((u, p); (v, q)) - \kappa(p, \pi_h q).$$

The finite element method (2.7) is stable if there is a constant  $C$  such that for all  $g$  in the dual space of  $V \times Q$ , the solution to (2.7) satisfies

$$(2.9) \quad \|u_h\|_1 + \|p_h\|_0 \leq C \|g\|_{-1,0},$$

where  $\|\cdot\|_{-1,0}$  denotes the dual norm of  $V \times Q$ :

$$\|g\|_{-1,0} = \sup_{(v,q) \in V \times Q} \frac{g(v,q)}{\|v\|_1 + \|q\|_0}.$$

Inequality (2.9) is equivalent to the following condition due to Babuška [2]: There is a constant  $C > 0$  such that for all  $(v, q) \in V_h \times Q_h$  with  $(q, 1) = 0$ :

$$(2.10) \quad \sup_{(w,r) \in V_h \times Q_h} \frac{\mathcal{B}_h((v,q);(w,r))}{\|w\|_1 + \|r\|_0} \geq C(\|v\|_1 + \|q\|_0).$$

On the other hand, (2.10) follows if one can prove for all  $q \in Q_h$  with  $(q, 1) = 0$ , the inequality (cf. [5], [6], [7])

$$(2.11) \quad \sup_{v \in V_h} \frac{(\operatorname{div} v, q)}{\|v\|_1} \geq C \|q\|_0,$$

where  $C > 0$  is independent of  $q$ . This is usually referred to as the Babuška-Brezzi stability inequality. In the present case (2.11) does not hold (see below), but we can nevertheless prove (2.10) for  $\kappa = 1$ .

Below we denote by  $\tilde{u}$  the usual interpolant of  $u$  in  $V_h$  and by  $\tilde{p}$  the  $L_2$ -projection of  $p$  into  $Q_h$ . We also need another interpolant  $\tilde{\tilde{p}}$  defined so that for each  $T \in \mathcal{C}_{2h}$ ,  $\tilde{\tilde{p}}|_T \in P_1(T)$  and

$$\int_T (p - \tilde{\tilde{p}}) \rho \, dx = 0, \quad \rho \in P_1(T).$$

**THEOREM 1.** *Let  $(u, p) \in V \times Q$  be a solution to (2.6), where  $g$  is a bounded linear functional on  $V \times Q$ , and let  $(u_h, p_h) \in V_h \times Q_h$  satisfy (2.7), where  $\mathcal{B}_h$  is defined by (2.8) with  $\kappa = 1$ . Further, assume that  $(p, 1) = (p_h, 1) = 0$ . Then we have the following stability and error estimates:*

$$\|u_h\|_1 + \|p_h\|_0 \leq C \|g\|_{-1,0},$$

and

$$h^{-1} \|u - u_h\|_0 + |u - u_h|_1 + \|p - p_h\|_0 \leq C (|u - \tilde{u}|_1 + \|p - \tilde{p}\|_0 + \|p - \tilde{\tilde{p}}\|_0),$$

where the interpolants are defined as above.

Before proving this theorem, let us note that we have the simple

**COROLLARY.** *Under the assumptions of Theorem 1, if the solution of (2.6) satisfies  $u \in [H^2(\Omega)]^2$ ,  $p \in H^1(\Omega)$ , then one has the error estimates*

$$h^{-1} \|u - u_h\|_0 + |u - u_h|_1 + \|p - p_h\|_0 \leq Ch (|u|_2 + |p|_1).$$

*Proof of Theorem 1.* Let us first prove the stability estimate in the form of (2.10). We need the following weak form of the Babuška-Brezzi stability estimate. For the proof, see [10]. Below,  $I$  denotes the identity operator.

**LEMMA 2.1.** *There is a positive constant  $C$  such that for all  $q \in Q_h$  with  $(q, 1) = 0$ ,*

$$\sup_{v \in V_h} \frac{(\operatorname{div} v, q)}{\|v\|_1} \geq C \|(I - \pi_h)q\|_0.$$

*Remark.* One can improve the above estimate at best as

$$\sup_{v \in V_h} \frac{(\operatorname{div} v, q)}{\|v\|_1} \geq C(\|(I - \pi_h)q\|_0 + |\pi_h q|_h),$$

where  $|\cdot|_h$  is a seminorm with two-dimensional null space  $N_h$  (including the unity), see [10]. If  $Q_h$  is replaced by its orthogonal complement  $N_h^\perp$  with respect to this null space, then the mapping

$$q \rightarrow \sup_{v \in V_h} \frac{(\operatorname{div} v, q)}{\|v\|_1}$$

obviously defines a norm in  $N_h^\perp$ . However, this norm is still weaker than the  $L_2$ -norm: one can only estimate it as [10]

$$\sup_{v \in V_h} \frac{(\operatorname{div} v, q)}{\|v\|_1} \geq Ch\|q\|_0, \quad q \in N_h^\perp.$$

Thus, (2.11) does not hold even if  $Q_h$  is replaced by  $N_h^\perp$ .  $\square$

Now let  $(v, q) \in V_h \times Q_h$  be given with  $(q, 1) = 0$ . Then by Lemma 2.1, there exists  $z \in V_h$  satisfying

$$(2.12a) \quad \|z\|_1 \leq C\|(I - \pi_h)q\|_0,$$

$$(2.12b) \quad (\operatorname{div} z, q) = \|(I - \pi_h)q\|_0^2.$$

Now choose  $w \in V_h$  and  $r \in Q_h$  as follows:

$$w = v - \delta z, \quad r = -q,$$

where  $\delta \in (0, 1)$  is a constant to be chosen below. Then we have by (2.12a),

$$(2.13) \quad \|w\|_1 + \|r\|_0 \leq C(\|v\|_1 + \|q\|_0).$$

Moreover, recalling (2.8) and (2.5), we have by (2.12a, b),

$$\begin{aligned} \mathcal{B}_h((v, q); (w, r)) &= |v|_1^2 + \|\pi_h q\|_0^2 + \delta(\operatorname{div} z, q) - \delta(\nabla v, \nabla z) \\ &\geq (1 - C\delta)|v|_1^2 + \|\pi_h q\|_0^2 + \frac{\delta}{2}\|(I - \pi_h)q\|_0^2 \end{aligned}$$

and thus choosing  $\delta = 1/2C$ ,

$$\mathcal{B}_h((v, q); (w, r)) \geq C(|v|_1^2 + \|q\|_0^2).$$

Together with (2.13) this proves (2.10) and, accordingly, the asserted stability estimate.

To prove the error estimate we proceed from the error equation

$$\begin{aligned} \mathcal{B}_h((u_h - \tilde{u}, p_h - \tilde{p}); (v, q)) &= \mathcal{B}((u - \tilde{u}, p - \tilde{p}); (v, q)) + (\tilde{p}, \pi_h q), \\ &\quad (v, q) \in V_h \times Q_h, \end{aligned}$$

which follows easily from (2.6) through (2.8). Since  $(p_h - \tilde{p}, 1) = 0$  we conclude applying the stability estimate (2.10) and the usual estimates on the right side that

$$\|u_h - \tilde{u}\|_1 + \|p_h - \tilde{p}\|_0 \leq C(|u - \tilde{u}|_1 + \|p - \tilde{p}\|_0) + C \sup_{\substack{q \in Q_h \\ \|q\|_0=1}} (\tilde{p}, \pi_h q).$$

To estimate the last term on the right side, note that the functions  $\xi_{\nu\mu}$  in (2.4) are orthogonal to polynomials of degree  $\leq 1$  on  $L_2(T_{\nu\mu})$ . Therefore,

$$(\tilde{p}, \pi_h q) = (p, \pi_h q) = (p - \tilde{p}, \pi_h q), \quad q \in Q_h.$$

Using this in the above inequality, the asserted estimates for  $|u - u_h|_1$  and  $\|p - p_h\|_0$  follow applying the triangle inequality.

The estimate for  $\|u - u_h\|_0$  is finally proved applying the ordinary duality argument. We omit the details, so the proof is complete.  $\square$

*Remark.* Although the assumption  $\kappa = 1$  was used heavily in the above proof, it is not necessary for obtaining error estimates. For example, it can be shown that in the whole range  $0 \leq \kappa \leq 1$ , one has the estimate

$$(2.14) \quad h^{-1}\|u - u_h\|_0 + |u - u_h|_1 + \|(I - \pi_h)p_h - p\|_0 \leq Ch(|u|_2 + |p|_1),$$

i.e., the usual optimal estimate holds provided that the pressure is first smoothened by removing the oscillating component  $\pi_h p_h$ . In the case  $\kappa = 0$ , (2.14) is proved in [12], [15]. However, one further assumption is required: the functional  $g$  in (2.6) has to be such that  $g(0, \pi_h q) = 0$  for all  $q \in Q_h$ . In the above situation this holds trivially, but when the algorithm is used in a multigrid process, this assumption is no longer valid in general. This is one reason for introducing the stabilizing term into the finite element functional  $\mathcal{B}_h$ .  $\square$

*Remark.* The term  $\|p - \tilde{p}\|_0$  on the right side of the error estimate of Theorem 1 may be interpreted as the extra consistency error caused by the added term  $-(p_h, \pi_h q)$  in  $\mathcal{B}_h$ . Since  $\tilde{p}$  is defined on the coarser mesh, one might suspect that the stabilization reduces the accuracy of the method severely, say, to the level of the corresponding method on the coarser grid. This could indeed be the case if  $\tilde{p}$  were piecewise constant on  $\mathcal{C}_{2h}$ . However, with the above definition we have also the higher-order estimates

$$\|p - \tilde{p}\|_{L_2(T)} \leq Ch^2 |p|_{H^2(T)}, \quad T \in \mathcal{C}_{2h},$$

so the extra consistency error is probably relatively small in practical situations.  $\square$

**3. Convergence of a Two-Grid Algorithm.** In this section we consider a two-grid procedure for solving the system (2.2a, b') with  $\kappa = 1$ . To define the algorithm, let us associate to the partitioning  $\mathcal{C}_{2h}$  the space  $V_{2h}$  of continuous piecewise bilinear functions on  $\mathcal{C}_{2h}$  and the space  $Q_{2h}$  of piecewise constant-valued functions on  $\mathcal{C}_{2h}$ . Further, assuming that  $m_1$  and  $m_2$  in (2.3) are divisible by four, we may associate with the coarse grid the projector  $\pi_{2h}$  defined as above replacing  $\mathcal{C}_h$  by  $\mathcal{C}_{2h}$ . With this notation, one step of a two-grid algorithm consists of the following: Given an initial guess  $(u_h^0, p_h^0)$  such that  $(p_h^0, 1) = 0$  and a linear smoothing operator  $G_h^{(m)}: V_h \times Q_h \rightarrow V_h \times Q_h$  involving  $m$  steps of some iteration process, compute first  $(u_h^{(m)}, p_h^{(m)})$  as

$$(3.1) \quad (u_h^{(m)}, p_h^{(m)}) = G_h^{(m)}(u_h^{(0)}, p_h^{(0)}).$$

Then compute a correction to  $(u_h^{(m)}, p_h^{(m)})$  on the coarser grid by solving the problem: Find  $(d_{2h}, e_{2h}) \in V_{2h} \times Q_{2h}$  such that  $(e_{2h}, 1) = 0$  and

$$(3.2) \quad \mathcal{B}_{2h}((d_{2h}, e_{2h}), (v, q)) = g(v, q) - \mathcal{B}((u_h^{(m)}, p_h^{(m)}); (v, q)), \\ (v, q) \in V_{2h} \times Q_{2h}.$$

Here  $\mathcal{B}_{2h}$  is defined as  $\mathcal{B}_h$  above, except that  $\pi_h$  is replaced by  $\pi_{2h}$ , and we have assumed (as will be the case below) that the smoothing operator preserves the pressure normalization, i.e.,  $(p_h^{(m)}, 1) = 0$ . The two-grid step is now completed by adding the correction to  $(u_h^{(m)}, p_h^{(m)})$ :

$$(3.3) \quad u_h^{(m+1)} = u_h^{(m)} + d_{2h}, \quad p_h^{(m+1)} = p_h^{(m)} + e_{2h}.$$

To define an appropriate smoothing operator we follow the idea of [8], [14] of scaling the variables. Let  $\{\varphi_i, i = 1, \dots, M\}$  (where  $M = 2(m_1 - 1)(m_2 - 1)$ ) be the usual basis of  $V_h$  associated with the nodes of the rectangular grid, and let  $\varphi_i, i = M + 1, \dots, N$  (where  $N = M + m_1 m_2$ ) be the basis of  $Q_h$  consisting of the characteristic functions of the rectangles in  $\mathcal{C}_h$ . We assume each  $\varphi_i$  to be normalized so that  $\|\varphi_i\|_0 = 1$ . Then if we associate with the solution  $(u_h, p_h)$  of (2.7) a vector  $\xi \in \mathbf{R}^N$  so that

$$(3.4) \quad (u_h, p_h) = \left( \sum_{i=1}^M \xi_i \varphi_i, h^{-1} \sum_{i=M+1}^N \xi_i \varphi_i \right),$$

we may rewrite (2.7) as a system

$$A\xi = b,$$

where  $A = (a_{ij})$  is a symmetric matrix with elements defined by

$$a_{ij} = \begin{cases} (\nabla \varphi_i, \nabla \varphi_i), & i, j \leq M, \\ -h^{-1}(\operatorname{div} \varphi_i, \varphi_j), & i \leq M, j > M, \\ -h^{-2}(\pi_h \varphi_i, \pi_h \varphi_j), & i, j > M. \end{cases}$$

Here we have dropped the normalization condition for the pressure, which means that one of the eigenvalues of  $A$  is zero.

In the sequel we consider only smoothing operators  $G_h^{(m)}$  with  $m$  even,  $m \geq 2$ , which result from applying  $m/2$  single iteration steps to the positive semidefinite system

$$(3.5) \quad A^2 \xi = Ab.$$

Further, we consider for simplicity only two ways of defining a single iteration step. These are:

(1) *The Jacobi method.* With a given initial guess for  $\xi^0$ , iterate according to the rule

$$(3.6) \quad \xi^{k+1} = \xi^k - \omega |\lambda_N|^{-2} (A^2 \xi^k - Ab),$$

where  $\lambda_N$  is the largest eigenvalue of  $A$  in absolute value, and  $\omega$  is a parameter,  $0 < \omega \leq 1$ .

(2) *The conjugate gradient method.* Iterate according to the rule

$$\begin{aligned} \xi^{k+1} &= \xi^k + \alpha_k d^k, & \alpha_k &= |r^k|^2 / |Ad^k|^2, \\ d^{k+1} &= -r^{k+1} + \beta_k d^k, & \beta_k &= |r^{k+1}|^2 / |r^k|^2, \end{aligned}$$

where  $d^0 = -r^0$ ,  $r^k = A^2 \xi^k - Ab$ , and  $|\cdot|$  denotes the Euclidean norm of  $\mathbf{R}^N$ .

*Remark.* If the algorithm (3.1) through (3.3) is used as a block in a multigrid process, an appropriate work estimate is the number of arithmetic operations

required in multiplying a given vector by the matrix  $A$ . If  $m_1 = m_2 = m$  in (2.3) the multiplication requires  $64m^2 + O(m)$  operations without the stabilizing term in the functional (i.e.,  $\kappa = 0$  in (2.2b')), and  $66m^2 + O(m)$  operations in the above stabilized form. Thus, the stabilization increases the computational work only by a negligible amount.  $\square$

The main result of this section formulated below states an estimate for the convergence rate of the algorithm (3.1) through (3.3). As in [14], we measure the convergence in a scaled  $L_2$ -norm  $\|\cdot\|_{0,h}$  defined as

$$\|(u, p)\|_{0,h}^2 = \|u\|_0^2 + h^2\|p\|_0^2.$$

**THEOREM 2.** *Let  $m$  be even and let  $(u_h^{m+1}, p_h^{m+1})$  be computed from  $(u_h^{(0)}, p_h^{(0)})$  according to (3.1) through (3.3), where  $G_h^{(m)}$  in (3.3) corresponds to  $m/2$  steps of either the Jacobi iteration or the conjugate gradient algorithm as described above, and  $(p_h^{(0)}, 1) = 0$ . Then if  $(u_h, p_h)$  is the solution to (2.7) with  $(p_h, 1) = 0$ , we have the estimate*

$$\|(u_h^{(m+1)} - u_h, p_h^{(m+1)} - p_h)\|_{0,h} \leq c(m) \|(u_h^{(0)} - u_h, p_h^{(0)} - p_h)\|_{0,h},$$

where  $c(m) \leq C(\omega m)^{-1/2}$  for the Jacobi iteration and  $c(m) \leq Cm^{-1}$  for the conjugate gradient algorithm.

The proof of Theorem 2 is split below into three lemmas. We introduce first some notation. Let  $0 = \lambda_1 < |\lambda_2| \leq \dots \leq |\lambda_N|$  be the eigenvalues of the matrix  $A$  and let  $\{\psi_1, \dots, \psi_N\} \subset \mathbf{R}^N$  be the corresponding system of eigenvectors, orthonormal with respect to the Euclidean inner product  $\langle \cdot, \cdot \rangle$ . Further, let  $v \in V_h$  and  $q \in Q_h$  be expanded in terms of the basis functions of the subspaces as

$$(3.7) \quad v = \sum_{i=1}^M \eta_i \varphi_i, \quad q = h^{-1} \sum_{i=M+1}^N \eta_i \varphi_i,$$

where  $\eta \in \mathbf{R}^N$  is further expanded as

$$\eta = \sum_{i=1}^N c_i \psi_i.$$

Then we may define a scale of (semi)norms  $\|\|\| \cdot \|\|\|_s$  for  $s \geq 0$  by setting

$$\|\|\|(v, q)\|\|\|_s^2 = \sum_{i=1}^N |\lambda_i|^s c_i^2.$$

Below we need only the norm  $\|\|\| \cdot \|\|\|_0$  and the seminorm  $\|\|\| \cdot \|\|\|_2$ . Note that we have

$$(3.8a) \quad \|\|\|(v, q)\|\|\|_0 = |\eta|,$$

$$(3.8b) \quad \|\|\|(v, q)\|\|\|_2 = \sqrt{\langle A^2 \eta, \eta \rangle} = |A\eta|,$$

where  $|\cdot|$  denotes as above the Euclidean norm. Moreover, by (3.7) and by the definition of the basis functions  $\varphi_i$ , we have

$$(3.9) \quad C^{-1}|\eta| \leq \|(v, q)\|_{0,h} \leq C|\eta|$$

for some constant  $C$ , i.e.,  $\|\cdot\|_{0,h}$  and  $\|\|\| \cdot \|\|\|_0$  are equivalent norms on  $V_h \times Q_h$ .

The first step in the proof of Theorem 2 is to obtain an estimate for the smoothing effect of the operator  $G_h^{(m)}$ . The proof of the following lemma is found essentially in



[14] (cf. also [4]). For the convenience of the reader, the proof is reproduced at the end of this section.

LEMMA 3.1. *If  $G_h^{(m)}$  in (3.1) is defined as above, we have*

$$\| (u_h^{(m)} - u_h, p_h^{(m)} - p_h) \|_2 \leq Ch^{-2} c(m) \| (u_h^{(0)} - u_h, p_h^{(0)} - p_h) \|_0,$$

where  $c(m)$  is as in Theorem 1.

The next step in the proof is now to verify that the solution of (3.2) satisfies the estimate

$$(3.10) \quad \| (d_{2h} - d, e_{2h} - e) \|_{0,h} \leq Ch^2 \| (d, e) \|_2,$$

where  $(d, e) = (u_h - u_h^{(m)}, p_h - p_h^{(m)})$ . The statement of Theorem 2 then follows by combining (3.3), (3.10) and Lemma 3.1. The proof of (3.10) is also given essentially in [14], where it is carried out in the spirit of [4] using duality techniques. Here we present another proof which appears to be a bit more transparent. It is based on the auxiliary seminorms.

$$|v|_{2,h}^2 = \sum_{K \in \mathcal{C}_h} |v|_{H^2(T)}^2 + h^{-1} \sum_{S \in \tau_h} \int_S \left| \left[ \frac{\partial v}{\partial n} \right] \right|^2 ds,$$

and

$$|q|_{1,h}^2 = \sum_{K \in \mathcal{C}_h} |q|_{H^1(T)}^2 + h^{-1} \sum_{S \in \tau_h} \int_S [q]^2.$$

Here  $\tau_h$  denotes the set of sides of rectangles in  $\mathcal{C}_h$  that are interior to  $\Omega$ ,  $[q]$  is the jump of  $q$  across  $S$  and  $[\partial v / \partial n]$  is the jump across  $S$  of the derivative of  $v$  with respect to the normal of  $S$ . These seminorms are associated, respectively, with the spaces

$$[H_h^2(\Omega)]^2 = \{ v \in V: v|_T \in [H^2(T)]^2, T \in \mathcal{C}_h \},$$

and

$$H_h^1(\Omega) = \{ q \in Q: q|_T \in H^1(T), T \in \mathcal{C}_h \}.$$

Using the above notation, we now split the proof of (3.10) into the following two lemmas.

LEMMA 3.2. *There is a constant  $C$  such that for all  $(v, q) \in V_h \times Q_h$ ,*

$$|v|_{2,h} + |q|_{1,h} \leq C \| (v, q) \|_2.$$

LEMMA 3.3. *The solution of (3.2) satisfies the error estimate*

$$\| (d_{2h} - d, e_{2h} - e) \|_{0,h} \leq Ch^2 (|d|_{2,h} + |e|_{1,h}),$$

where  $(d, e) = (u_h - u_h^{(m)}, p_h - p_h^{(m)})$ .

*Proof of Lemma 3.1.* Consider first the Jacobi method, where according to (3.4) through (3.6),

$$\xi^m - \xi = \left( I - \omega |\lambda_N|^{-2} A^2 \right)^{m/2} (\xi^0 - \xi).$$

By (3.8), the asserted inequality is equivalent to the estimate

$$|\xi^m - \xi|_A \leq Ch^{-2}c(m)|\xi^0 - \xi|,$$

where  $|\eta|_A = |A\eta|$ . Now if  $\xi^0 - \xi = \sum_{i=1}^N c_i \psi_i$ , we have

$$\begin{aligned} |\xi^m - \xi|_A^2 &= \left\langle A^2(I - \omega|\lambda_N|^{-2}A^2)^m(\xi^0 - \xi), \xi^0 - \xi \right\rangle \\ &= \sum_{i=1}^N \lambda_i^2 \left[ 1 - \omega \left( \frac{\lambda_i}{\lambda_N} \right)^2 \right]^m c_i^2 \\ &\leq \omega^{-1} |\lambda_N|^2 \max_{-1 \leq x \leq 1} x^2 (1 - x^2)^m |\xi^0 - \xi|^2. \end{aligned}$$

To estimate  $|\lambda_N|$ , note that by (3.9) and by definition of the matrix  $A$ ,

$$|\lambda_N| = \max_{\eta \in \mathbb{R}^N} \frac{|\langle A\eta, \eta \rangle|}{|\eta|^2} \leq C \max_{(v,q) \in V_h \times Q_h} \frac{|\mathcal{B}_h((v,q); (v,q))|}{\|(v,q)\|_{0,h}^2}.$$

By a simple inverse estimate,

$$\begin{aligned} |\mathcal{B}_h((v,q); (v,q))| &\leq C(|v|_1^2 + \|q\|_0^2) \\ &\leq Ch^{-2} \|(v,q)\|_{0,h}^2, \quad (v,q) \in V_h \times Q_h. \end{aligned}$$

Therefore,  $|\lambda_N| \leq Ch^{-2}$ , and since

$$\max_{-1 \leq x \leq 1} |x|(1 - x^2)^{m/2} \leq C/\sqrt{m},$$

the asserted estimate follows.

Consider next the conjugate gradient method. In this case we have the identity (cf. [1])

$$|\xi^m - \xi|_A = \min_{r \in \Lambda_{m/2}} |r(A^2)(\xi^0 - \xi)|_A,$$

where  $\Lambda_{m/2}$  denotes the set of polynomials  $r$  of degree  $\leq m/2$  such that  $r(0) = 1$ . Recalling that  $|\lambda_N| \leq Ch^{-2}$ , this gives the upper bound

$$|\xi^m - \xi|_A \leq Ch^{-2} \min_{r \in \Lambda_{m/2}} \max_{-1 \leq x \leq 1} |xr(x^2)| |\xi^0 - \xi|.$$

Since  $m$  is even, it is possible to choose

$$r(x^2) = \frac{1}{m+1} x^{-1} T_{m+1}(x),$$

where  $T_{m+1}$  is the Chebyshev polynomial of degree  $m+1$ . This choice yields immediately the asserted estimate, so the lemma is proved.  $\square$

*Proof of Lemma 3.2.* Let  $(v,q) \in V_h \times Q_h$  and let  $\eta$  be as in (3.7). Then recalling (3.9) and the definition of the matrix  $A$ , we have

$$\|(v,q)\|_2 = |A\eta| = \sup_{\zeta \in \mathbb{R}^N} \frac{\langle A\eta, \zeta \rangle}{|\zeta|} \geq C \sup_{(w,r) \in V_h \times Q_h} \frac{\mathcal{B}_h((v,q); (w,r))}{\|(w,r)\|_{0,h}}.$$

Therefore, if  $(\omega, \rho) \in V_h \times Q_h$  is such that

$$(3.11) \quad \mathcal{B}_h((v,q), (w,r)) = (\omega, w) + h^2(\rho, r), \quad (w,r) \in V_h \times Q_h,$$

we have

$$\|(\omega, \rho)\|_{0,h} \leq C \|(v,q)\|_2,$$

and so the asserted estimate follows if we can show that

$$(3.12) \quad |v|_{2,h} + |q|_{1,h} \leq C \|(\omega, \rho)\|_{0,h},$$

which is a stability estimate for (3.11).

To prove (3.12), let  $(v, q) = (v^{(1)}, q^{(1)}) + (v^{(2)}, q^{(2)})$ , where  $(v^{(i)}, q^{(i)}) \in V_h \times Q_h$  satisfies  $(q^{(2)}, 1) = 0$  and

$$\mathcal{B}_h((v^{(1)}, q^{(1)}), (w, r)) = (\omega, w), \quad (w, r) \in V_h \times Q_h,$$

and

$$\mathcal{B}_h((v^{(2)}, q^{(2)}), (w, r)) = h^2(\rho, r), \quad (w, r) \in V_h \times Q_h.$$

Then  $(v^{(1)}, q^{(1)})$  is the approximate solution of the Stokes problem (2.1) with  $f = \omega$ . Denote by  $(u, p)$  the corresponding exact solution normalized so that  $(p, 1) = (q^{(1)}, 1)$  and let  $(\tilde{u}, \tilde{p})$  be the interpolant of  $(u, p)$  in  $V_h \times Q_h$ . Then applying first the triangle inequality,

$$\begin{aligned} |v^{(1)}|_{2,h} + |q^{(1)}|_{1,h} &\leq |v^{(1)} - \tilde{u}|_{2,h} + |q^{(1)} - \tilde{p}|_{1,h} \\ &\quad + |u - \tilde{u}|_{2,h} + |p - \tilde{p}|_{1,h} + |u|_{2,h} + |p|_{1,h} \end{aligned}$$

and on the right side the further estimates

$$\begin{aligned} |v^{(1)} - \tilde{u}|_{2,h} + |q^{(1)} - \tilde{p}|_{1,h} &\leq Ch^{-1}(|v^{(1)} - \tilde{u}|_1 + \|q^{(1)} - \tilde{p}\|_0) \\ &\leq C_1(|u|_2 + |p|_1), \\ |u - \tilde{u}|_{2,h} + |p - \tilde{p}|_{1,h} &\leq C(|u|_2 + |p|_1) \end{aligned}$$

(where we used standard inverse and interpolation error estimates (cf. also [13]), together with the error estimate of Theorem 1), we conclude, noting that  $|u|_{2,h} = |u|_2$  and  $|p|_{1,h} = |p|_1$  and applying the standard regularity estimate, that

$$|v^{(1)}|_{2,h} + |q^{(1)}|_{1,h} \leq C(|u|_2 + |p|_1) \leq C_1 \|\omega\|_0.$$

In estimating  $(v^{(2)}, q^{(2)})$  we need only to combine the stability estimate of Theorem 1 with inverse estimates to obtain

$$|v^{(2)}|_{2,h} + |q^{(2)}|_{1,h} \leq Ch^{-1}(|v^{(2)}|_1 + \|q^{(2)}\|_0) \leq C_1 h \|\rho\|_0.$$

Combining now the estimates for  $(v^{(i)}, q^{(i)})$ ,  $i = 1, 2$ , we obtain (3.12), and the lemma is thus proved.

*Proof of Lemma 3.3.* Comparing (2.7) and (3.2) and noting that  $\pi_h q = 0$  for  $q \in Q_{2h}$ , we obtain

$$\begin{aligned} \mathcal{B}_{2h}((d_{2h}, e_{2h}); (v, q)) &= \mathcal{B}_{2h}((d, e); (v, q)) + (e, \pi_{2h} q), \\ &\quad (v, q) \in V_{2h} \times Q_{2h}. \end{aligned}$$

Arguing as in the proof of Theorem 1, we obtain the error estimate

$$\|(d_{2h} - d, e_{2h} - e)\|_{0,h} \leq Ch(|d - \tilde{d}|_1 + \|e - \tilde{e}\|_0 + \|e - \tilde{e}\|_0),$$

where the interpolants  $\tilde{d} \in V_{2h}$ ,  $\tilde{e} \in Q_{2h}$  and  $\tilde{\tilde{e}}$  are the coarse-mesh analogies of those in Theorem 1. The asserted estimate obviously follows if we can prove the interpolation error estimates

$$(3.13) \quad |d - \tilde{d}|_1 \leq Ch|d|_{2,h}, \quad d \in V_h,$$

and

$$(3.14) \quad \|e - \tilde{e}\|_0 + \|e - \tilde{e}\|_1 \leq Ch|e|_{1,h}, \quad e \in Q_h.$$

To prove (3.13), let  $K_1, K_2 \in \mathcal{C}_h$  be two rectangles with a common side  $S$ . Set  $K = K_1 \cup K_2 \cup S$  and define the seminorm  $|v|_{2,h,K}$  for  $v \in V_h$  by

$$|v|_{2,h,K}^2 = |v|_{H^2(K_1)}^2 + |v|_{H^2(K_2)}^2 + h^{-1} \int_S \left| \left[ \frac{\partial v}{\partial n} \right] \right|^2 ds,$$

where we use the same notation as in defining the seminorm  $|\cdot|_{2,h}$  above. It is easy to see that the null space of  $|\cdot|_{2,h,K}$  is contained in the null space of the seminorm  $v \rightarrow |v - \tilde{v}|_{H^1(K)}$ , where  $\tilde{v}$  is the interpolant of  $v$  in  $V_{2h}$ . Therefore, by a simple scaling argument and by the equivalence of norms in a finite-dimensional space, there is a constant  $C$  independent of  $K$  such that

$$|v - \tilde{v}|_{H^1(K)} \leq Ch|v|_{2,h,K}, \quad v \in V_h.$$

When summed over all pairs of adjacent rectangles in  $\mathcal{C}_h$ , this proves (3.13). Estimate (3.14) is proved in a similar manner, and so the assertion of the lemma follows.  $\square$

**4. Numerical Results.** The algorithm of Section 3 was tested by solving numerically the Stokes equations on the unit square, with  $f$  in (2.1) chosen so that the exact solution is given by

$$\begin{aligned} u_1(x) &= x_1^2(1-x_1)^2(2x_2 - 6x_2^2 + 4x_2^3), \\ u_2(x) &= x_2^2(1-x_2)^2(-2x_1 + 6x_1^2 - 4x_1^3), \\ p(x) &= x_1^2 - x_2^2. \end{aligned}$$

A three-level algorithm with  $W$ -cycle (cf. [13]) was applied to solve the problem on a  $16 \times 16$  grid. The relaxation steps were distributed symmetrically with respect to coarse grid corrections, and on the coarsest  $4 \times 4$  grid the discrete equations were solved directly as usual.

In the numerical tests the scaling parameter  $h$  in (3.4) was first replaced by  $\gamma h$  and  $\gamma$  was varied so as to obtain the best convergence rates. The optimal value was found to be close to 0.5. Similarly, the value of the parameter  $\omega$  in the Jacobi relaxation was chosen close to the optimum, i.e.,  $\omega \approx 1$ . In the subsequent computations the parameters  $\gamma$  and  $\omega$  were held fixed and asymptotic convergence rates in the norm  $\|\cdot\|_{0,h}$  were computed as a function of the remaining parameters, i.e., the stabilization parameter  $\kappa$  and the number of relaxation steps per level.

Table 1 shows the dependence of the convergence rate on the stabilization parameter  $\kappa$  in the case of eight conjugate gradient relaxations per level. Similar results were obtained with other values of  $m$  and with the Jacobi relaxation. Thus we conclude that the stabilization, although unnecessary in the direct solution, is of essential importance in the multigrid solution.

Table 2 shows the dependence of the convergence rate on the relaxation method and on the number of relaxation steps per level. As expected, the conjugate gradient relaxation gives faster convergence for large  $m$ . However, for small  $m$  the Jacobi relaxation seems superior as it is simpler.

We should finally point out that very likely neither of the relaxation methods considered in this paper is the best possible one. The development of smoothing algorithms that give better convergence rates for small  $m$  no doubt deserves further studies. For previous work in this direction, cf. Brandt and Dinar [16].

TABLE 1

*Asymptotic convergence factors of a three-level algorithm for various values of the stabilization parameter. Eight conjugate gradient relaxations per level.*

$\kappa = 0$	0.1	0.25	0.5	1.0	2.5
.94	.50	.22	.22	.23	.48

TABLE 2

*Asymptotic convergence factors of a three-level algorithm for various types of relaxation.  $\kappa = 0.5$ .*

	$m =$	2	4	8
Jacobi		.87	.81	.71
Conjugate gradient		.87	.62	.22

Institute of Mathematics  
Helsinki University of Technology  
SF-02150 Espoo 15, Finland

1. O. AXELSSON, "Solution of linear systems of equations: iterative methods," in *Sparse Matrix Techniques*, Lecture Notes in Math., Vol. 572 (V. A. Barker, ed.), Springer-Verlag, Berlin and New York, 1977.
2. I. BABUŠKA, "Error bounds for finite element methods," *Numer. Math.*, v. 16, 1971, pp. 322–333.
3. I. BABUŠKA, J. OSBORN & J. PITKÄRANTA, "Analysis of mixed methods using mesh dependent norms," *Math. Comp.*, v. 35, 1980, pp. 1039–1062.
4. R. E. BANK & T. DUPONT, "An optimal order process for solving elliptic finite element equations," *Math. Comp.*, v. 36, 1981, pp. 35–51.
5. F. BREZZI, "On the existence, uniqueness and approximation of saddle-point problems arising from Lagrange multipliers," *RAIRO Ser. Rouge*, v. 8, no. R2, 1974, pp. 129–151.
6. M. CROUZEIX & P.-A. RAVIART, "Conforming and nonconforming finite element methods for solving the stationary Stokes equations," *RAIRO Ser. Rouge*, v. 7, no. R3, 1973, pp. 33–76.
7. V. GIRAULT & P.-A. RAVIART, *Finite Element Approximation of the Navier-Stokes Equations*, Lecture Notes in Math., Vol. 749, Springer-Verlag, Berlin and New York, 1979.
8. W. HACKBUSCH, *Analysis and Multigrid Solutions of Mixed Finite Element and Mixed Difference Equations*, Preprint, Universität Bochum, 1980.
9. T. J. HUGHES, W. K. LIU & A. BROOKS, "Finite element analysis of incompressible viscous flows by the penalty function formulation," *J. Comput. Phys.*, v. 30, 1979, pp. 1–60.
10. C. JOHNSON & J. PITKÄRANTA, "Analysis of some mixed finite element methods related to reduced integration," *Math. Comp.*, v. 38, 1982, pp. 375–400.
11. J. PITKÄRANTA, "On a mixed finite element method for the Stokes problem in  $\mathbb{R}^3$ ," *RAIRO Anal. Numér.*, v. 16, 1982, pp. 275–291.

12. J. PITKÄRANTA & R. STENBERG, *Error Bounds for the Approximation of the Stokes Problem Using Bilinear/Constant Elements on Irregular Quadrilateral Meshes*, Proceedings of the MAFELAP-1984 conference. (To appear.)
13. K. STÜBEN & U. TROTTEBERG, "Multigrid methods: fundamental algorithms, model problem analysis and applications," in *Multigrid Methods*, Lecture Notes in Math., Vol. 960 (W. Hackbusch & U. Trottenberg, eds.), Springer-Verlag, Berlin and New York, 1982.
14. R. VERFÜRTH, *A Multilevel Algorithm for Mixed Problems*, Preprint, Universität Bochum, 1982.
15. J. M. BOLAND & R. A. NICOLAIDES, *Stable and Semistable Low Order Finite Elements for Viscous Flows*, Preprint, University of Connecticut, 1983.
16. A. BRANDT & N. DINAR, "Multigrid solutions to elliptic flow problems," in *Numerical Methods for Partial Differential Equations* (S. V. Parter, ed.), Academic Press, New York, 1979.