Application of the satisfying Babuska-Brezzi method to a two dimensional diffusion problem

Damman, C.M.; Baaijens, F.P.T.

Published in:
Applied Mathematical Modelling

DOI:
10.1016/0307-904X(93)90106-Q

Published: 01/01/1993

Document Version
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the author’s version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 05. Dec. 2018
Application of the satisfying Babuška–Brezzi method to a two-dimensional diffusion problem

Carolien M. Kootwijk

Philips Research Laboratories, Eindhoven, The Netherlands

Frank P. T. Baaijens

Philips Research Laboratories and Eindhoven University of Technology, Faculty of Mechanical Engineering, Eindhoven, The Netherlands

The satisfying Babuška–Brezzi (SBB) method is applied to a two-dimensional diffusion problem. Several interpolation schemes, both continuous and discontinuous, are investigated numerically. The superconvergence phenomenon observed in the one-dimensional case is not recovered in two dimensions. However, SBB does stabilize the numerical solution substantially, particularly for low-order interpolations.

Keywords: mixed finite element method, SBB method, diffusion problem

Introduction

Mixed finite element methods are becoming increasingly popular. They offer additional degrees of freedom in designing new elements, but nonstandard variational principles are needed to overcome certain difficulties that arise when the traditional Galerkin finite element method is applied to mixed problems. A good illustration is found in Hughes and Franca.¹

In the one-dimensional situation, Franca² reported superconvergence of the so-called satisfying Babuška–Brezzi (SBB) method for the one-dimensional diffusion problem posed in a mixed setting. Even when using discontinuous interpolations for the flux variable, continuous solutions were obtained.

The purpose of this work is to investigate whether similar results can be achieved for two-dimensional problems. This would be particularly interesting for, e.g., injection molding computations. For such problems the so-called pressure problem (which is a generalized Reynolds equation³,⁴) is to be solved, while the velocity field is related to the flux of the pressure field. Mostly linear triangular elements are used for these analyses, giving discontinuous velocity fields. This discontinuity is held responsible for the occasional numerical instabilities that occur in computing the temperature distribution. This temperature field depends on the velocity distribution via advective terms in the energy equation. Therefore, Sitters³ suggested introduction of an element with continuous pressure gradients. Another application would be the analysis of viscoelastic flow⁵ where a mixed stress–velocity–pressure formulation is used. Singularities in the flow domain may yield oscillatory stress fields that may possibly be stabilized by SBB.

The outline of this report is as follows. In the next section, the theory of SBB is recalled. Following this, a two-dimensional problem is introduced, and several elements are proposed.⁶ Finally, numerical results are presented.

Theory

The general mixed form of a boundary value problem can be written as

\[ A\sigma + B^*u = f \]  
\[ B\sigma + Cu = g \]

plus boundary condition terms. \( A, B, \) and \( B^* \) are differential operators, with \( B \) and \( B^* \) conjugate, and \( \sigma, u, f, \) and \( g \) are scalars, vectors, or second-order tensors.

Now, consider a domain \( \Omega \) in \( \mathbb{R}^p \), where \( p \) denotes the dimension of the space. To derive the weak form of the problem, two sets of trial solutions and two sets of weighting functions are needed. The sets of trial solutions are \( \mathcal{S} \) for \( \sigma \), respectively. \( U \) for \( u \) and the sets of

Address reprint requests to Prof. Baaijens at Eindhoven University of Technology, Faculty of Mechanical Engineering, P.O. Box 513, 5600 MB Eindhoven, The Netherlands.

Received 1 August 1991; revised 8 June 1992; accepted 20 July 1992
weighting functions are \( T \) for the weighting function \( \tau \), respectively. \( W \) for the weighting function \( w \). To derive the weak form, equation (1) is multiplied with an arbitrary weighting function \( \tau \in T \) and equation (2) with an arbitrary weighting function \( w \in W \) and integrating over \( \Omega \):

**Problem 1 (MX).** Given \( f: \Omega \to \mathbb{R}^m \) and \( g: \Omega \to \mathbb{R}^n \), find \( \{\sigma, u\} \in S \times U \) such that

\[
\begin{align*}
(\tau, A\sigma + B^*u - f) &= 0 \quad \forall \tau \in T \tag{3} \\
(w, B\sigma + Cu - g) &= 0 \quad \forall w \in W \tag{4}
\end{align*}
\]

plus natural boundary conditions.

The notation \( (v, w) \) is defined by

\[
(v, w) = \int_{\Omega} v \cdot w \, d\Omega
\]

where \( v \) and \( w \) may both be scalars, vectors, or second-order tensors, and "." accounts for the appropriate inner product.

A finite element approach is obtained by performing an integration by parts and by choosing the solution from \( S, U, T, \) and \( W, \) finite dimensional approximations of \( S, U, T, \) and \( W, \) respectively. This gives the Galerkin formulation:

**Problem 2 (MG).** Given \( f: \Omega \to \mathbb{R}^m \) and \( g: \Omega \to \mathbb{R}^n \), find \( \{\sigma_h, u_h\} \in S_h \times U_h \) such that

\[
\begin{align*}
(a_h(\tau_h, \sigma_h) + b_h(\tau_h, u_h)) &= f_h(\tau_h) \quad \forall \tau_h \in T_h \tag{5} \\
b_h(\sigma_h, w_h) + c_h(w_h, u_h) &= g_h(w_h) \quad \forall w_h \in W_h \tag{6}
\end{align*}
\]

plus natural boundary conditions. Starting with (MX), the SBB formulation of the problem has the form:

**Problem 3 (SBB).** Given \( f: \Omega \to \mathbb{R}^m \) and \( g: \Omega \to \mathbb{R}^n \), find \( \{\sigma, u\} \in S \times U \) such that

\[
\begin{align*}
(a(\tau, \sigma) + b^*(u, \tau) - f) &= 0 \quad \forall \tau \in T \tag{7} \\
b(\sigma, w) + c(w, u) &= g(w) \quad \forall w \in W \tag{8}
\end{align*}
\]

plus natural boundary conditions.

in which \( a \) is a nonnegative stability constant. Discretization, performing an integration by parts, and choosing \( a = \alpha \delta h^2 \), with \( \delta > 0 \) and \( \alpha \geq 0 \), leads to the general form of the SBB formulation:

**Problem 4 (SBBa).** Given \( f: \Omega \to \mathbb{R}^m \) and \( g: \Omega \to \mathbb{R}^n \), find \( \{\sigma_h, u_h\} \in S_h \times U_h \) such that

\[
\begin{align*}
a_h(\tau_h, \sigma_h) &= a(\tau_h, \sigma_h) - (\delta h^2 B \tau_h, B \sigma_h)_h \\
b_{1h}(\tau_h, u_h) &= b(\tau_h, u_h) - (\delta h^2 B \tau_h, C u_h)_h \\
b_{2h}(w_h, \sigma_h) &= b(\sigma_h, w_h) \\
c_h(w_h, u_h) &= c(w_h, u_h)
\end{align*}
\]

where

\[
\begin{align*}
a_h(\tau_h, \sigma_h) &= a(\tau_h, \sigma_h) - (\delta h^2 B \tau_h, B \sigma_h)_h \\
b_{1h}(\tau_h, u_h) &= b(\tau_h, u_h) - (\delta h^2 B \tau_h, C u_h)_h \\
b_{2h}(w_h, \sigma_h) &= b(\sigma_h, w_h) \\
c_h(w_h, u_h) &= c(w_h, u_h)
\end{align*}
\]

A two-dimensional problem

In this section, the SBB formulation is applied to a two-dimensional problem. Several element choices are investigated.

Consider a domain \( \Omega \) in \( \mathbb{R}^2 \). The domain \( \Omega \) has boundary \( \Gamma \), with unit outward normal \( \vec{n} \), that is divided into \( \Gamma_u \) and \( \Gamma_p \) such that \( \Gamma = \Gamma_u \cup \Gamma_p \) and \( \Gamma_u \cap \Gamma_p = \emptyset \). Among \( \Gamma_u \) the essential boundary conditions are prescribed, and along \( \Gamma_p \) the natural boundary conditions are given. Now consider the following two-dimensional problem

\[
\nabla \cdot (c \nabla u) = - f
\]

with

\[
\nabla \cdot (c \nabla u) = - f
\]

and \( u = u(x, y), c = c(x, y), \) and \( f = f(x, y) \) while \( \{\vec{e}_x, \vec{e}_y\} \) is a Cartesian vector base in \( \mathbb{R}^2 \).

A physical interpretation of this model problem is for instance the steady heat conduction problem, in which \( \nabla u \) represents the heat flux, \( c \) the thermal conductivity, and \( f \) a source term. It has the same mathematical structure as the generalized Reynolds problem used in injection molding simulations.\(^4\)

The mixed, strong form of the boundary value problem is

**Problem 5 (MS).** Given \( f: \Omega \to \mathbb{R} \) and \( c: \Omega \to \mathbb{R} \) and a constant \( p \), find \( u: \Omega \to \mathbb{R} ) \) and \( \tilde{\sigma}: \Omega \to \mathbb{R}^2 \) such that

\[
\begin{align*}
\sigma \cdot \nabla u &= 0 \\
\nabla \cdot \tilde{\sigma} &= f \\
u &= 0 \quad \text{on } \Gamma_u \\
\tilde{\sigma} \cdot \vec{n} &= p \quad \text{on } \Gamma_p
\end{align*}
\]

Application of SBB method: C. M. Kootwijk and F. P. T. Baaijens

Comparing (MS) with equations (1) and (2) gives

\[ A\sigma = \frac{-\tilde{\sigma}}{c} \]  
(19)

\[ B^T\mathbf{u} = \tilde{\mathbf{v}}u \]  
(20)

\[ B\sigma = -\tilde{\mathbf{v}} \cdot \tilde{\sigma} \]  
(21)

\[ C\mathbf{u} = 0 \]  
(22)

\[ f = 0 \]  
(23)

\[ g = f \]  
(24)

To define the weak form of (MS), the following classes of functions are defined:

**Trial solutions**

\[ U = \{ u | u \in H^1(\Omega), u = 0 \text{ on } \Gamma_u \} \]

**Weighting functions**

\[ W = \{ w | w \in H^1(\Omega), w = 0 \text{ on } \Gamma_w \} \]

where \( H^1(\Omega) = \{ u | u, u' \in L^2(\Omega) \} \)

and \( L^2(\Omega) \) is the space of square integrable functions on the domain \( \Omega \)

\[ L^2(\Omega) = \left\{ u \left| \int_{\Omega} u^2 d\Omega < \infty \right. \right\} \]

The notation \( \tilde{\tau} \in [L^2(\Omega)]^2 \) is used to indicate that each component of the vector \( \tilde{\tau} \), with respect to some appropriate vector base, must be in \( L^2(\Omega) \).

The weak form of (MS) is defined by

**Problem 6 (MW).** Given \( f: \Omega \to R \) and \( c: \Omega \to R \) and a constant \( p \), find \( \{ \tilde{\sigma}_h, \tilde{\mathbf{u}}_h \} \in S_h \times U_h \) such that

\[ a(\tilde{\tau}, \tilde{\sigma}) + b(\tilde{\tau}, \tilde{\mathbf{u}}) = 0 \quad \forall \tilde{\tau} \in T \]

\[ b(\tilde{\sigma}, w) = f(w) \quad \forall w \in W \]

with

\[ a(\tilde{\tau}, \tilde{\sigma}) = -\int_{\Omega} \frac{\tilde{\tau} \cdot \tilde{\sigma}}{c} d\Omega \]  
(25)

\[ b(\tilde{\tau}, u) = \int_{\Omega} \tilde{\tau} \cdot \tilde{\mathbf{v}} u d\Omega \]  
(26)

\[ f(w) = (w, f) + \int_{\Gamma} wp d\Gamma \]  
(27)

Taking \( U_h \subset U \), \( W_h \subset W \), \( S_h \subset S \), and \( T_h \subset T \) to be spanned by piecewise polynomial interpolations with continuous functions for \( U_h \) and \( W_h \) and with possibly discontinuous functions for \( S_h \) and \( T_h \), the Galerkin approximation is given by

**Problem 7 (MG).** Given \( f: \Omega \to R \) and \( c: \Omega \to R \) and constants \( u_0 \) and \( \rho \), find \( \{ \tilde{\sigma}_h, \tilde{\mathbf{u}}_h \} \in S_h \times U_h \) such that

\[ a(\tilde{\tau}_n, \tilde{\sigma}_h) + b(\tilde{\tau}_n, \tilde{\mathbf{u}}_h) = 0 \quad \forall \tilde{\tau}_n \in T_n \]

\[ b(\tilde{\sigma}_h, w_h) = f(w_h) \quad \forall w_h \in W_h \]

According to the general SBB form and with equations (19)–(24), the SBB formulation of the two-dimensional boundary value problem is

**Problem 8 (SBB).** Given \( f: \Omega \to R \) and \( c: \Omega \to R \) and a constant \( p \), find \( \{ \tilde{\sigma}_h, \tilde{\mathbf{u}}_h \} \in S_h \times U_h \) such that

\[ a(\tilde{\tau}_n, \tilde{\sigma}_h) + b(\tilde{\tau}_n, \tilde{\mathbf{u}}_h) = 0 \quad \forall \tilde{\tau}_n \in T_n \]

\[ b(\tilde{\sigma}_h, w_h) = f(w_h) \quad \forall w_h \in W_h \]

where

\[ a(\tilde{\tau}_n, \tilde{\sigma}_h) = a(\tilde{\tau}_n, \tilde{\sigma}) - (\delta h)^2 \tilde{\mathbf{v}} \cdot \tilde{\tau}_n \cdot \tilde{\mathbf{v}} \cdot \tilde{\sigma}_h \]

\[ g_h(\tilde{\tau}_n) = (\delta h)^2 \tilde{\mathbf{v}} \cdot \tilde{\tau}_n \cdot f \]

Notice that \( \delta = 1 \) is chosen.

In matrix form, (SBB) can be written as (see Appendix)

\[ -A_\delta \sigma + B^T \mathbf{u} = G_\delta \]  
(28)

\[ -B \sigma = F \]  
(29)

where

\[ \sigma^T = [\sigma_{11}, \sigma_{12}, \ldots, \sigma_{nm}, \sigma_{nm}] \]

\[ \mathbf{u}^T = [u_1, \ldots, u_n] \]

and \( m \) and \( n \) are the total number of \( \tilde{\sigma} \)-, respectively, \( \mathbf{u} \)-nodes in the assembly. \( A_\delta \) and \( B \) are the matrices associated with \( u_h \), \( \tilde{\sigma}_h \), and \( b(\tilde{\tau}_n, \tilde{\mathbf{u}}_h) \) and \( G_\delta \) and \( F \) are load vectors associated with \( g_h(\tilde{\tau}_n) \) and \( f(w_h) \).

If \( \tilde{\sigma}_h \) is interpolated discontinuously, provided \( (A_\delta)^{-1} \) exists, the \( \sigma \)-degrees of freedom can be eliminated on the element level to get the following equation in \( \mathbf{u} \)

\[ \mathbf{Ku} = \mathbf{F} \]  
(30)

where

\[ K = \sum_{e=1}^{ne} K^e \quad \text{with} \quad K^e = B^e (A_\delta)^{-1} (B^e)^T \]  
(31)

\[ F = \sum_{e=1}^{ne} F^e \quad \text{with} \quad F^e = F^e + B^e (A_\delta)^{-1} G^e \]  
(32)

The symbol \( A_{\delta_{12}} \) represents the assembly process. Now, \( \sigma \) can be calculated on the element level by

\[ \sigma^e = (A_\delta)^{-1} (B^e)^T \mathbf{u}^e - (A_\delta)^{-1} G^e \]  
(33)

When \( \tilde{\sigma}_h \) is interpolated continuously, both \( \mathbf{u} \) and \( \sigma \) must be solved from the following matrix equation:

\[ \begin{bmatrix} -A_\delta & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} \sigma \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} G_\delta \\ \mathbf{F} \end{bmatrix} \]  
(34)
The following elements are considered:
- triangular elements with three-node equal-order interpolation, six-node equal-order interpolation, seven-node equal-order interpolation, and seven-node interpolation for \( u_h \) and three-node interpolation for \( c_h \)
- rectangular elements with nine-node equal-order interpolation

The calculations are made with the SEPRAN finite element package. There are two methods to represent \( \tilde{c}_h \), whereby method one can be divided into two representations

1. \( \tilde{c}_h \) is interpolated discontinuously and (a) represented per element and (b) represented at each nodal point by averaging contributions of adjacent elements
2. \( \tilde{c}_h \) is interpolated continuously and represented at each nodal point

Method 1 uses equation (30) to calculate \( u \) and equation (33) to calculate \( \sigma \), whereby representation 1b displays the mean of the value of elements sharing the same nodal point. Method 2 uses equation (34) to calculate both \( u \) and \( \sigma \). The different element types are shown in Figure 1.

Numerical results

The following problem is analyzed to compare the SBB formulation with the Galerkin formulation, which is obtained by choosing \( S = 0 \).

\[
\frac{\partial \sigma_x}{\partial x} + \frac{\partial \sigma_y}{\partial y} = -f
\]

\[
\sigma_x = \frac{\partial u}{\partial x}
\]

\[
\sigma_y = \frac{\partial u}{\partial y}
\]

with \( u : \Omega \rightarrow R \) and

\[
\Omega = \{x, y|x, y \in R, 0 \leq x \leq 1, 0 \leq y \leq 1\}
\]

Because the problem is symmetric, the results are only considered for \( u_h \) and \( \sigma_{xx} \).

First, two problems with discontinuities in \( f \) are considered (with element size 0.1):

1. \( f = 100 \) if \( \{0.4 \leq x \leq 0.5 \text{ and } 0.4 \leq y \leq 0.5\} \) and \( f = 0 \) elsewhere, with boundary conditions \( u(0, y) = u(1, y) = u(x, 0) = u(x, 1) = 0 \) (see Figure 3).
2. \( f = 100 \) if \( \{0.4 \leq x \leq 0.5 \text{ and } 0 \leq y \leq 1\} \) and \( f = 0 \) elsewhere, with boundary conditions \( u(0, y) = u(1, y) = u(x, 0) = u(x, 1) = 0 \) (see Figure 3).
elsewhere, with boundary conditions $u(0, y) = u(1, y) = 0$ (see Figure 3).

These discontinuities in $f$ cause instabilities in $\sigma_{xf}$ for the three-node and the six-node triangular elements. This is shown in Figures 4 and 5, where $\sigma_{xf}$ is plotted as a function of $(x, y)$. In these figures $\sigma_{xf}$ is interpolated continuously on a $10 \times 10$ mesh, either by using a six-node triangular element (Figure 4) or a three-node triangular element (Figure 5). The instabilities get smaller if SBB is used.

For the other tests $f = 100$ is chosen on $\Omega$ with boundary conditions:

$$u(0, y) = u(1, y) = u(x, 0) = u(x, 1) = 0$$

so $\Gamma_u = \Gamma$ and $\Gamma_p = \phi$. The exact solution of this problem is

$$u(x, y) = 50 \left\{ y(1 - y) - \frac{8}{\pi^2} \sum_{n=0}^{\infty} \frac{1}{(2n + 1)^2} \frac{\cosh[(2n + 1)\pi(x - \frac{1}{2})]}{\cosh[(n + \frac{1}{2})\pi]} \sin[(2n + 1)\pi x] \right\}$$

$$\sigma_x(x, y) = 50 \left\{ -\frac{8}{\pi^2} \sum_{n=0}^{\infty} \frac{1}{(2n + 1)^2} \frac{\sinh[(2n + 1)\pi(x - \frac{1}{2})]}{\cosh[(n + \frac{1}{2})\pi]} \sin[(2n + 1)\pi y] \right\}$$

$$\sigma_y(x, y) = 50 \left\{ 1 - 2y - \frac{8}{\pi^2} \sum_{n=0}^{\infty} \frac{1}{(2n + 1)^2} \frac{\cosh[(2n + 1)\pi(x - \frac{1}{2})]}{\cosh[(n + \frac{1}{2})\pi]} \cos[(2n + 1)\pi y] \right\}$$

Figure 6 shows the discontinuities of $\sigma_{xf}$ at the interface of the elements, when $\sigma_{xf}$ is interpolated discontinuously and $\Omega$ is divided into eight triangular elements (see Figure 2). Remarkably, without the application of SBB, the discontinuously interpolated elements (with respect to $\sigma$) respond with a behavior one order lower than the actual interpolation function, e.g., constant in case of a...
Figure 6. Discontinuities of $\sigma_x$ at the interface of the elements for the six-node triangular element, when $\delta = 0$, respectively $\delta = 100$.

Figure 7. Error $e_{\sigma_x}$ for the three-node triangular element with equal-order interpolation.

Figure 8. Error $e_{\sigma_x}$ for the six-node triangular element with equal-order interpolation.
Application of SBB method: C. M. Kootwijk and F. P. T. Baaijens

linear interpolation and linear in case of quadratic interpolation. Application of SBB improves the order of response to the actual interpolation order. This can, for instance, be seen in Figure 6.

This is less obvious for the seven-node triangular element and nine-node rectangular element, but the response is smoothed somewhat by application of SBB. The triangular element with seven-node interpolation for \( u \) and three-node interpolation for \( \sigma \), is not influenced by the SBB formulation.

In contrast with the one-dimensional case, the discontinuities at the interface of the elements are not removed by the application of the SBB method for either of the element types.

Figure 7 shows the \( L_2 \) error of \( \sigma \) for the three-node triangular element. The error \( e_{\sigma} \) decreases somewhat by the use of SBB if \( \sigma \) is interpolated discontinuously (method 1a and 1b). If \( \sigma \) is interpolated continuously (method 2), the error \( e_{\sigma} \) increases when the SBB method is applied.

The error \( e_{\sigma} \) for the six-node triangular element is shown in Figure 8. The error \( e_{\sigma} \) decreases by the use of SBB, but the convergence rate stays the same if \( \sigma \) is interpolated discontinuously, respectively, continuously.

Figure 9 shows the \( L_2 \) error of \( \sigma \) for the nine-node rectangular element. The SBB formulation reduces the error \( e_{\sigma} \), slightly, if \( \sigma \) is interpolated discontinuously (method 1). If \( \sigma \) is interpolated continuously (method 2), the error \( e_{\sigma} \) is larger for \( \delta = 100 \).

**Figure 7.** Error \( e_{\sigma} \) for the three-node triangular element with equal-order interpolation.

**Figure 8.** Error \( e_{\sigma} \) for the six-node triangular element with equal-order interpolation.

**Figure 9.** Error \( e_{\sigma} \) for the nine-node rectangular element with equal-order interpolation.

**Figure 10.** Error \( e_{\sigma} \) as a function of \( \delta \) for the three-node triangular element with equal-order interpolation.
Application of SBB method: C. M. Kootwijk and F. P. T. Baaijens

Figure 11. Errors $e_u$ and $e_{\sigma_n}$ as a function of $\delta$ for the six-node triangular element with equal-order interpolation.

Figure 12. Errors $e_u$ and $e_{\sigma_n}$ as a function of $\delta$ for the seven-node triangular element with equal-order interpolation and the triangular element with seven-node interpolation for $u_h$ and three-node interpolation for $\sigma_n$.

Figures 10–13 show the $L_2$ errors of $u_h$ and $\sigma_n$ for the various elements as a function of $\delta$, when $\Omega$ is divided into eight triangular elements or four rectangular elements. One can see from these figures that, if the SBB formulation gives an improvement in the $L_2$ errors, this improvement is obtained if $\delta \geq 0.1$.

The nine-node rectangular element (Figure 13) gives different results than the other elements. The best value of $\delta$ for this element is 0.1.

The $L_2$ errors are not influenced by the value of $\delta$ when the seven-node triangular element or the triangular element with seven-node interpolation for $u_h$ and three-node interpolation for $\sigma_n$ is used (see Figure 12).

The above results are summarized in Table 1.

<table>
<thead>
<tr>
<th>Interpolation</th>
<th>$\delta$</th>
<th>$u$</th>
<th>Optimal $\delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_iP_i$</td>
<td>+</td>
<td>-</td>
<td>$&gt; 10^{-1}$</td>
</tr>
<tr>
<td>$P_iP_i'$</td>
<td>+</td>
<td></td>
<td>$&gt; 10^{-1}$</td>
</tr>
<tr>
<td>$P_iP_i''$</td>
<td>+</td>
<td>+/−</td>
<td>$10^{-1}$</td>
</tr>
<tr>
<td>$O_iO_i'$</td>
<td>+</td>
<td>+/−</td>
<td>$10^{-1}$</td>
</tr>
<tr>
<td>$Q_iQ_i'$</td>
<td>+</td>
<td>+/−</td>
<td>$5 \times 10^{-2}$</td>
</tr>
<tr>
<td>$Q_iQ_i''$</td>
<td>+/−</td>
<td>+/−</td>
<td>$5 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

++ very good, + good, | moderate, − bad, +/− good/bad; $P_i$ linear interpolation of order $i$, $O_i$ bilinear interpolation of order $i$, $Q_i$ discontinuous interpolation, ($|$) extended quadratic interpolation.

Figures 10-13 show the $L_2$ errors of $u_h$ and $\sigma_n$ for the various elements as a function of $\delta$, when $\Omega$ is divided into eight triangular elements or four rectangular elements. One can see from these figures that, if the SBB formulation gives an improvement in the $L_2$ errors, this improvement is obtained if $\delta \geq 0.1$.

The nine-node rectangular element (Figure 13) gives different results than the other elements. The best value of $\delta$ for this element is 0.1.

The $L_2$ errors are not influenced by the value of $\delta$ when the seven-node triangular element or the triangular element with seven-node interpolation for $u_h$ and three-node interpolation for $\sigma_n$ is used (see Figure 12).

The above results are summarized in Table 1.
Figure 13. Errors $e_u$ and $e_{w_h}$ as a function of $\delta$ for the nine-node triangular element with equal-order interpolation.

Conclusions

Instabilities that occur near discontinuities in the solution of the gradient are decreased when SBB is applied. The SBB formulation makes the approximation of $\sigma_{nh}$ over an element better when $\sigma_{nh}$ is interpolated discontinuously. For some combinations of element choice and interpolation of $\sigma_{nh}$, the SBB formulation has a good influence on the convergence of the errors. Further, there exists an optimal value of $\delta$ at which the $L_2$ error of $\sigma_{nh}$ is minimal.

The superconvergence phenomenon observed in the one-dimensional case is not recovered.

A critical point in the use of SBB is the choice of the parameter $\delta$. A more mathematical approach is needed to establish these values.

References


Appendix: Matrix form of the SBB formulation for the two-dimensional problem

Let $n_{\omega}$ be the number of local nodes of $\Omega$ at which $u_h$ and $w_h$ are interpolated. Then $u_h$ and $w_h$ can be written

as

$$u_h|\Omega^e = \sum_{a=1}^{n_{\omega}} \phi_a(x) u_a$$

$$w_h|\Omega^e = \sum_{a=1}^{n_{\omega}} \phi_a(x) w_a$$

or

$$u_h|\Omega^e = \xi^T(x) u^e$$

$$w_h|\Omega^e = \xi^T(x) w^e$$

with

$$\xi^T = [\phi_1 \ldots \phi_{n_{\omega}}]$$

$$u^e = [u^e_1 \ldots u^e_{n_{\omega}}]$$

$$w^e = [w^e_1 \ldots w^e_{n_{\omega}}]$$

where $\phi_a(x)$ is the shape function associated with local node $a$. The shape functions for $\sigma_{nh}$ and $\tau_{nh}$ may be of another order than for $u_h$ and $w_h$ and are called $\psi_a(x)$

$$\sigma_{nh}|\Omega^e = \sum_{a=1}^{m_{\omega}} \phi_a(x) \sigma_{a}$$

$$\tau_{nh}|\Omega^e = \sum_{a=1}^{m_{\omega}} \phi_a(x) \tau_{a}$$

or

$$\sigma_{nh}|\Omega^e = \phi^T(x) \sigma^e$$

$$\tau_{nh}|\Omega^e = \phi^T(x) \tau^e$$

with

$$\phi^T = [\phi_1 \ldots \phi_{m_{\omega}}]$$

$$\sigma^e = [\sigma^e_1 \ldots \sigma^e_{m_{\omega}}]$$

$$\tau^e = [\tau^e_1 \ldots \tau^e_{m_{\omega}}]$$
where \( m_{\text{nd}} \) is the number of local nodes at which \( \vec{\sigma}_n \) and \( \vec{\tau}_n \) are interpolated. The vectors \( \vec{\sigma}_n \) and \( \vec{\tau}_n \) can also be written as

\[
\vec{\sigma}_n = \sigma_x \vec{\varepsilon}_x + \sigma_y \vec{\varepsilon}_y \\
\vec{\tau}_n = \tau_x \vec{\varepsilon}_x + \tau_y \vec{\varepsilon}_y
\]

whereby \( \sigma_x, \sigma_y, \tau_x, \) and \( \tau_y \) are interpolated as follows

\[
\begin{bmatrix} \sigma_x \\ \sigma_y \end{bmatrix} = \Phi^T \sigma^e
\]  

(A5)

\[
a_n(\vec{\tau}_n, \vec{\sigma}_n) = - \int \vec{\tau}_n \cdot \frac{\vec{\sigma}_n}{c} \Omega - \sum_{e=1}^{n_{\text{el}}} \int \delta(hr) \vec{\varepsilon}(\vec{\nabla} \cdot \vec{\tau}_n)(\vec{\nabla} \cdot \vec{\sigma}_n) d\Omega \\
= - \sum_{e=1}^{n_{\text{el}}} \int \left( \vec{\tau}_n \cdot \frac{\vec{\sigma}_n}{c} + \delta(hr) \vec{\varepsilon}(\vec{\nabla} \cdot \vec{\tau}_n)(\vec{\nabla} \cdot \vec{\sigma}_n) \right) d\Omega \\
= - \sum_{e=1}^{n_{\text{el}}} \tau^T \Lambda \sigma^e
\]  

(A7)

\[
b(\vec{\tau}_n, u_h) = \int \vec{\tau}_n \cdot \vec{\nabla} u_h d\Omega \\
- \sum_{e=1}^{n_{\text{el}}} \int \vec{\tau}_n \cdot \vec{\nabla} u_h d\Omega \\
= \sum_{e=1}^{n_{\text{el}}} \tau^T B u^e
\]  

(A8)

\[
g_n(\vec{\tau}_n) = \sum_{e=1}^{n_{\text{el}}} \int \delta(hr) \vec{\varepsilon}(\vec{\nabla} \cdot \vec{\tau}_n) f d\Omega \\
= \sum_{e=1}^{n_{\text{el}}} \int \delta(hr) \vec{\varepsilon}(\vec{\nabla} \cdot \vec{\tau}_n) f d\Omega \\
= \sum_{e=1}^{n_{\text{el}}} \tau^T G
\]  

(A9)

\[
f(w_h) = \int w_h f d\Omega + \int w_h p d\Gamma \\
= \sum_{e=1}^{n_{\text{el}}} \int w_h f d\Omega + \int w_h p d\Gamma \\
= \sum_{e=1}^{n_{\text{el}}} w^T T \Phi + \int w_h p d\Gamma
\]  

(A10)

where

\[
\Lambda_{\sigma \tau} = \begin{bmatrix} \int_{\Omega} \delta(hr)^2 \frac{d\phi_a}{dx} \frac{d\phi_b}{dx} d\Omega \\
\int_{\Omega} \delta(hr)^2 \frac{d\phi_a}{dx} \frac{d\phi_b}{dy} d\Omega \\
\int_{\Omega} \delta(hr)^2 \frac{d\phi_a}{dy} \frac{d\phi_b}{dx} d\Omega \\
\int_{\Omega} \left( \phi_a \frac{d\phi_b}{c} + \delta(hr)^2 \frac{d\phi_a}{dy} \frac{d\phi_b}{dy} \right) d\Omega \end{bmatrix}
\]  

(A11)
Now, the equations of (SBB) can be written as

\[
\begin{align*}
B_{ek}^e &= \int \frac{d\varphi_e}{dx} \varphi_k d\Omega + \int \frac{d\varphi_e}{dy} \varphi_k d\Omega \quad (A12) \\
C_{ek}^e &= \int \gamma(h)^e \frac{d\varphi_e}{dx} f d\Omega \\
F_e^e &= \int w_e f d\Omega \quad (A14)
\end{align*}
\]

\[
G_{ek}^e = \int \gamma(h)^e \frac{d\varphi_e}{dy} f d\Omega
\]

Now, the equations of (SBB) can be written as

\[
-\sum_{e=1}^{n_c} \tau^T A^e e \sigma^e - \sum_{e=1}^{n_c} \tau^T B^e u^e = \sum_{e=1}^{n_c} \tau^T G^e \sigma^e
\]

\[
\sum_{e=1}^{n_c} w^T B^e e = \sum_{e=1}^{n_c} w^T F^e + \int_{\Gamma_p} w h p d\Gamma
\]

If \( u, \), \( \omega, \), \( \sigma, \) and \( \tau, \) are the global counterparts of the local \( u^e, \), \( \omega^e, \), \( \sigma^e, \) and \( \tau^e, \) respectively

\[
\begin{align*}
\bar{u} &= [u_1 \cdots u_n] \\
\bar{w} &= [w_1 \cdots w_m] \\
\bar{\sigma} &= [\sigma_{x_1} \sigma_{y_1} \cdots \sigma_{x_m} \sigma_{y_m}] \\
\bar{\tau} &= [\tau_{x_1} \tau_{y_1} \cdots \tau_{x_m} \tau_{y_m}]
\end{align*}
\]

where \( n \) is the total number of the \( u \) nodes and \( m \) is the total number of the \( \sigma \) nodes in the assembly, equations (A17) and (A18) can be assembled to

\[
-\tau^T A_\delta \sigma + \bar{\tau} B^T u = \bar{\tau} G_\delta
\]

\[
\bar{w}^T B \bar{\sigma} = \bar{w}^T F + \int_{\Gamma_p} w h p d\Gamma
\]

\( \int_{\Gamma_2} w p d\Gamma \) is a natural boundary condition that can be incorporated in \( \bar{w}^T F \). Equations (A17) and (A18) must hold for all admissible \( \bar{w} \) and \( \tau \), hence

\[
-\Delta_\delta \sigma + \bar{B}^T u = \bar{G}_\delta
\]

\[
\bar{B} \sigma = \bar{F}
\]