

**REMARKS ON MIXED FINITE
ELEMENT METHODS FOR PROBLEMS
WITH ROUGH COEFFICIENTS**

RICHARD S. FALK AND JOHN E. OSBORN

ABSTRACT. This paper considers the finite element approximation of elliptic boundary value problems in divergence form with rough coefficients. The solution of such problems will, in general, be rough, and it is well-known that the usual (Ritz or displacement) finite element method will be inaccurate in general. The purpose of the paper is to help clarify the issue of whether the use of mixed variational principles leads to finite element schemes, i.e., to mixed methods, that are more accurate than the Ritz or displacement method for such problems. For one dimensional problems, it is well-known that certain mixed methods are more accurate and robust than the Ritz method for problems with rough coefficients. Our results for two dimensional problems are mostly of a negative character. Through an examination of examples, we show that certain standard mixed methods fail to provide accurate approximations for problems with rough coefficients except in some special situations.

1. Introduction.

This paper is concerned with the finite element approximation of elliptic boundary value problems in divergence form with rough coefficients. The solutions of such problems will in general be rough and it is well known that the usual (Ritz or displacement) finite element method based on piecewise linear approximating functions is inaccurate in general. The purpose of the paper is to help clarify the issue of whether the use of mixed variational principles leads to finite element schemes, i.e., to mixed methods, that are more accurate than the Ritz

1991 *Mathematics Subject Classification.* 65N30.

Key words and phrases. finite elements, mixed methods.

The work of the first author was supported by National Science Foundation Grant DMS-9106051.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$ - $\mathcal{T}\mathcal{E}\mathcal{X}$

or displacement method for such problems. Mixed variational principles arise naturally in the mathematical formulation of many physical problems. For example, the laws of linear elasticity may be described in terms of a displacement variational formulation, involving only displacements, or in terms of a mixed variational formulation, involving both stresses and displacements, whose equations express the stress-strain relation and the balance of forces. In purely mathematical terms, one can obtain a mixed formulation from a displacement formulation by introducing new variables for some of the derivatives or certain linear combinations of derivatives of the unknown function. A mixed approximation method is obtained by basing a finite element method on the mixed variational formulation incorporating both the original and new variables. A difficulty in making precise and incisive statements about whether such methods are more accurate is that one can construct a variety of reasonable mixed methods.

For the one dimensional problem, it is well known that certain mixed methods are more accurate and robust than the Ritz method for problems with rough coefficients (cf. Babuška and Osborn [2]). In the case of two dimensions, rigorous results of this type do not exist except in special cases, and generally the situation is much less clear. Despite this fact, one dimensional results are sometimes used to justify the use of mixed methods for two dimensional problems with rough coefficients.

The results of this paper are mostly of a negative nature. Through an examination of examples, we show that certain standard mixed methods fail to give good approximations for problems with rough coefficients, unless the new mixed variable has some added regularity over what might be expected from merely differentiating the original variable. Unlike the situation for the one dimensional problem, this will not generally be the case in two dimensions. When this added regularity does occur, we are able to show that certain mixed methods do provide accurate and robust approximations. Thus, our analysis shows the importance of regularity in assessing when a mixed method will be effective.

To make our discussion precise, we consider a model boundary value problem and several specific mixed methods that have been discussed in the literature. The description of the boundary value problem is presented in § 2 along with an abstract error estimate and a regularity result due to Bernstein that will play a key role in

the subsequent discussion. In § 3, we present a brief review of what has been established about mixed methods for problems with rough coefficients in the one dimensional case. In § 4, the main section of the paper, we describe the mixed methods we will consider and then show that for two dimensional problems, the use of mixed methods cannot be expected to provide improved accuracy for problems with rough coefficients except in some special circumstances. In § 5, we make some comments about the use of special finite element methods, i.e., finite element methods using special approximating functions that depend on the coefficients, in the context of mixed methods.

In the course of our discussion, we shall also consider the relation of mixed methods to generalized displacement methods and whether the use of the piecewise harmonic average of the coefficients leads to improved accuracy (as compared to the use of the ordinary piecewise average) in the approximation of two dimensional problems.

For a comprehensive discussion of mixed finite element methods, the reader is advised to consult the book of Brezzi and Fortin [5].

2. Preliminaries.

We shall consider in this paper the approximation of the model problem

$$(2.1) \quad -\operatorname{div} A \mathbf{grad} u = f \quad \text{in } \Omega,$$

$$(2.2) \quad u = 0 \quad \text{on } \partial\Omega,$$

where Ω is a bounded domain in \mathbb{R}^2 and $A = (a_{ij}(x, y))$ is a symmetric matrix with entries $\in L^\infty(\Omega)$ that satisfies

$$(2.3) \quad \nu \sum_{i=1}^2 \xi_i^2 \leq \sum_{i,j=1}^2 a_{ij}(x, y) \xi_i \xi_j \leq \mu \sum_{i=1}^2 \xi_i^2, \quad \forall (x, y) \in \Omega, \forall \xi \in \mathbb{R}^2.$$

The class of mixed finite element formulations we shall consider are based on variational formulations of (2.1)-(2.2) that fit into the following abstract framework. Let V , W , and H be three real Hilbert spaces with norms $\|\cdot\|_V$, $\|\cdot\|_W$, and $\|\cdot\|_H$, respectively, and assume $V \subset H$ with $\|\tau\|_H \leq K\|\tau\|_V$ for all $\tau \in V$. Let $a(\cdot, \cdot)$ and $b(\cdot, \cdot)$ be continuous bilinear forms on $H \times H$ and $V \times W$, respectively:

$$\begin{aligned} |a(\boldsymbol{\sigma}, \boldsymbol{\tau})| &\leq \|a\| \|\boldsymbol{\sigma}\|_H \|\boldsymbol{\tau}\|_H && \text{for all } \boldsymbol{\sigma}, \boldsymbol{\tau} \in H, \\ |b(\boldsymbol{\sigma}, u)| &\leq \|b\| \|\boldsymbol{\sigma}\|_V \|u\|_W && \text{for all } \boldsymbol{\sigma} \in V, u \in W. \end{aligned}$$

We then consider the following variational problem:

Problem P: Given $f \in W'$, find $(\boldsymbol{\sigma}, u) \in V \times W$ satisfying

$$(2.4) \quad a(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, u) = 0 \quad \text{for all } \boldsymbol{\tau} \in V,$$

$$(2.5) \quad b(\boldsymbol{\sigma}, v) = f(v) \quad \text{for all } v \in W.$$

In order to obtain a mixed finite element method for (2.4)-(2.5), we suppose we are given finite element spaces $V_h \subset V$ and $W_h \subset W$, and then consider the following approximate problem:

Problem P_h: Find $(\boldsymbol{\sigma}_h, u_h) \in V_h \times W_h$ satisfying

$$a(\boldsymbol{\sigma}_h, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, u_h) = 0 \quad \text{for all } \boldsymbol{\tau} \in V_h,$$

$$b(\boldsymbol{\sigma}_h, v) = f(v) \quad \text{for all } v \in W_h.$$

As will be seen from the examples in the following sections, many standard mixed finite element methods for the approximation of (2.1)-(2.2) fit this framework. Although many of our results will be of a negative nature, there are some cases in which positive results for the approximation of problems with rough coefficients can be obtained. In these cases, our error analysis will require several additional assumptions, which we now state. For all h ,

$$(2.6) \quad Z_h \subseteq Z,$$

where

$$Z = \{\boldsymbol{\tau} \in V : b(\boldsymbol{\tau}, v) = 0 \quad \text{for all } v \in W\},$$

$$Z_h = \{\boldsymbol{\tau} \in V_h : b(\boldsymbol{\tau}, v) = 0 \quad \text{for all } v \in W_h\}.$$

There is a constant $\alpha > 0$, independent of h , such that

$$(2.7) \quad a(\boldsymbol{\tau}, \boldsymbol{\tau}) \geq \alpha \|\boldsymbol{\tau}\|_H^2 \quad \text{for all } \boldsymbol{\tau} \in Z_h.$$

There is a constant $\beta > 0$, independent of h , such that for all $v \in W_h$,

$$(2.8) \quad \sup_{\boldsymbol{\tau} \in V_h} \frac{b(\boldsymbol{\tau}, v)}{\|\boldsymbol{\tau}\|_V} \geq \beta \|v\|_W.$$

We note that if (2.8) holds (cf. [6, Proposition 1]), then there is an operator $\pi_h : V \rightarrow V_h$ satisfying

$$(2.9) \quad b(\boldsymbol{\tau} - \pi_h \boldsymbol{\tau}, v) = 0 \quad \text{for all } \boldsymbol{\tau} \in V, \quad v \in W_h.$$

If in addition, (2.6) holds, (cf. [6, Proposition 3]), then there is an operator $\Sigma_h : W \rightarrow W_h$ satisfying

$$(2.10) \quad b(\boldsymbol{\tau}, v - \Sigma_h v) = 0 \quad \text{for all } \boldsymbol{\tau} \in V_h, \quad v \in W.$$

The following abstract error estimates follow easily when all the above assumptions are valid. The results are special cases of a more general theory to be found in [5] and [6].

Theorem 2.1. *If $(\boldsymbol{\sigma}, u)$ and $(\boldsymbol{\sigma}_h, u_h)$ are the solutions of Problems P and P_h respectively and hypotheses (2.6), (2.7), and (2.8) hold, then*

$$(2.11) \quad \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_H \leq \left(1 + \frac{\|a\|}{\alpha}\right) \|\boldsymbol{\sigma} - \pi_h \boldsymbol{\sigma}\|_H,$$

$$(2.12) \quad \|\Sigma_h u - u_h\|_W \leq \frac{K\|a\|}{\beta} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_H,$$

where π_h and Σ_h are any operators satisfying (2.9) and (2.10), respectively.

Proof. It follows easily from the definitions that for all $\boldsymbol{\tau} \in V_h$,

$$a(\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\tau}) = a(\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}, \boldsymbol{\tau}) + b(\boldsymbol{\tau}, u_h - u),$$

and letting $\boldsymbol{\tau} = \pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h$, we get

$$a(\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h) = a(\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}, \pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h) + b(\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, u_h - u).$$

Also,

$$b(\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, u_h - u) = b(\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, u_h - \Sigma_h u) = b(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, u_h - \Sigma_h u) = 0,$$

and so

$$a(\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h) = a(\pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}, \pi_h \boldsymbol{\sigma} - \boldsymbol{\sigma}_h).$$

Estimate (2.11) now follows from (2.7) and the triangle inequality. To obtain the second result, we apply (2.8) to get

$$\beta \|\Sigma_h u - u_h\|_W \leq \sup_{\boldsymbol{\tau} \in V_h} \frac{|b(\boldsymbol{\tau}, \Sigma_h u - u_h)|}{\|\boldsymbol{\tau}\|_V}.$$

It again follows easily from the definitions and our assumptions that

$$\begin{aligned} |b(\boldsymbol{\tau}, \Sigma_h u - u_h)| &= |b(\boldsymbol{\tau}, u - u_h)| \\ &= |a(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}, \boldsymbol{\tau})| \leq \|a\| \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_H \|\boldsymbol{\tau}\|_H. \end{aligned}$$

Estimate (2.12) now follows immediately.

Note that although (2.12) is a somewhat crude estimate for general problems, it will enable us to obtain error estimates for some special classes of problems with rough coefficients. Also observe that to obtain positive rates of convergence from this theorem, we will need to know that $\boldsymbol{\sigma}$ is smoother than implied by inclusion in the function space H .

Since smoothness of the solution will play a key role in our discussion of the effectiveness of mixed methods for the approximation of problems with rough coefficients, we shall frequently rely on the following theorem of Bernstein [4], [8, Section 3.17] on the regularity of solutions of elliptic equations in nondivergence form in two dimensions. We now state this theorem, which will be used in § 4 in our discussion of mixed methods.

Consider the problem

$$\begin{aligned} (2.13) \quad & -a_{11}(x, y) \frac{\partial^2 u}{\partial x^2} - 2a_{12}(x, y) \frac{\partial^2 u}{\partial x \partial y} - a_{22}(x, y) \frac{\partial^2 u}{\partial y^2} = f \quad \text{in } \Omega, \\ (2.14) \quad & u = 0 \quad \text{on } \partial\Omega, \end{aligned}$$

where Ω is a bounded, convex domain in \mathbb{R}^2 with a piecewise C^2 boundary $\partial\Omega$, and where the functions $a_{ij} \in L^\infty(\Omega)$ satisfy

$$(2.15) \quad \nu \sum_{i=1}^2 \xi_i^2 \leq \sum_{i,j=1}^2 a_{ij}(x, y) \xi_i \xi_j \leq \mu \sum_{i=1}^2 \xi_i^2 \quad \forall (x, y) \in \Omega, \forall \boldsymbol{\xi} \in \mathbb{R}^2,$$

with $a_{21} = a_{12}$ and where ν and μ are positive constants.

Theorem 2.2. (Bernstein) *For each $f \in L^2(\Omega)$, Problem (2.13)-(2.14) has a unique solution $u \in H^2(\Omega) \cap H_0^1(\Omega)$. Furthermore there is a constant $C = C(\nu, \mu)$, depending on ν and μ but independent of f , such that*

$$\|u\|_{H^2(\Omega)} \leq C \|f\|_{L^2(\Omega)}.$$

Our hypothesis on Ω is not identical to the one in [8]. To prove that the result is valid for a domain of the type we are considering, one can use the a priori estimates in [7, Section 3.1]. The Bernstein Theorem says that nondivergence form equations have solutions in $H^2(\Omega)$ even though the coefficients are very rough.

3. A survey of results for the one dimensional problem.

In this section we discuss the application of mixed methods to the approximate solution of the one dimensional version of problem (2.1)-(2.2), namely

$$(3.1) \quad -(a(x)u'(x))' = f(x) \quad 0 < x < 1,$$

$$(3.2) \quad u(0) = 0, \quad u(1) = 0,$$

where $a(x) \in L^\infty(0, 1)$ and satisfies

$$0 < \nu \leq a(x) \leq \mu.$$

We shall consider three variational formulations of (3.1)-(3.2) that fit the abstract framework of the previous section. These are:

I. Find $\sigma \in L^2(0, 1)$, $u \in H_0^1(0, 1)$ satisfying

$$\begin{aligned} - \int_0^1 a\sigma v' dx &= - \int_0^1 f v dx && \text{for all } v \in H_0^1(0, 1), \\ \int_0^1 a\sigma \tau dx - \int_0^1 a u' \tau dx &= 0 && \text{for all } \tau \in L^2(0, 1). \end{aligned}$$

II. Find $\sigma \in L^2(0, 1)$, $u \in H_0^1(0, 1)$ satisfying

$$\begin{aligned} - \int_0^1 \sigma v' dx &= - \int_0^1 f v dx && \text{for all } v \in H_0^1(0, 1), \\ \int_0^1 \frac{\sigma \tau}{a} dx - \int_0^1 u' \tau dx &= 0 && \text{for all } \tau \in L^2(0, 1). \end{aligned}$$

III. Find $\sigma \in H^1(0, 1)$, $u \in L^2(0, 1)$ satisfying

$$\begin{aligned} \int_0^1 \sigma' v dx &= - \int_0^1 f v dx && \text{for all } v \in L^2(\Omega), \\ \int_0^1 \frac{\sigma \tau}{a} dx + \int_0^1 u \tau' dx &= 0 && \text{for all } \tau \in H^1(0, 1). \end{aligned}$$

These are standard mixed formulations, which arise by introducing the new variable $\sigma = u'$ (in I) or $\sigma = au'$ (in II and III).

We then consider certain simple natural finite element discretizations of these problems. Let $\mathcal{T}_h = \{0 = x_0 < x_1 < \dots < x_n = 1\}$ be a mesh on $I = [0, 1]$ satisfying $x_j - x_{j-1} \leq h$ for $j = 1, \dots, n$, and let $I_j = (x_{j-1}, x_j)$ and $h_j = x_j - x_{j-1}$. We discretize formulations I and II by choosing

$$\begin{aligned} V_h &= \{\tau : \tau|_{I_j} = \text{constant for } j = 1, \dots, n\}, \\ W_h &= \{u \in H_0^1(0, 1) : u|_{I_j} = \text{linear for } j = 1, \dots, n\}. \end{aligned}$$

This leads to the approximate problems:

I_h. Find $\sigma_h \in V_h$, $u_h \in W_h$ satisfying

$$\begin{aligned} - \int_0^1 a \sigma_h v' dx &= - \int_0^1 f v dx && \text{for all } v \in W_h, \\ \int_0^1 a \sigma_h \tau dx - \int_0^1 a u_h' \tau dx &= 0 && \text{for all } \tau \in V_h. \end{aligned}$$

II_h. Find $\sigma_h \in V_h$, $u_h \in W_h$ satisfying

$$\begin{aligned} - \int_0^1 \sigma_h v' dx &= - \int_0^1 f v dx && \text{for all } v \in W_h, \\ \int_0^1 \frac{\sigma_h \tau}{a} dx - \int_0^1 u_h' \tau dx &= 0 && \text{for all } \tau \in V_h. \end{aligned}$$

We discretize formulation III by choosing

$$\begin{aligned} V_h &= \{\tau \in H^1(0, 1) : \tau|_{I_j} = \text{linear for } j = 1, \dots, n\}, \\ W_h &= \{v \in L^2(0, 1) : v|_{I_j} = \text{constant for } j = 1, \dots, n\}. \end{aligned}$$

This leads to the approximate problem:

III_h. Find $\sigma_h \in V_h$, $u \in W_h$ satisfying

$$\begin{aligned} \int_0^1 \sigma_h' v dx &= - \int_0^1 f v dx && \text{for all } v \in W_h, \\ \int_0^1 \frac{\sigma_h \tau}{a} dx + \int_0^1 u_h \tau' dx &= 0 && \text{for all } \tau \in V_h. \end{aligned}$$

To analyze the approximation I_h , we observe from the second equation that

$$\sigma_h = u'_h.$$

Substituting u'_h for σ_h in the first equation, we find that $u_h \in W_h$ satisfies

$$(3.3) \quad \int_0^1 a u'_h v' dx = \int_0^1 f v dx \quad \text{for all } v \in W_h.$$

Thus u_h is the usual Ritz approximation to u using continuous piecewise linear approximating functions. As mentioned in the Introduction, the Ritz method does not in general provide good approximations when $a(x)$ is rough, and hence (σ_h, u_h) is not a good approximation to (σ, u) . We will comment further on this point below. For later use, let us note that (3.3) can be written as:

$$\int_0^1 a^h u'_h v' dx = \int_0^1 f v dx \quad \text{for all } v \in W_h,$$

where $a^h(x)$ is the piecewise average of $a(x)$, i.e.,

$$a^h |_{I_j} = h_j^{-1} \int_{I_j} a dx.$$

To analyze approximation Π_h , we may use Theorem 2.1 with $V = H = L^2(\Omega)$, $W = H_0^1(\Omega)$,

$$a(\sigma, \tau) = \int_0^1 \frac{1}{a} \sigma \tau dx, \quad b(\sigma, v) = - \int_0^1 \sigma v' dx,$$

and V_h and W_h as defined above. We obtain

$$(3.4) \quad \|\sigma - \sigma_h\|_{L^2(0,1)} \leq C(\nu, \mu) \|\sigma - \pi_h \sigma\|_{L^2(0,1)},$$

$$(3.5) \quad \|(\Sigma_h u - u_h)'\|_{L^2(0,1)} \leq C(\nu, \mu) \|\sigma - \sigma_h\|_{L^2(0,1)},$$

where ν and μ are the lower and upper bounds for $a(x)$. The key ingredients in the application of Theorem 2.1 are to show that $Z_h = Z$ is the set of global constants and that (2.7) and (2.8) are satisfied. We also see that we may take π_h to be the L^2 projection and Σ_h to be the piecewise linear interpolant.

Recall that $\sigma = au'$, so $\sigma' = -f$, showing that σ is smoother than implied by $\sigma \in H = L^2(\Omega)$. Thus from (3.4) we have

$$(3.6) \quad \|\sigma - \sigma_h\|_{L^2(0,1)} \leq C(\nu, \mu)h\|f\|_{L^2(0,1)}.$$

Since $u \notin H^2(0,1)$ we cannot use the triangle inequality, (3.5), and (3.6) to get an $O(h)$ estimate for $\|u - u_h\|_{H^1(0,1)}$. However, since

$$\|\Sigma_h u - u_h\|_{L^2(0,1)} \leq C\|(\Sigma_h u - u_h)'\|_{L^2(0,1)},$$

it easily follows from the triangle inequality and (3.6) that

$$(3.7) \quad \|u - u_h\|_{L^2(0,1)} \leq C(\nu, \mu)h\|f\|_{L^2(0,1)}.$$

Estimates (3.6) and (3.7) show that method II_h is accurate and robust for one dimensional problems with rough coefficients.

By eliminating σ_h from Problem I_h , we saw that u_h was the Ritz approximation to u and $\sigma_h = u_h'$. In a similar way, we can eliminate σ_h from Problem II_h . Using the second equation, we find that

$$\sigma_h|_{I_j} = u_h' \left(\frac{1}{h_j} \int_{I_j} \frac{dx}{a(x)} \right)^{-1},$$

and substituting this in the first equation, we find that $u_h \in W_h$ satisfies

$$(3.8) \quad \int_0^1 a_h u_h' v' dx = \int_0^1 f v dx \quad \text{for all } v \in W_h,$$

where $a_h(x)$ is the piecewise harmonic average of $a(x)$, i.e.,

$$a_h|_{I_j} = \left(\frac{1}{h_j} \int_{I_j} \frac{dx}{a(x)} \right)^{-1}.$$

Problem (3.8) is called a Generalized Displacement Method (for the determination of u_h) (cf. [2]).

As mentioned above, we will now comment further on the fact that the Ritz approximation is inaccurate. Let N be a large even integer and suppose

$$a(x) = \begin{cases} 2, & \frac{k-1}{N} < x < \frac{k}{N}, \quad k = 1, 3, \dots, N-1 \\ 1, & \frac{k-1}{N} < x < \frac{k}{N}, \quad k = 2, 4, \dots, N, \end{cases}$$

and consider Problem (3.1)-(3.2) with this $a(x)$. Let (σ_h, u_h) be the approximate solution defined by Problem II_h and temporarily let $(\tilde{\sigma}_h, \tilde{u}_h)$ be the approximate solution defined by Problem I_h . We know that u_h is characterized by (3.8) and $\sigma_h = a_h u_h'$, and that \tilde{u}_h is characterized by (3.3) (and is thus the Ritz approximation to u) and $\tilde{\sigma}_h = \tilde{u}_h'$. Suppose h is an even multiple of N^{-1} . Then $a^h(x) = 3/2$ and $a_h(x) = 4/3$. Thus \tilde{u}_h is also the Ritz approximation to the problem:

$$-\frac{3}{2}\tilde{w}'' = f \quad 0 < x < 1, \quad \tilde{w}(0) = 0, \quad \tilde{w}(1) = 0,$$

and u_h is the Ritz approximation to the problem:

$$-\frac{4}{3}w'' = f \quad 0 < x < 1, \quad w(0) = 0, \quad w(1) = 0.$$

Now we know from (3.7) that u_h is a very accurate approximation to u , and from standard error estimates for the Ritz method that \tilde{u}_h is a very accurate approximation to w and that u is a very accurate approximation to \tilde{w} . However, it is easy to see that w is not close to \tilde{w} (since they solve different differential equations). Hence, we conclude that \tilde{u}_h is not a good approximation to u . In fact \tilde{u}_h stays away from u (and close to \tilde{w}) as long as h is an even multiple of N^{-1} . As soon as h is a fraction of N^{-1} (i.e., $h = N^{-1}k^{-1}$, $k = 1, 2, \dots$), the Ritz approximation \tilde{u}_h gets close to u . However, if N is very large, we may not be able to take h as small as N^{-1} in practical computations.

To analyze the approximation III_h , we may again apply Theorem 2.1, obtaining

$$\begin{aligned} \|\sigma - \sigma_h\|_{L^2(0,1)} &\leq C(\nu, \mu)h\|f\|_{L^2(0,1)}, \\ \|u - u_h\|_{L^2(0,1)} &\leq C(\nu, \mu)h\|f\|_{L^2(0,1)}. \end{aligned}$$

Standard inverse estimates then imply for quasi-uniform meshes that

$$\|\sigma - \sigma_h\|_{H^1(0,1)} \leq C(\nu, \mu)\|f\|_{L^2(0,1)}.$$

The key ingredients in the proof are again to show that $Z_h = Z$ is the set of global constants, and that we may now take π_h to be the piecewise linear interpolant and Σ_h to be the L^2 projection. Once again hypotheses (2.7) and (2.8) are easily verified.

4. Two dimensional results.

We begin this section by considering a variational formulation of Problem (2.1)-(2.2) that is the two dimensional analogue of formulation I of § 3, and which arises by letting $\boldsymbol{\sigma} = \mathbf{grad} u$.

I. Find $\boldsymbol{\sigma} \in \mathbf{L}^2(\Omega)$, $u \in H_0^1(\Omega)$ satisfying

$$\begin{aligned} - \int_{\Omega} A \boldsymbol{\sigma} \cdot \mathbf{grad} v \, dx \, dy &= - \int_{\Omega} f v \, dx \, dy \quad \text{for all } v \in H_0^1(\Omega), \\ \int_{\Omega} A \boldsymbol{\sigma} \cdot \boldsymbol{\tau} \, dx \, dy - \int_{\Omega} A \mathbf{grad} u \cdot \boldsymbol{\tau} \, dx \, dy &= 0 \quad \text{for all } \boldsymbol{\tau} \in \mathbf{L}^2(\Omega). \end{aligned}$$

To obtain a simple finite element discretization, we let \mathcal{T}_h , $0 < h < 1$, be a triangulation of Ω with triangles T of diameter less than or equal to h , and assume that $\{\mathcal{T}_h\}$ satisfies the minimal angle condition. We then discretize Problem I according to the framework described in § 2 by choosing

$$\begin{aligned} V_h &= \{ \boldsymbol{\tau} \in \mathbf{L}^2(\Omega) : \boldsymbol{\tau}|_T = \text{constant for all } T \in \mathcal{T}_h \}, \\ W_h &= \{ v \in H_0^1(\Omega) : v|_T = \text{linear for all } T \in \mathcal{T}_h \}. \end{aligned}$$

This leads to the approximate problem:

I_h. Find $\boldsymbol{\sigma}_h \in V_h$, $u_h \in W_h$ satisfying

$$\begin{aligned} - \int_{\Omega} A \boldsymbol{\sigma}_h \cdot \mathbf{grad} v \, dx \, dy &= - \int_{\Omega} f v \, dx \, dy \quad \text{for all } v \in W_h, \\ \int_{\Omega} A \boldsymbol{\sigma}_h \cdot \boldsymbol{\tau} \, dx \, dy - \int_{\Omega} A \mathbf{grad} u_h \cdot \boldsymbol{\tau} \, dx \, dy &= 0 \quad \text{for all } \boldsymbol{\tau} \in V_h. \end{aligned}$$

As in the one dimensional case, we immediately see that $u_h \in W_h$ satisfies:

$$\int_{\Omega} A^h \mathbf{grad} u_h \cdot \mathbf{grad} v \, dx \, dy = \int_{\Omega} f v \, dx \, dy \quad \text{for all } v \in W_h,$$

and

$$\boldsymbol{\sigma}_h = \mathbf{grad} u_h,$$

where A^h is the piecewise average of A , i.e.,

$$A^h|_T = \frac{1}{|T|} \int_T A \, dx \, dy \quad \text{for all } T \in \mathcal{T}_h.$$

Thus u_h is the usual Ritz approximation to u based on continuous, piecewise linear approximating functions.

As mentioned above, and as elaborated on for the one-dimensional case in § 3, the Ritz method does not in general provide good approximations to problems with rough coefficients. Thus, method I_h cannot generally be expected to provide good approximations for problems with rough coefficients. However, this does not mean that it works poorly on all problems with rough coefficients.

To see this, consider the application of method I_h to the boundary value problem:

$$(4.1) \quad -a_{11}(y) \frac{\partial^2 u}{\partial x^2} - 2a_{12} \frac{\partial^2 u}{\partial x \partial y} - a_{22}(x) \frac{\partial^2 u}{\partial y^2} = f \quad \text{in } \Omega \equiv (0, 1) \times (0, 1),$$

$$(4.2) \quad u = 0 \quad \text{on } \partial\Omega,$$

where a_{11} depends only on y , a_{12} is constant, a_{22} depends only on x , and a_{ij} satisfies (2.3) (or (2.15)). Although the equation (4.1) is in nondivergence form (so we can apply the Bernstein Theorem), in this special case it can also be written in the divergence form of equation (2.1). Hence, method I_h and the Ritz method are applicable. Using the Bernstein result and the usual analysis of the Ritz method we find that

$$(4.3) \quad \|u - u_h\|_{L^2(\Omega)} + h\|u - u_h\|_{H^1(\Omega)} \leq Ch^2\|f\|_{L^2(\Omega)},$$

where C depends only on ν and μ in (2.3). In the context of the mixed method I_h , (4.3) becomes:

$$\|u - u_h\|_{H^1(\Omega)} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega)} \leq Ch\|f\|_{L^2(\Omega)}.$$

Hence, we have an example of a mixed method that provides accurate approximations for some problems with rough coefficients, but not others. Of course our mixed method is equivalent to a displacement method that provides similar approximations.

Next, consider the two dimensional analogue of formulation II of § 3, which arises by letting $\boldsymbol{\sigma} = A \mathbf{grad} u$.

II. Find $\boldsymbol{\sigma} \in \mathbf{L}^2(\Omega)$, $u \in H_0^1(\Omega)$ satisfying

$$-\int_{\Omega} \boldsymbol{\sigma} \cdot \mathbf{grad} v \, dx \, dy = -\int_{\Omega} f v \, dx \, dy \quad \text{for all } v \in H_0^1(\Omega),$$

$$\int_{\Omega} A^{-1} \boldsymbol{\sigma} \cdot \boldsymbol{\tau} \, dx \, dy - \int_{\Omega} \mathbf{grad} u \cdot \boldsymbol{\tau} \, dx \, dy = 0 \quad \text{for all } \boldsymbol{\tau} \in \mathbf{L}^2(\Omega).$$

Choosing V_h and W_h as in method I_h leads to the approximate problem:

II_h . Find $\boldsymbol{\sigma}_h \in V_h$, $u_h \in W_h$ satisfying

$$\begin{aligned} - \int_{\Omega} \boldsymbol{\sigma}_h \cdot \mathbf{grad} v \, dx \, dy &= - \int_{\Omega} f v \, dx \, dy && \text{for all } v \in W_h, \\ \int_{\Omega} A^{-1} \boldsymbol{\sigma}_h \cdot \boldsymbol{\tau} \, dx \, dy - \int_{\Omega} \mathbf{grad} u_h \cdot \boldsymbol{\tau} \, dx \, dy &= 0 && \text{for all } \boldsymbol{\tau} \in V_h. \end{aligned}$$

In § 3, we saw that this method is accurate and robust for problems with rough coefficients in one dimension. We will now see that the situation is very different in the two dimensional case. We first note that, as in the one dimensional case, $u_h \in W_h$ satisfies:

$$(4.4) \quad \int_{\Omega} A_h \mathbf{grad} u_h \cdot \mathbf{grad} v \, dx \, dy = \int_{\Omega} f v \, dx \, dy \quad \text{for all } v \in W_h,$$

and

$$\boldsymbol{\sigma}_h = A_h \mathbf{grad} u_h,$$

where A_h is the piecewise harmonic average of A , i.e.,

$$A_h|_T = \left(\frac{1}{|T|} \int_T A^{-1} \, dx \, dy \right)^{-1} \quad \text{for all } T \in \mathcal{T}_h.$$

As mentioned in § 3, (4.4) is referred to as a Generalized Displacement Method.

Consider the application of method II_h to the boundary value problem:

$$(4.5) \quad - \frac{\partial}{\partial x} \left(a(x) \frac{\partial u}{\partial x} \right) - \frac{\partial}{\partial y} \left(a(x) \frac{\partial u}{\partial y} \right) = f \quad \text{in } \Omega \equiv (0, 1) \times (0, 1),$$

$$(4.6) \quad u = 0 \quad \text{on } \partial\Omega,$$

where $a = a(x)$ is a function of x only. In variational formulation II , $\boldsymbol{\sigma} = A \mathbf{grad} u$, so in the case of (4.5), we see that $\boldsymbol{\sigma} = a(x) \mathbf{grad} u$. We now show that $\boldsymbol{\sigma}$ cannot be smooth if $a(x)$ is rough.

Toward this end, we introduce the change of variables

$$\bar{x} = \int_0^x \frac{1}{a} \, ds, \quad \bar{y} = y.$$

Setting $\bar{u}(\bar{x}, \bar{y}) = u(x, y)$, the boundary value problem (4.5)-(4.6) becomes:

$$(4.7) \quad -\frac{\partial^2 \bar{u}}{\partial \bar{x}^2} - \bar{a}^2 \frac{\partial^2 \bar{u}}{\partial \bar{y}^2} = \bar{a} \bar{f} \quad \text{in } \bar{\Omega},$$

$$(4.8) \quad \bar{u} = 0 \quad \text{on } \partial \bar{\Omega},$$

where $\bar{\Omega} = (0, \int_0^1 a^{-1} ds) \times (0, 1)$. Since (4.7) is in nondivergence form, it follows from the Bernstein result that $\bar{u}(\bar{x}, \bar{y}) \in H^2(\bar{\Omega})$ and hence that $\bar{u}_{\bar{x}} = au_x$, $\bar{u}_{\bar{y}} = u_y \in H^1(\Omega)$. Thus, au_y and hence $\boldsymbol{\sigma} = (au_x, au_y)$ cannot be smooth. A specific example illustrating this is obtained by letting

$$f(x, y) = -\frac{2}{a(x)}y(y-1) - 2a(x) \int_0^x \frac{ds}{a} \left[\int_0^1 \frac{ds}{a} - \int_0^x \frac{ds}{a} \right].$$

With this f ,

$$u(x, y) = \int_0^x \frac{ds}{a} \left[\int_0^1 \frac{ds}{a} - \int_0^x \frac{ds}{a} \right] y(1-y),$$

and we see that

$$\begin{aligned} \boldsymbol{\sigma} = a \mathbf{grad} u = & \left(\left[\int_0^1 \frac{ds}{a} - 2 \int_0^x \frac{ds}{a} \right] y(1-y), \right. \\ & \left. a(x) \int_0^x \frac{ds}{a} \left[\int_0^1 \frac{ds}{a} - \int_0^x \frac{ds}{a} \right] (1-2y) \right). \end{aligned}$$

In particular, σ_2 cannot be smooth if $a(x)$ is rough.

The importance of showing that $\boldsymbol{\sigma}$ is not smooth is that this implies that it cannot be accurately approximated by piecewise constant functions. Hence, $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega)}$ cannot have a good rate of convergence. Since the difference between this approximation scheme and the previous one amounts to the replacement of the piecewise average of the coefficient by the piecewise harmonic average, we also see that unlike the situation in one dimension, the use of the harmonic average in two dimensions is not sufficient to give accurate and robust approximations to problems with rough coefficients.

We now turn to a two dimensional analogue of formulation III of § 3.

III. Find $\boldsymbol{\sigma} \in \mathbf{H}(\text{div}, \Omega)$, $u \in L^2(\Omega)$ satisfying

$$\begin{aligned} \int_{\Omega} \text{div } \boldsymbol{\sigma} v \, dx \, dy &= - \int_{\Omega} f v \, dx \, dy \quad \text{for all } v \in L^2(\Omega), \\ \int_{\Omega} A^{-1} \boldsymbol{\sigma} \cdot \boldsymbol{\tau} \, dx \, dy + \int_{\Omega} u \text{div } \boldsymbol{\tau} \, dx \, dy &= 0 \quad \text{for all } \boldsymbol{\tau} \in \mathbf{H}(\text{div}, \Omega). \end{aligned}$$

Choosing V_h to be the space of lowest order Raviart-Thomas elements and W_h to be the space of piecewise constants, we obtain method III_h. First we apply method III_h to Problem (4.5)-(4.6). We have (as in formulation II) $\boldsymbol{\sigma} = A \mathbf{grad} u = a(x) \mathbf{grad} u$, and we have seen that $\boldsymbol{\sigma}$ is not smooth if $a(x)$ is rough. Since we will be approximating $\boldsymbol{\sigma}$ by a special type of discontinuous linear element, and this works well only if $\boldsymbol{\sigma}$ is smooth, $\boldsymbol{\sigma}_h$ again cannot be expected to be an accurate approximation to $\boldsymbol{\sigma}$.

It is also of interest to consider the boundary value problem:

$$\begin{aligned} -\frac{\partial^2 u}{\partial x^2} - a(x) \frac{\partial^2 u}{\partial y^2} &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

which is a special case of (4.1) - (4.2). From the Bernstein result, we see that $u \in H^2(\Omega)$ and hence $u_y \in H^1(\Omega)$. Thus on the one hand the Ritz method works well, while on the other hand, since $a(x)$ is rough, $\boldsymbol{\sigma} = A \mathbf{grad} u = (u_x, a(x)u_y)$ cannot be smooth. In particular, $\boldsymbol{\sigma} = A \mathbf{grad} u \notin \mathbf{H}^1(\Omega)$. A specific example illustrating this is obtained by letting

$$f = 2y(1-y) + 2xa(x)(1-x), \quad \Omega = (0, 1) \times (0, 1).$$

With this f , $u(x, y) = xy(1-x)(1-y)$. Thus

$$\boldsymbol{\sigma} = \left(\frac{\partial u}{\partial x}, a \frac{\partial u}{\partial y} \right) = ((2x-1)y(y-1), a(x)(2y-1)x(x-1)),$$

and we see that $\boldsymbol{\sigma}_2 = a(x)(2y-1)x(x-1)$ is clearly not smooth. Thus, we cannot approximate $\boldsymbol{\sigma}$ to order h by any element in the subspace and hence $\boldsymbol{\sigma}_h$ cannot be an order h approximation to $\boldsymbol{\sigma}$. Thus we have a problem for which a standard mixed method produces a worse

approximation to the “stress” variable $\boldsymbol{\sigma}$ than the approximation obtained by the Ritz method by forming $A \mathbf{grad} u_h$.

Despite these negative results, there is a situation in two dimensions in which some mixed methods work well and that is when $\boldsymbol{\sigma} = A \mathbf{grad} u$ is a smoother variable than might be expected, i.e., when $u \in H^1(\Omega)$, $u \notin H^2(\Omega)$, but $\boldsymbol{\sigma} \in \mathbf{H}^1(\Omega)$. In this case, we can approximate the boundary value problem (2.1)-(2.2) by method III_h using the lowest order Raviart-Thomas elements, and use Theorem 2.1 to derive the error estimates:

$$\begin{aligned} \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega)} &\leq Ch \|\boldsymbol{\sigma}\|_{\mathbf{H}^1(\Omega)}, \\ \|u - u_h\|_{L^2(\Omega)} &\leq Ch (\|u\|_{H^1(\Omega)} + \|\boldsymbol{\sigma}\|_{\mathbf{H}^1(\Omega)}). \end{aligned}$$

Note that for this method, it is well known that all our hypotheses are valid.

We have already seen how a special regularity result of this type occurs in a one dimensional problem and it is not difficult to extend examples of this type to boundary value problems, which although formally two-dimensional, have a coefficient and a solution that depend only on a single variable. For example, consider the boundary value problem

$$\begin{aligned} -\frac{\partial}{\partial x} \left(a(x) \frac{\partial u}{\partial x} \right) - \frac{\partial}{\partial y} \left(a(x) \frac{\partial u}{\partial y} \right) &= f(x) \quad \text{in } \Omega \equiv (0, 1) \times (0, 1), \\ u(0, y) = u(1, y) &= 0, \quad 0 < y < 1, \\ \frac{\partial u}{\partial y}(x, 0) = \frac{\partial u}{\partial y}(x, 1) &= 0, \quad 0 < x < 1. \end{aligned}$$

Since $a = a(x)$ and $f = f(x)$ are functions of x only, u is also a function only of x and satisfies the boundary value problem (3.1)-(3.2). Hence, the one dimensional regularity results apply and the variable $\boldsymbol{\sigma} = (au_x, au_y) = (au_x, 0)$ will be smooth.

A somewhat more interesting class of examples can be constructed in the following way. Consider the boundary value problem (2.1)-(2.2), where Ω is convex and $A = aI$. Let t be the solution of

$$-\Delta t = f \quad \text{in } \Omega, \quad t = 0 \quad \text{on } \partial\Omega.$$

By standard regularity results, if $f \in L^2(\Omega)$, then $t \in H^2(\Omega)$. Denote by T_- and T_+ the minimum and maximum values of t over $\bar{\Omega}$. Let

$\phi(t) \in W^{1,\infty}(T_-, T_+)$ and satisfy $\phi(0) = 0$, $\phi'(t) \geq k > 0$. For example, if $(T_- + T_+)/2 > 0$, set

$$\phi(t) = \begin{cases} t, & T_- \leq t \leq (T_- + T_+)/2 \\ 2t - (T_- + T_+)/2, & (T_- + T_+)/2 \leq t \leq T_+ \end{cases}.$$

Setting

$$u(x, y) = \phi(t(x, y)), \quad a(x, y) = 1/\phi'(t(x, y)),$$

we observe that $a \in L^\infty(\Omega)$, u solves the boundary value problem (2.1)-(2.2), $u \in H^1(\Omega)$, and $u \notin H^2(\Omega)$, but $\boldsymbol{\sigma} \equiv a \mathbf{grad} u = \mathbf{grad} t \in \mathbf{H}^1(\Omega)$. A simple special case in which all quantities can be made explicit is when Ω is the unit disk and $f \equiv 4$. Then $t(x, y) = 1 - x^2 - y^2$, $T_- = 0$, and $T_+ = 1$. Note that examples of this type are quite special, with a strong connection between the coefficient a and the right hand side f .

A third example for which we can prove a regularity result of this type is obtained if we consider (2.1)-(2.2) and assume $\Omega = (0, 1) \times (0, 1)$ and A is a diagonal matrix with $A_{11} = a_{11}(x)$ and $A_{22} = a_{22}(y)$. We are thus considering the boundary value problem

$$(4.9) \quad -\frac{\partial}{\partial x} \left(a_{11}(x) \frac{\partial u}{\partial x} \right) - \frac{\partial}{\partial y} \left(a_{22}(y) \frac{\partial u}{\partial y} \right) = f \quad \text{in } \Omega,$$

$$(4.10) \quad u = 0 \quad \text{on } \partial\Omega.$$

If we introduce the change of variables

$$(4.11) \quad \bar{x} = \int_0^x \frac{1}{a_{11}(s)} ds, \quad \bar{y} = \int_0^y \frac{1}{a_{22}(t)} dt,$$

and set $\bar{u}(\bar{x}, \bar{y}) = u(x, y)$, (4.9)-(4.10) becomes

$$(4.12) \quad -\bar{a}_{22}(\bar{y}) \frac{\partial^2 \bar{u}}{\partial \bar{x}^2} - \bar{a}_{11}(\bar{x}) \frac{\partial^2 \bar{u}}{\partial \bar{y}^2} = \bar{a}_{11}(\bar{x}) \bar{a}_{22}(\bar{y}) \bar{f} \quad \text{in } \bar{\Omega},$$

$$(4.13) \quad \bar{u} = 0 \quad \text{on } \partial\bar{\Omega},$$

where $\bar{\Omega} = (0, \int_0^1 a_{11}^{-1} ds) \times (0, \int_0^1 a_{22}^{-1} dt)$. Since (4.12)-(4.13) is in nondivergence form, it follows directly from the Bernstein theorem that $\bar{u} \in H^2(\bar{\Omega})$. Hence

$$\boldsymbol{\sigma} = A \mathbf{grad} u = (a_{11}(x) \frac{\partial u}{\partial x}, a_{22}(y) \frac{\partial u}{\partial y}) = \left(\frac{\partial \bar{u}}{\partial \bar{x}}, \frac{\partial \bar{u}}{\partial \bar{y}} \right) \in \mathbf{H}^1(\Omega),$$

and we have the desired regularity result.

We have shown that except in very special situations, Method III_h does not provide accurate approximations to $\boldsymbol{\sigma}$. It is of interest to ask whether it provides accurate approximations to u . We will end this section by giving a negative answer to this question by comparing $\|u - u_h\|_{L^2(\Omega)}$ and $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega)}^2$. Let $g \in L^2(\Omega)$, let $(\boldsymbol{\gamma}_g, w_g) \in V \times W$ be the solution of Problem P with right hand side $-\int_{\Omega} g v \, dx dy$, and let $(\boldsymbol{\gamma}_{gh}, w_{gh})$ be the solution of P_h with this right hand side. Then from the definitions of Problems P and P_h and of Σ_h , we obtain

$$\begin{aligned}
 \int_{\Omega} (u_h - u)g \, dx dy &= b(\boldsymbol{\gamma}_g, u - u_h) \\
 &= b(\boldsymbol{\gamma}_g - \boldsymbol{\gamma}_{gh}, u - u_h) + b(\boldsymbol{\gamma}_{gh}, u - u_h) \\
 &= b(\boldsymbol{\gamma}_g - \boldsymbol{\gamma}_{gh}, u - \Sigma_h u) + a(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}, \boldsymbol{\gamma}_{gh}) \\
 &= b(\boldsymbol{\gamma}_g - \boldsymbol{\gamma}_{gh}, u - \Sigma_h u) + a(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}, \boldsymbol{\gamma}_{gh} - \boldsymbol{\gamma}_g) + a(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}, \boldsymbol{\gamma}_g) \\
 (4.14) \quad &= b(\boldsymbol{\gamma}_g - \boldsymbol{\gamma}_{gh}, u - \Sigma_h u) + a(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}, \boldsymbol{\gamma}_{gh} - \boldsymbol{\gamma}_g) + b(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, w_g) \\
 &= a(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}, \boldsymbol{\gamma}_{gh} - \boldsymbol{\gamma}_g) + b(\boldsymbol{\gamma}_g, u - \Sigma_h u) + b(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, w_g - \Sigma_h w_g) \\
 &= a(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}, \boldsymbol{\gamma}_{gh} - \boldsymbol{\gamma}_g) + b(\boldsymbol{\gamma}_g, u - \Sigma_h u) + b(\boldsymbol{\sigma}, w_g - \Sigma_h w_g) \\
 &= a(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}, \boldsymbol{\gamma}_{gh} - \boldsymbol{\gamma}_g) - \int_{\Omega} g(u - \Sigma_h u) \, dx dy \\
 &\quad - \int_{\Omega} f(w_g - \Sigma_h w_g) \, dx dy.
 \end{aligned}$$

Letting $g = f$ in (4.14) we have

$$\begin{aligned}
 \int_{\Omega} (u_h - u)f \, dx dy &= a(\boldsymbol{\sigma}_h - \boldsymbol{\sigma}, \boldsymbol{\sigma}_h - \boldsymbol{\sigma}) - 2 \int_{\Omega} f(u - \Sigma_h u) \, dx dy \\
 &\geq \nu \|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}\|_{L^2(\Omega)}^2 - 2 \left| \int_{\Omega} f(u - \Sigma_h u) \, dx dy \right|.
 \end{aligned}$$

Since Σ_h can be taken to be the $L^2(\Omega)$ -projection onto W_h , we have

$$\int_{\Omega} (u_h - u)f \, dx dy \geq \nu \|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}\|_{L^2(\Omega)}^2 - Ch \|u\|_{H^1(\Omega)} \|f\|_{L^2(\Omega)}.$$

Thus

$$(4.15) \quad \|u_h - u\|_{L^2(\Omega)} \geq \frac{\int_{\Omega} (u_h - u)f \, dx dy}{\|f\|_{L^2(\Omega)}} \geq \frac{\nu \|\boldsymbol{\sigma}_h - \boldsymbol{\sigma}\|_{L^2(\Omega)}^2}{\|f\|_{L^2(\Omega)}} - Ch \|u\|_{H^1(\Omega)}.$$

For the case of a rough coefficient, in which we do not expect σ to have much regularity beyond $\mathbf{L}^2(\Omega)$, (4.15) essentially implies that the error $\|u_h - u\|_{L^2(\Omega)}$ cannot be smaller than the square of the error $\|\sigma_h - \sigma\|_{L^2(\Omega)}$. In particular, in the case when σ_h converges to σ , but not with any positive rate of convergence, then u_h will converge to u , but (4.15) implies that this also will not occur with any positive rate of convergence. On the other hand, one can show that $\|u - u_h\|_{L^2(\Omega)}$ is essentially no bigger than $o(\|\sigma - \sigma_h\|_{L^2(\Omega)})$.

5. Comments on special finite element methods.

Another approach to developing finite element methods that work well for problems with rough coefficients is to use special approximating functions—referred to as special elements—that depend on the coefficients. This idea is most easily illustrated by considering the approximation of the one dimensional boundary value problem (3.1)-(3.2). For this problem, it was shown in [2] that the method defined by

Find $\tilde{u}_h \in \tilde{W}_h \equiv \{\tilde{v} \in H_0^1(0, 1) : \tilde{v}|_{I_j} \in \text{span}[1, \int_0^x a^{-1}(s) ds], j = 1, \dots, n\}$ satisfying

$$(5.1) \quad \int_0^1 a \tilde{u}'_h \tilde{v}' dx = \int_0^1 f \tilde{v} dx, \quad \text{for all } \tilde{v} \in \tilde{W}_h,$$

is accurate and robust. Here the variational formulation is the usual displacement formulation, but the approximating functions are linear combinations of 1 and $\int_0^x a^{-1} ds$.

One way to motivate the special element method (5.1) is by an appropriate change of variables. If we let $\bar{x} = \int_0^x a^{-1} ds$ and set $\bar{u}(\bar{x}) = u(x)$, the boundary value problem (3.1)-(3.2) is transformed to

$$(5.2) \quad -\bar{u}''(\bar{x}) = \bar{a}(\bar{x}) \bar{f}(\bar{x}) \quad 0 < \bar{x} < d, \quad \bar{u}(0) = \bar{u}(d) = 0,$$

where $d = \int_0^1 a^{-1} ds$. The solution \bar{u} of (5.2) is in $H^2(0, d)$ and thus the Ritz method using piecewise linear approximating functions gives a good approximation. Transforming back to the x variable, we are led to the special element method (5.1). Note that the mesh $\{x_j\}$ on $[0, 1]$ is obtained by transforming the mesh $\{\bar{x}_j\}$ on $[0, d]$, and that the mesh parameters h and \bar{h} satisfy $\mu^{-1}h \leq \bar{h} \leq \nu^{-1}h$, where μ and ν are lower and upper bounds on $a(x)$. For the remainder of this section, we will suppress the bar on \bar{h} .

It is also possible to apply the technique of constructing special elements via a change of variables in the context of mixed methods. To illustrate this, we again consider (3.1)-(3.2) and its transformed version (5.2). Since $\bar{u} \in H^2(0, d)$, a good mixed method for (5.2) is I_h . For (5.2), the variational formulation I is:

Find $\bar{\sigma} \in L^2(0, d)$, $\bar{u} \in H_0^1(0, d)$ satisfying

$$(5.3) \quad - \int_0^d \bar{\sigma} \bar{v}' d\bar{x} = - \int_0^d \bar{a} \bar{f} \bar{v} d\bar{x} \quad \text{for all } \bar{v} \in H_0^1(0, d),$$

$$(5.4) \quad \int_0^d \bar{\sigma} \bar{\tau} d\bar{x} - \int_0^d \bar{u}' \bar{\tau} d\bar{x} = 0 \quad \text{for all } \bar{\tau} \in L^2(0, d).$$

Method I_h is obtained from I by approximating \bar{u} by piecewise linears and $\bar{\sigma} = d\bar{u}/d\bar{x}$ by piecewise constants. The result of transforming I_h is:

Find $\tilde{\sigma}_h \in V_h \equiv \text{space of piecewise constants}$, $\tilde{u}_h \in \tilde{W}_h$ satisfying

$$(5.5) \quad - \int_0^1 \tilde{\sigma}_h \tilde{v}' dx = - \int_0^1 f \tilde{v} dx \quad \text{for all } \tilde{v} \in \tilde{W}_h,$$

$$(5.6) \quad \int_0^1 a^{-1} \tilde{\sigma}_h \tau dx - \int_0^1 \tilde{u}_h' \tau dx = 0 \quad \text{for all } \tau \in V_h.$$

This is a special element mixed method, and it is equivalent to the special element displacement method (5.1). Specifically, the functions \tilde{u}_h produced by the two methods are identical and

$$\tilde{\sigma}_h(x) \equiv \bar{\sigma}_h(\bar{x}) \equiv d\bar{u}_h(\bar{x})/d\bar{x} = a(x)d\tilde{u}_h(x)/dx.$$

To see this, observe that the Ritz method applied to (5.2) is equivalent to method I_h applied to (5.3)-(5.4) and then use the change of variables.

We next show that (5.5)-(5.6) is closely related to method II_h , a standard method using ordinary elements which was shown in § 3 to be accurate for problems with rough coefficients. To see this, we first note that

$$\int_0^1 \tilde{\sigma}_h \tilde{v}' dx = \int_0^1 \tilde{\sigma}_h(\tilde{v})'_I dx, \quad \int_0^1 \tilde{u}_h' \tau dx = \int_0^1 (\tilde{u}_h)'_I \tau dx,$$

where $(\tilde{v})_I$ denotes the piecewise linear interpolant of \tilde{v} . If we use these relations and replace $\int_0^1 f\tilde{v} dx$ by $\int_0^1 f(\tilde{v})_I dx$ (the latter expression can be viewed as an approximation of the former one) in (5.5), (5.6), we obtain Π_h .

Now Π_h turns out to be equivalent to the following modification of the special element method (5.1) (first analyzed in [2]):

Find $\tilde{w}_h \in \tilde{W}_h$ satisfying

$$(5.7) \quad \int_0^1 a\tilde{w}'_h \tilde{v}' dx = \int_0^1 f(\tilde{v})_I dx \quad \text{for all } \tilde{v} \in \tilde{W}_h.$$

Specifically, if (σ_h, u_h) is the solution of Π_h , then $\sigma_h = a\tilde{w}'_h$ and $u_h = (\tilde{w}_h)_I$. To see this, suppose (σ_h, u_h) is the solution of Π_h and define \tilde{z}_h to be the \tilde{W}_h -interpolant of u_h . (Note this implies that $u_h = (\tilde{z}_h)_I$.) Then from the second equation in Π_h , we obtain

$$\int_0^1 a^{-1}(\sigma_h - a\tilde{z}'_h)\tau dx = 0$$

for all piecewise constant τ , which, together with the fact that $a\tilde{z}'_h$ is a piecewise constant, implies $\sigma_h = a\tilde{z}'_h$. Substituting $a\tilde{z}'_h$ for σ_h in the first equation in Π_h , we get

$$\int_0^1 a\tilde{z}'_h \tilde{v}' dx = \int_0^1 a\tilde{z}'_h (\tilde{v})'_I dx = \int_0^1 f(\tilde{v})_I dx \quad \text{for all } \tilde{v} \in \tilde{W}_h.$$

From the uniqueness of the solution of (5.7), we conclude that $\tilde{z}_h = \tilde{w}_h$ and hence that $\sigma_h = a\tilde{w}'_h$ and $u_h = (\tilde{w}_h)_I$.

We next consider the use of special elements for the two dimensional model problem (4.9)-(4.10). A special finite element method for this problem can be obtained by making the change of variable (4.11). This leads to the boundary value problem (4.12)-(4.13) for $\bar{u}(\bar{x}, \bar{y}) = u(x, y)$. Since $\bar{u} \in H^2(\bar{\Omega})$, the usual Ritz method (applied in the bar variables) gives optimal results. Applying this method and transforming the equations back to the original variables, one obtains a displacement method using special elements derived from piecewise linear elements using the transformation: $u(x, y) = \bar{u}(\bar{x}, \bar{y})$. The special elements are now piecewise linear combinations of

$$(5.8) \quad 1, \quad \int_0^x \frac{1}{a_{11}(s)} ds, \quad \int_0^y \frac{1}{a_{22}(s)} ds.$$

Note that the triangulation of Ω is by curvilinear triangles obtained by transforming a triangulation of $\bar{\Omega}$ by ordinary triangles. For a further discussion of special element displacement methods, see Babuška, Caloz, and Osborn [1] and Babuška and Osborn [3].

As noted above for one dimensional problems, one can apply the technique of constructing special element methods via a change of variables in the context of mixed methods. To illustrate this for two dimensional problems, we again consider (4.9)-(4.10) and its transformed version (4.12)-(4.13). Since $\bar{u} \in H^2(\bar{\Omega})$, a good mixed method for (4.12)-(4.13) is I_h . For this problem, the variational formulation I is:

Find $\bar{\sigma} \in \mathbf{L}^2(\bar{\Omega})$, $\bar{u} \in H_0^1(\bar{\Omega})$ satisfying

$$(5.9) \quad - \int_{\bar{\Omega}} \bar{B} \bar{\sigma} \cdot \mathbf{grad} \bar{v} \, d\bar{x} \, d\bar{y} = - \int_{\bar{\Omega}} \bar{a} \bar{f} \bar{v} \, d\bar{x} \, d\bar{y} \quad \text{for all } v \in H_0^1(\bar{\Omega}),$$

$$(5.10) \quad \int_{\bar{\Omega}} \bar{B} \bar{\sigma} \cdot \bar{\tau} \, d\bar{x} \, d\bar{y} - \int_{\bar{\Omega}} \bar{B} \mathbf{grad} \bar{u} \cdot \bar{\tau} \, d\bar{x} \, d\bar{y} = 0 \quad \text{for all } \bar{\tau} \in \mathbf{L}^2(\bar{\Omega}),$$

where

$$\bar{B} = \begin{pmatrix} a_{22}(\bar{y}) & 0 \\ 0 & a_{11}(\bar{x}) \end{pmatrix}.$$

Method I_h is obtained from I by approximating \bar{u} by piecewise linears and $\bar{\sigma} = (\bar{u}_{\bar{x}}, \bar{u}_{\bar{y}})$ by piecewise constants.

Now both I and I_h can be transformed back to the (x, y) variables. Problem I, i.e., (5.9)-(5.10) becomes:

Find $\sigma \in \mathbf{L}^2(\Omega)$, $u \in H_0^1(\Omega)$ satisfying

$$(5.11) \quad - \int_{\Omega} \sigma \cdot \mathbf{grad} v \, dx \, dy = - \int_{\Omega} f v \, dx \, dy \quad \text{for all } v \in H_0^1(\Omega),$$

$$(5.12) \quad \int_{\Omega} A^{-1} \sigma \cdot \tau \, dx \, dy - \int_{\Omega} \mathbf{grad} u \cdot \tau \, dx \, dy = 0 \quad \text{for all } \tau \in \mathbf{L}^2(\Omega),$$

where

$$A = \begin{pmatrix} a_{11}(x) & 0 \\ 0 & a_{22}(y) \end{pmatrix}.$$

Note that (5.11)-(5.12) is mixed formulation II applied to equations (4.9)-(4.10). The result of transforming I_h amounts to approximating u and $\boldsymbol{\sigma}$ in (5.11)-(5.12) by the elements (5.8) and piecewise constant functions, respectively. Using the equivalence of Problems (5.11)-(5.12) with (5.9)-(5.10) and of (5.9)-(5.10) with (4.12)-(4.13), we immediately obtain the error estimate:

$$\|u - u_h\|_{H^1(\Omega)} + \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega)} \leq Ch\|f\|_{L^2(\Omega)},$$

where C again depends only on the lower and upper bounds for $a_{11}(x)$ and $a_{22}(y)$.

For the one dimensional problem, we saw that there was an equivalence between a modified special element displacement method and a standard mixed method. A reasonable question to ask is whether a similar equivalence holds in two dimensions. To answer this, first recall that method I_h applied to the transformed problem (4.12)-(4.13) is equivalent to the usual Ritz method applied to that problem. Transforming the Ritz method back to the original variables gives the special element displacement method described previously. Hence, this special element method is equivalent to the special element mixed method based on mixed formulation II described above. Specifically, the functions u_h produced by the two methods are identical and $\boldsymbol{\sigma}_h = A \mathbf{grad} u_h$. However, unlike the one dimensional case, there is no simple way to replace the special elements by ordinary elements to produce a standard mixed method. In particular, it is no longer possible to introduce a piecewise linear function z_h such that

$$\int_{\Omega} \mathbf{grad}(z_h - u_h) \cdot \boldsymbol{\tau} \, dx = 0 \quad \text{for all piecewise constant } \boldsymbol{\tau}.$$

We remark that it is also possible to formulate accurate special element methods that use ordinary rather than curvilinear triangles. In particular, if we discretize (5.11)-(5.12) using ordinary triangles, approximate u by special elements, approximate $\boldsymbol{\sigma}$ by piecewise constants, and use piecewise linears and piecewise constants for v and $\boldsymbol{\tau}$, respectively, we get a special element mixed method that is accurate for problems with rough coefficients. This method is equivalent to a special element displacement method known to be accurate for problems with rough coefficients (cf. [1] and [3]). Note that both of these methods use different test and trial functions.

It is also illuminating to consider a similar technique applied to the the approximation of the boundary value problem (4.5)-(4.6), i.e.,

$$-\frac{\partial}{\partial x} \left(a(x) \frac{\partial u}{\partial x} \right) - \frac{\partial}{\partial y} \left(a(x) \frac{\partial u}{\partial y} \right) = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega,$$

where $a = a(x)$ is a function of x only. To obtain a good approximation scheme for this problem, we make the change of variables

$$\bar{x} = \int_0^x \frac{1}{a} ds, \quad \bar{y} = y,$$

and obtain for $\bar{u}(\bar{x}, \bar{y}) = u(x, y)$ the boundary value problem

$$-\frac{\partial \bar{u}}{\partial \bar{x}^2} - \bar{a}^2 \frac{\partial \bar{u}}{\partial \bar{y}^2} = \bar{a} \bar{f} \quad \text{in } \bar{\Omega}, \quad \bar{u} = 0 \quad \text{on } \partial\bar{\Omega},$$

where $\bar{\Omega} = (0, \int_0^1 a^{-1} ds) \times (0, 1)$. Since by the Bernstein result, $\bar{u} \in H^2(\bar{\Omega})$, we again apply the Ritz method (in the bar variables) and then transform the equations back to the original variables. We thus obtain a displacement method whose special elements are now linear combinations of 1, $\int_0^x a^{-1} ds$, and y . In the context of mixed methods, we have

$$\left(\frac{\partial \bar{u}}{\partial \bar{x}}, \frac{\partial \bar{u}}{\partial \bar{y}} \right) \equiv \bar{\sigma}(\bar{x}, \bar{y}) \equiv \sigma(x, y) = \left(a \frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \right).$$

Because $\bar{u} \in H^2(\bar{\Omega})$, we see that $\bar{\sigma} \in \mathbf{H}^1(\bar{\Omega})$. Thus, the smooth vector variable for the boundary value problem (4.5)-(4.6) is the vector $\sigma = (au_x, u_y)$. Again, a good mixed method for this problem is method I_h , for which the corresponding variational formulation I is:

Find $\bar{\sigma} \in \mathbf{L}^2(\bar{\Omega})$, $\bar{u} \in H_0^1(\bar{\Omega})$ satisfying

$$\begin{aligned} - \int_{\bar{\Omega}} \bar{A} \bar{\sigma} \cdot \mathbf{grad} \bar{v} \, d\bar{x} d\bar{y} &= - \int_{\bar{\Omega}} \bar{a} \bar{f} \bar{v} \, d\bar{x} d\bar{y} \quad \text{for all } \bar{v} \in H_0^1(\bar{\Omega}), \\ \int_{\bar{\Omega}} \bar{A} \bar{\sigma} \cdot \bar{\tau} \, d\bar{x} d\bar{y} - \int_{\bar{\Omega}} \bar{A} \mathbf{grad} \bar{u} \cdot \bar{\tau} \, d\bar{x} d\bar{y} &= 0 \quad \text{for all } \bar{\tau} \in \mathbf{L}^2(\bar{\Omega}), \end{aligned}$$

where

$$\bar{A} = \begin{pmatrix} 1 & 0 \\ 0 & \bar{a}^2 \end{pmatrix}.$$

Transforming back to the (x, y) variables, we obtain the variational problem:

Find $\boldsymbol{\sigma} \in \mathbf{L}^2(\Omega)$, $u \in H_0^1(\Omega)$ satisfying

$$(5.13) \quad \int_{\Omega} \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix} \boldsymbol{\sigma} \cdot \mathbf{grad} v \, dx dy = \int_{\Omega} f v \, dx dy \quad \text{for all } v \in H_0^1(\Omega),$$

$$(5.14) \quad \int_{\Omega} \begin{pmatrix} a^{-1} & 0 \\ 0 & a \end{pmatrix} \boldsymbol{\sigma} \cdot \boldsymbol{\tau} \, dx dy - \int_{\Omega} \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix} \mathbf{grad} u \cdot \boldsymbol{\tau} \, dx dy = 0$$

for all $\boldsymbol{\tau} \in \mathbf{L}^2(\Omega)$.

The result of transforming I_h leads to approximation of $\boldsymbol{\sigma}$ by piecewise constants and of u by the special elements used in the special element displacement method derived above for this problem. Although it is clear from the derivation that the special element displacement and mixed methods are equivalent, it is interesting to note that (5.13)-(5.14) is a mixed formulation of the boundary value problem that we have not previously considered, since the new variable introduced is $\boldsymbol{\sigma} = (au_x, u_y)$. If one views the development of a good finite element method for problems with rough coefficients as the search for an appropriate variational principal combined with a good choice of approximating functions, then this mixed method indicates that a wide variety of mixed variational principles may prove useful.

REFERENCES

1. I. Babuška, G. Caloz, and J. Osborn, *Special finite element methods for a class of second order elliptic problems with rough coefficients*, Technical Note BN-1130, Institute for Physical Science and Technology, University of Maryland (1992).
2. I. Babuška and J. Osborn, *Generalized finite element methods: their performance and their relation to mixed methods*, SIAM J. Numer. Anal. **20** (1983), 510-536.
3. I. Babuška and J. Osborn, *Finite element methods for the solution of problems with rough data*, Singularities and Constructive Methods for Their Treatment, Springer-Verlag Lecture Notes in Mathematics 1121 (P. Grisvard, W. Wendland, and J. R. Whiteman, Editors), Springer-Verlag, Berlin-New York, 1985, pp. 1-18.
4. S. N. Bernstein, *Sur la généralisation du problème de Dirichlet*, Math. Ann. **62** (1906), 253-272; 69 (1910), pp. 82-136.

5. F. Brezzi and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer-Verlag, New York, 1991.
6. R. Falk and J. Osborn, *Error estimates for mixed methods*, R.A.I.R.O. **14** (1980), 249-277.
7. P. Grisvard, *Elliptic Problems in Non-smooth Domains*, Pitman Advanced Publishing Program, Boston-London-Melbourne, 1985.
8. O. Ladyzhenskaia and N. Ural'tseva, *Linear and quasilinear elliptic equations*, Academic Press, New York, 1968.

DEPARTMENT OF MATHEMATICS, RUTGERS UNIVERSITY, NEW BRUNSWICK,
NJ 08903

E-mail address: falk@math.rutgers.edu

DEPARTMENT OF MATHEMATICS AND INSTITUTE FOR PHYSICAL SCIENCE AND
TECHNOLOGY, UNIVERSITY OF MARYLAND, COLLEGE PARK, MD 20742

E-mail address: jeo@julia.umd.edu