A THEORY OF MIXED FINITE ELEMENT APPROXIMATIONS
OF NON-SELF-ADJOINT BOUNDARY VALUE PROBLEMS

J. T. Oden

Seventh U. S. National Congress of Applied Mechanics
Boulder, Colorado
June, 1974
Some features of this paper may seem to be a bit too technical for a general exposition on some aspect of mechanics, as this should be, but the subject of the paper is at the heart of much of the work on computer applications in solid mechanics now-a-days -- namely, mixed finite element approximations. It is well known that by using independent approximations for stresses and displacements in linear elasticity problems, for example, improved accuracy in the stresses can sometimes be obtained. For this reason, mixed models have been used in a variety of recent applications in plasticity, viscoelasticity, and fracture mechanics.

The real utility of such mixed models, however, has always been a cloudy issue. It is not difficult to find examples of mixed finite-element approximations which lead to unstable or ill-conditioned equations or which, despite a great deal of additional labor, produce results which are inferior to, or no more accurate than, those obtained using conventional "displacement" methods. The need for an analysis of some of
the intrinsic mathematical properties of mixed approximations has, consequently, been recognized for some time. A first step toward fulfilling that need for a special class of self-adjoint problems has been provided in some recent papers [1,2]. There a theory is developed for the decomposition of an operator of the form $T^*T$ into a single pair of canonical equations which are approximated using Galerkin methods. In the present paper, that theory is expanded to include non-self-adjoint problems and to the more delicate question of multiple decompositions in which any number of dependent variables is approximated simultaneously. In some instances, the results are quite surprising.

2. Decomposible and Non-Self-Adjoint Operators

Most linear problems in solid mechanics involve self-adjoint operators of some type, and are, therefore, especially receptive to approximation by variational methods. Finite element or Galerkin approximations of self-adjoint elliptic problems, for example, lead to symmetric and usually well-conditioned matrices. Frequently, the variational statements of such problems involve coercive bilinear forms, guaranteeing the existence of unique solutions which depend continuously on the data, and, therefore, containing an inherent degree of stability. In recent years, the mathematical theory of finite-element approximations of linear, self-adjoint, elliptic boundary-value problems has been developed.
quite extensively, and convergence criteria, error estimates, criteria for selection of basis functions, and criteria for numerical stability are fairly well known.

The corresponding theory of approximation of non-self-adjoint problems is another matter, particularly when the operators involved are not elliptic. The variational formulation of such problems does not involve the conveniently stabilizing symmetric, coercive bilinear forms, and the corresponding Galerkin approximations lead to unsymmetrical matrices. Numerical experiments indicate that finite-element models of non-self-adjoint problems are, generally speaking, more sensitive to round-off errors than those of self-adjoint problems. Convergence, when it exists, is not monotonic, and specific information on rates-of-convergence does not appear to be available. In short, the non-self-adjoint problem has a number of complicating features that have stood in the way of the development of a complete theory of associated finite-element/Galerkin approximations.

In the present paper, we describe a fairly general theory of finite-element approximation of a class of linear mixed boundary value of the form

\[ Au = f \quad \text{in } \Omega \]

\[ B_1 u = g_1 \text{ on } \partial \Omega_1, \quad B_2 Tu = g_2 \text{ on } \partial \Omega_2 \]

(2.1)

where \( \Omega \) is an open bounded domain in \( n \)-dimensional space \( \mathbb{R}^n \), \( \partial \Omega \) is its boundary, and \( A \) is any linear operator repre-
sentable as the composition of two linear operators \( S \) and \( T \):

\[
A = ST
\]  

(2.2)

The boundary conditions may be mixed; \( \Omega \) is the union of two mutually disjoint sections, \( \Omega_1 \) and \( \Omega_2 \), and the boundary data is given on each \( \Omega_i \) in terms of operators \( B_i \) and \( B_2T \) corresponding to \( A \) in some way that will allow the problem to have a unique solution for reasonable choices of \( f \) and \( g_i \). We shall cite some specific classes of problems subsequently.

The form (2.2) is, of course, quite general. By taking \( S = I \) and \( T = \partial / \partial x \) (\( n = 2 \)), we obtain a general class of first-order boundary value problems. By setting \( S = T^* \), \( T^* \) being the adjoint of \( T \), we obtain the general self-adjoint problem in which \( A = T^*T \). What is significant about the form (2.2) is that it can be decomposed into a pair of lower-order problems, which suggest so-called mixed Galerkin approximations; e.g. \( Au = f \) is equivalent to

\[
Tu = y
\]

(2.3)

\[
Sy = f
\]

By considering a general decomposition of the type (2.3) in which \( S \) is not necessarily the adjoint of \( T \), we uncover a technique for studying the convergence and accuracy of finite element approximations which has more important ramifications than merely the study of non-self-adjoint problems.
We are referring to approximations based on multiple decomposition of a given problem. Such decompositions have always been a popular thing to do in mechanics. For example, we often choose to study, instead of Navier's equations in elasticity,

\[
\frac{\partial}{\partial x_1} (E_{ijkm} \frac{\partial u_K}{\partial x_m}) + \rho f^j = 0
\]

the associated system of equations

\[
u(k,m) = \epsilon_{km}, \quad E^{ijkm} \epsilon_{km} = \sigma^{ij}, \quad \frac{\partial \sigma^{ij}}{\partial x_1} = -\rho f^j
\]

Or, instead of the Bernoulli-Euler beam equation

\[
Du = p \quad (D = \frac{d}{dx})
\]

we consider any of the systems

\[
\begin{align*}
Du &= \sigma \\
D^2u &= M \\
D\sigma &= M \\
D^2\sigma &= p \\
DM &= V \\
DV &= p
\end{align*}
\]

e tc; or in the longitudinal motion of an elastic rod we may replace

\[
\rho \frac{\partial^2 u}{\partial t^2} - \frac{\partial}{\partial x} (E \frac{\partial u}{\partial x}) = \rho f
\]

by the system
\[
\begin{aligned}
\frac{\partial u}{\partial t} &= v \\
\frac{\partial v}{\partial t} &= a \\
\frac{\partial u}{\partial x} &= \varepsilon \\
\frac{\partial \varepsilon}{\partial x} &= \sigma \\
\frac{\partial \sigma}{\partial x} &= \rho(a - f)
\end{aligned}
\]

etc.

In finite-element-Galerkin approximations of such systems, exactly how does the error in one component of the solution influence that of the others? Does, in fact, decomposing the system make it possible to always obtain more accurate results for various derivatives of the primal dependent variable u? Does it pay to use higher order polynomials for displacements than stresses, as some have suggested, or is the converse true? What can be said about the numerical stability of such mixed schemes? We shall attempt to provide some answers for these and related questions for a class of linear boundary-value problems in the developments to follow.

3. **Prolongations, Projections, and Interpolations**

Since Galerkin methods are based on the idea of projections, it is fitting that we record a few properties of orthogonal projections of Hilbert spaces onto finite-dimensional subspaces. The present section is devoted to a brief account of the ideas of biorthogonal bases described in [3,4] and to
a review and slight extension of Aubin's work on prolongations [5].

3.1 Some Notation. Most of what we have to say pertains to projections of elements in a general Hilbert space $H$, but in applications we generally have in mind the Sobolev spaces $H^m(\Omega)$, defined as the completion of the space $C^m(\Omega)$ of functions infinitely differentiable on a bounded domain $\Omega$ in $\mathbb{R}^n$, in the norm

$$||u||_m^2 = \int_\Omega \sum_{|\alpha| \leq m} (D^\alpha u)^2 \, dx$$

(3.1)

where $dx = dx_1 dx_2 \cdots dx_n$ and we have used multi-index notation; i.e., $\alpha$ is an ordered $n$-tuple of non-negative integers, $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ and

$$|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n; \quad x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$$

$$D^\alpha = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}}$$

(3.2)

It can be shown that $H^m(\Omega)$ coincides with the space of functions with weak (generalized) derivatives of order $\leq m$; it is also a Hilbert space, with an inner product given by

$$(u,v)_m = \int_\Omega \sum_{|\alpha| \leq m} D^\alpha u \, D^\alpha v \, dx$$

(3.3)
The dual of $H^m(\Omega)$ is denoted $H^{-m}(\Omega)$, and $H^0(\Omega) = L_2(\Omega)$ is called the pivot space, since it is self dual and

$$\cdots \subset H^2(\Omega) \subset H^1(\Omega) \subset H^0(\Omega) \subset H^{-1}(\Omega) \subset \cdots$$ (3.4)

We also remark that the completion of the space $C_0^\infty(\Omega)$ of infinitely differentiable functions with compact support in $\Omega$ in the norm (3.1) is denoted $H_0^m(\Omega)$; also, the semi-norm

$$|u|_m^2 = \int_{\Omega} \sum_{|\alpha|=m} (D^\alpha u)^2 dx$$ (3.5)

is also useful in describing properties of functions in the quotient space $H^m(\Omega)/P^{m-1}(\Omega)$, where $P^{m-1}(\Omega)$ is the space of polynomials of degree $\leq m - 1$ on $\Omega$.

3.2 Restriction Operators. The most obvious thing about finite element or finite-difference approximations is that they involve fully discrete representations of functions; that is, they involve replacing a differential equation by an algebraic one in which the values of the dependent variable or its derivatives of some order at a finite number of points are to be calculated. In other words, the approximate problem is solved in a space of finite dimension.

Let $R_h^G$ denote a subset of $G$-dimensional euclidean space (the subscript $h$ is to imply that elements in $R_h^G$ may depend upon some mesh parameter $h$ such as the diameter of the largest finite element). Then the mapping $R_h$ which sends any
element \( u \) in \( H^m(\Omega) \) into a set of \( G \) real numbers in \( \mathbb{R}^G \) is called a restriction of \( H^m(\Omega) \) to \( \mathbb{R}^G \):

\[
\mathbb{R}^G_h u = (u^1, u^2, \ldots, u^G) = u
\]  

(3.6)

There are many ways to construct such restrictions, but we shall consider a class of restrictions suggested by finite element methods. We begin by representing \( \Omega \) as the union of \( E \) closed subdomains (finite elements) \( \Omega_e \) (\( \Omega_e \) is the closure of \( \Omega_e \)) such that \( \Omega_e \cap \Omega_f = \phi \) for \( e \neq f \), and we define locally over each \( \Omega_e \) a set of basis functions \( \psi^{(e)}(x) \) which are generally polynomials. Upon connecting all elements together to form a finite element model of \( \Omega \), we simultaneously construct a system of global basis functions.

\[
\varphi_1(x), \varphi_2(x), \ldots, \varphi_G(x) \in H^m(\Omega)
\]  

(3.7)

This procedure is well known and is described fully in, for example, [3]. For approximations in \( H^m(\Omega) \) we shall generally take

\[
\varphi_1(x) \in p^k(\Omega), \quad k \geq m
\]  

(3.8)

where, again \( p^k(\Omega) \) is the space of polynomials of degree \( \leq k \) on \( \Omega \).

The set of \( G \) linearly independent functions \( \{\varphi_i\}_{i=1}^G \) form the basis of a \( G \)-dimensional subspace \( M_h \) of \( H^m(\Omega) \). In general, the set \( \{\varphi_i\} \) is not orthogonal with respect to any inner product. However, a variety of biorthogonal basis can be constructed as follows:
1. For any choice of a Sobolev inner-product (3.3) of order \( \leq m \), construct the Gram matrix

\[
G_{ij}^{(s)} = (\varphi_i, \varphi_j)_s, \quad s \leq m
\]

\[1, j = 1, 2, \ldots, G\] (3.9)

2. Compute the corresponding inverse,

\[
G_{ij}^{(s)} = (G_{ij}^{(s)})^{-1}
\] (3.10)

3. Construct the conjugate basis functions of order \( s \)

(Cf [1,2]),

\[
\varphi_{i(s)}^{(1)}(x) = \sum_{j=1}^{G} G_{ij}^{(s)} \varphi_j(x)
\] (3.11)

4. The conjugate (or dual) basis functions so constructed have the property that

\[ (\varphi_{(s)}^{1}, \varphi_j)_s = \delta_j^{1}, \quad s \leq m \]

\[1, j = 1, 2, \ldots, G\] (3.12)

5. Finally, the \( s \)-th restriction of \( u \in H^m(\Omega) \) to \( \mathbb{R}^G_h \) is defined by

\[
R_h^{(s)} u = (u_{(s)}^1, u_{(s)}^2, \ldots, u_{(s)}^G)
\] (3.13)

wherein

\[
u_{(s)}^{1} = -(u, \varphi_{(s)}^{1})_s
\] (3.14)

For example, the L₂-restriction is simply
$$R_h^{(0)}u \equiv R_hu = \int_{\Omega} u \varphi_j^{(0)} dx = (u, \varphi_j)_0$$

where \((\varphi_j, \varphi_i)_0 = \delta_j^i\).

### 3.3 Prolongations and Projections

Now every function \(U(x)\) in the subspace \(M_h\) is, by definition, a linear combination of the functions \(\varphi_i(x)\); i.e., each \(U(x) \in M_h\) is of the form

$$U(x) = \sum_{i=1}^{G} a^i \varphi_i(x) \quad (3.15)$$

It is important to realize that the coefficients \(a^i\) are elements of \(R_h^G\). Thus, \((3.15)\) defines a mapping \(P_h\) from \(R_h^G\) onto \(M_h\). This mapping is called a prolongation of \(R_h^G\) into \(H^m(\Omega)\), and we write

$$P_h(u^1, u^2, \ldots, u^G) = \sum_{i=1}^{G} u^i \varphi_i(x) \quad (3.16)$$

Clearly, the prolongation \(P_h\) is defined by a given choice of basis functions \(\{\varphi_i(x)\}_{i=1}^{G}\), but it maps different restrictions \(\{u_i(s)\}_{i=1}^{G}\) into different elements \(U(x)\) in \(M_h\).

The compositions

$$\Pi_h(s) = P_h \varphi_h(s), \ s \leq m \quad (3.17)$$

define projections of \(H^m(\Omega)\) onto \(M_h\). Indeed,
\[ \Pi_h^{(s)} u = \sum_{i=1}^{G} (u, \varphi_i^{(s)})_s \varphi_i(x) \]  

(3.18)

For example, the \(L_2\)-projection of \(H^m(\Omega)\) onto \(M\) is

\[ \Pi_h^{(o)} u = \sum_{i=1}^{G} \int_{\Omega} u \varphi_i^{(o)} dx \varphi_i(x) \]

It is natural to seek optimal prolongations or projections for a given restriction (see, for example, [5]).

Seek, for example, the restriction \(u_r^1\) such that

\[ ||u - P_h u_r^1||_s = \inf_{U \in M_h} ||u - U||_s \]

(3.19)

The answer to this problem is, of course,

\[ u_r^1 = (u, \varphi_i^{(s)})_s = u_s^1 \]

(3.20)

Since (3.11) holds and

\[ \frac{a}{a u_k^1} ||u - U||_s = 0 = -2(u, \varphi_k)_s + 2 \sum_{j=1}^{G} u_j^1 G_j^k \]

Moreover, \(u - \Pi_h^{(s)} u\) is orthogonal to \(M_h\) with respect to the inner product \((\cdot, \cdot)_s\).

3.4 Finite Element Interpolation. It is well known that the finite element method has to do with a special but very useful method for constructing the basis functions \(\varphi_i\) of
In general, global functions are generated according to the formula,

$$\varphi_{i}^{\varrho}(\chi) = \bigcup_{e=1}^{E} \sum_{n=1}^{N_{e}} \varrho_{\Omega}^{(e)} \varphi_{N_{n}}^{(e)}(\chi)$$  \hspace{1cm} (3.21)

wherein $E$ is the number of finite elements, $N_{e}$ the number of nodes of element $\Omega_{e}$, $\Omega_{N}$ is the Boolean transformation matrix corresponding to $\Omega_{e}$, and $\varphi_{N}^{(e)}(\chi)$ are local interpolation functions, usually polynomials, such that

$$\varphi_{N}^{(e)}(\chi) = 0, \chi \not\in \Omega_{e} \text{ and } D^{\varrho} \varphi_{N}^{(e)}(\chi^{N}) = \delta_{N}^{M} \delta_{\varrho}^{\varrho}$$  \hspace{1cm} (3.22)

where $\chi_{e}^{N}$ is a mode of $\Omega_{e}$ and $|\varrho|, |\varrho| \leq m, m = \text{integer}$ (see [3] for additional details). By relabelling the $\varphi_{i}^{\varrho}$ of (3.21) we obtain a set $\{\varphi_{i}\}_{i=1}^{G} (3.8)$; they have compact support in $\Omega$, but their corresponding conjugates $\varphi_{i}^{\varrho}$ do not.

We shall assume that the following conditions hold:

1. $\Omega$ is an open bounded domain in $\mathbb{R}^{n}$ that satisfies the cone condition [6];

2. let $u(\chi) \in H^{j}(\Omega)$, and let $\{\varphi_{i}\}_{i=1}^{G}$ be a basis of a $G$-dimensional space $M_{h} \subset H^{k+1}(\Omega)$, the elements of which are piecewise polynomials of order $\leq k$, where $j \geq k + 1$ and $k > n/2$;

3. let $\Pi_{h}$ be a bounded linear operator from $H^{j}(\Omega)$ into $H^{m}(\Omega)$, $0 \leq m \leq k + 1$, such that

$$\Pi_{h} u = u \forall \ u \in P_{k}(\Omega)$$
4. let the dimension G of $M_h$ be increased by regular refinements of finite element meshes [3]; then, it can be shown [7] that as $h$ tends to zero,

$$||E_u||_m \leq c h^{k+1-m} |u|_{k+1}$$  \hspace{1cm} (3.23)

where $c$ is a constant independent of $h$, $|u|_{k+1}$ is defined in (2.5), and $E_u$ is the interpolation error

$$E_u = u - \Pi_h u$$  \hspace{1cm} (3.24)

Interpolation error estimates such as (3.23) are fundamental to finite element convergence theory.

4. Mixed Galerkin Approximations

We now return to the general non-self adjoint boundary-value problem

$$STu = f \quad \text{in } \Omega$$  \hspace{1cm} (4.1)

$$B_1 u = g_1 \text{ on } \partial\Omega_1, \quad B_2 Tu = g_2 \text{ on } \partial\Omega_2$$

which we split into the equivalent pair of equations

$$Tu = y \quad \text{in } \Omega, \quad B_1 u = g_1 \text{ on } \partial\Omega_1$$

$$Sy = f \quad \text{in } \Omega, \quad B_2 y = g_2 \text{ on } \partial\Omega_2$$

Here $ST = A$ is a continuous mapping of a Hilbert space $U$ into a larger space $F$ which contains $f$ (generally $U = H^m(\Omega)$, $F = H^0(\Omega) = L^2(\Omega)$). The operator $T$ is a linear mapping of $U$ into a Hilbert space $Y$ which is also contained in a larger
space $\mathcal{G}$ (often $\mathcal{G} = F$ or $\mathcal{H}F$) and $S$ maps $\psi$ into $F.$

**Definition 4.1.** Let $\mathcal{M}_h^G$ and $\mathcal{N}_\ell^H$ be $G$-dimensional and $H$-dimensional subspaces of $\mathcal{U}$ and $\mathcal{V},$ respectively, spanned by functions $\{\phi_i\}_{i=1}^G$ and $\{\psi_I\}_{I=1}^H$ generated using the finite-element method. Let $(R_h^G, P_h, R_h)$ and $(R_\ell^H, P_\ell, R_\ell)$ be the corresponding finite element approximation of the spaces $\mathcal{U}$ and $\mathcal{V};$ i.e.,

\[
\begin{align*}
R_h: & \quad \mathcal{U} \to \mathbb{R}_h^G, \quad R_h w = w_h \quad \forall w \in \mathcal{U} \\
P_h: & \quad R_h^G \to \mathcal{M}_h^G, \quad P_h w_h = w \in \mathcal{M}_h^G \subset \mathcal{U} \\
R_\ell: & \quad \mathcal{V} \to \mathbb{R}_\ell^H, \quad R_\ell w = w_\ell \quad \forall w \in \mathcal{V} \\
P_\ell: & \quad R_\ell^H \to \mathbb{N}_\ell^H, \quad P_\ell w_\ell = w \in \mathbb{N}_\ell^H \subset \mathcal{V}
\end{align*}
\]  

then the discrete, mixed finite-element approximation of (4.2) is

\[
R_\ell TP_h u_h = v_\ell \\
R_h SP_\ell v_\ell = R_h f
\]

\[
R_\ell, \omega_1 B_1 P_h, \omega_1 u_h = R_\ell, \omega_1 g_1; \quad R_h, \omega_2 B_2 P_\ell, \omega_2 v_\ell = R_h, \omega_2 g_2
\]

where $R_\ell, \omega_1$ etc. indicate extensions of the operators $R_\ell, P_h,$ etc. to $\omega.$

Upon introducing the projections

\[
\Pi_h = P_h R_h, \quad \Pi_\ell = P_\ell R_\ell
\]  

(4.6)
we see that

\[ \Pi_T U = V \]  \hspace{1cm} (4.7)  
\[ \Pi_h S V = \Pi_h f \]

where

\[ U = P_h u_h \quad \text{and} \quad V = P_{\ell} v_{\ell} \]  \hspace{1cm} (4.8)

Following the nomenclature of Aubin [5], the system (4.5) is called the exterior approximation of (4.2) while (4.7) is called the interior approximation of (4.2) (since \( U \) and \( V \) are in \( U \) and \( V \)). Upon elimination of \( V \) from (4.7), we obtain

\[ \Pi_h S \Pi_T U = \Pi_h f \]

from which we solve for the finite-element approximation \( U \) of the solution of (4.1).

**Definition 4.2.** The discrete operator \( A_h : R^G_h \to R^H_{\ell} \) given by

\[ A_h = R_h S P_{\ell} R_{\ell} T P_h \]  \hspace{1cm} (4.10)

is called the stiffness matrix of the mixed finite element approximation (4.5), and the vector

\[ f_h = R_h f \]  \hspace{1cm} (4.11)

is called the consistent load vector.

Again the external approximation of (4.1) is

\[ A_{h\ell} u_h = f_h \]  \hspace{1cm} (4.12)
while the **internal approximation** is

\[ Q_h \ell U = P_h f_h = \Pi_\ell f \]  

(4.13)

where

\[ Q_h U = P_h A_h \ell u_h = \Pi_h \text{SN}_\ell TU \]  

(4.14)

**Definition 4.3.** The functions

\[ e_u = u - U, \quad e_v = v - V \]  

(4.15)

where \((u,v)\) is the solution of (4.2) and \((U,V)\) is the solution of (4.7), are called the **approximation errors** of the mixed finite-element/Galerkin scheme.

It is a common feature of Galerkin methods that the approximate problem can be posed in a larger space than that containing the exact solution. For example, in the problem

\[ D^3 u = f, \quad 0 < x < a \]

where \(f \in H^0(0,a) = L_2(0,a)\), for reasonable choices of boundary conditions and data \(u \in H^3(0,a)\). In the decomposition,

\[ D^2 u = v, \quad Dv = f \]  

(4.16)

we may actually consider the pair

\[ (Du, Dv)_o = (v, \varphi)_o, \quad (Dv, \psi)_o = (f, \psi)_o \]  

(4.17)

for appropriate choices of weights \(\varphi\) and \(\psi\) and homogeneous boundary conditions. Then approximations to \(u\) and \(v\) can be taken from \(H^1(0,a)\) (or possibly \(H^1_0(0,a)\)). This is perfectly
admissible since $H^3(0,a) \subset H^2(0,a) \subset H^1(0,a)$, but it greatly influences the type of convergence we obtain.

With these observations in mind, we denote by $\tilde{u}$ some larger space, containing the "solution space" $U$, into which the operator $A$ is naturally extended by the decomposition and the variational formulation of the problem:

$$u \subseteq \tilde{u}$$

(4.18)

Naturally, we must also consider the possibility that $U = \tilde{U}$.

**Definition 4.4.** The mixed finite-element scheme (4.7) is said to be **convergent in energy** if

$$\lim_{h,\ell \to 0} \| e_u \|_{\tilde{u}} = 0, \quad \lim_{h,\ell \to 0} \| S e_v \|_F = 0$$

(4.19)

where $U \subseteq \tilde{U}$.

**Definition 4.5.** The mixed finite-element approximation (4.7) of (4.2) shall be referred to as **consistent relative** to $\tilde{u}$, $\tilde{U} \supset U$, if

$$\lim_{h \to 0} \| E_w \|_{\tilde{u}} = 0 \quad \text{and} \quad \lim_{\ell \to 0} \| S E_w \|_F = 0$$

(4.20)

and if there exists a constant $M_o > 0$, independent of $h$ and $\ell$, such that

$$\| S \Pi_\ell T w \|_F \leq M_o \| w \|_{\tilde{u}}$$

(4.21)

for every $w \in \tilde{u}$ and $w \in \mathcal{V}$, where $E_w$ and $E_w$ are the interpolation errors [see (3.24)]
\[ E_w = w - \Pi_h w, \quad E_\ell = w - \Pi_\ell w \quad (4.22) \]

**Definition 4.6.** The mixed finite-element approximation of (4.2) given by (4.5) shall be referred to as **stable** if there exists a constant \( M_1 > 0 \), independent of \( h \) and \( \ell \), such that

\[ \| A_{h\ell} w_h \|_G \geq M_1 \| w_h \|_G \quad (4.23) \]

or, equivalently, (4.7) is **stable relative to** \( \bar{u} \), \( U \subset \bar{u} \), if

\[ \| Q_{h\ell} \Pi_h w \|_F \geq M_1 \| \Pi_h w \|_{\bar{u}} \quad (4.24) \]

\( \forall w \in \bar{u} \), wherein \( A_{h\ell} \) and \( Q_h \) are given by (4.10) and (4.14) respectively.

It often happens that the operator \( S \) in (4.1) has a bounded inverse defined on its range. Then a stronger type of stability can be obtained. Recalling that \( T \) maps \( U \) into \( V \) and \( S \) maps \( V \) into \( F \), we observe that the variational problem can be posed in larger spaces \( \bar{U} \supset U \) and \( \bar{V} \supset V \) (often \( \bar{V} = \bar{u} \), as in the example (4.17)). It then makes sense to introduce an additional definition:

**Definition 4.7.** The mixed finite element scheme (4.7) shall be referred to as **strongly stable** relative to \( \bar{V} \supset V \) if it is stable relative to \( \bar{u} \) and if a constant \( N_1 > 0 \) exists such that

\[ \| \Pi_h S \Pi_\ell w \|_F \geq N_1 \| \Pi_\ell w \|_{\bar{V}} \quad (4.25) \]

\( \forall w \in \bar{V} \).
The stage is now set for a look at convergence theory.

5. Convergence and Accuracy

The notions of consistency and stability of difference schemes for linear differential equations are well known, and their relationship to convergence is summed up in the equivalence theorem of Lax. That our definitions given in the previous section lead to the same sort of equivalence theorem for finite element approximations is established in the following theorem.

**Theorem 5.1.** Let the mixed finite-element scheme (4.5) be stable and consistent relative to \( \tilde{U} \supset U \). Then it is convergent in energy.

**Proof:** Suppose that \( u \) is the exact solution to (4.1) and \( \pi_h u \) is its orthogonal projection in \( M_h^Q \). Since the scheme is stable, we have from (4.24)

\[
\| U - \pi_h u \|_\tilde{U} \leq M_1^{-1} \| \pi_h S \|_F \| U - \pi_h u \|_F
\]

\[
= M_1^{-1} \| \pi_h f - \pi_h S(\pi h u - Tu + \xi) \|_F
\]

\[
= M_1^{-1} \| \pi_h S(\xi - \pi h u) \|_F + M_1^{-1} \| \pi h u \|_F
\]

Since \( \| \pi h w \|_F \leq \| w \|_F \) and the scheme is consistent, we can introduce (4.21) to obtain
\[ ||U - \Pi_h u||_\tilde{u} \leq \frac{1}{M_1} ||S\varepsilon_v||_F + \frac{M_0}{M_1} ||E_u||_\tilde{u} \]

However,

\[ ||e_u||_{\tilde{u}} = ||u - U||_{\tilde{u}} \]

\[ = ||u - \Pi_h u - (U - \Pi_h u)||_{\tilde{u}} \]

\[ = ||E_u||_{\tilde{u}} + ||U - \Pi_h u||_{\tilde{u}} \]

Hence,

\[ ||e_u||_{\tilde{u}} \leq (1 + \frac{M_0}{M_1}) ||E_u||_{\tilde{u}} + \frac{1}{M_1} ||S\varepsilon_v||_F \quad (5.1) \]

Since \[ ||E_u||_{\tilde{u}} \] and \[ ||S\varepsilon_v||_F \] vanish as \( h \) and \( \ell \) tend to zero for consistent schemes, so does \[ ||e_u||_{\tilde{u}} \], and the first number of \((4.13)\) holds. To establish that the second member of \((4.19)\) also holds, observe that

\[ ||S\varepsilon_v||_F = ||S\varepsilon - S\varepsilon||_F \]

\[ = ||S\varepsilon - S\Pi_\ell T U - S\Pi_\ell S\varepsilon + S\Pi_\ell S\Pi_\ell T U||_F \]

\[ = ||S(Y - \Pi_\ell Y) + S\Pi_\ell T(u - U)||_F \]

\[ \leq ||S\varepsilon_v||_F + ||S\Pi_\ell T e_u||_F \]

\[ \leq ||S\varepsilon_v||_F + M_0 ||e_u||_{\tilde{u}} \]

Therefore,

\[ ||S\varepsilon_v||_F \leq (1 + \frac{M_0}{M_1}) ||S\varepsilon_v||_F + M_0 (1 + \frac{M_0}{M_1}) ||E_u||_{\tilde{u}} \]
which, in view of (4.20), completes the proof.

Specific rates of convergence are established in the following theorem.

**Theorem 5.2.** Let the conditions of Theorem 5.2 hold, and let the subspaces \( M_h^G \subset \tilde{u}, M_L^H \subset \tilde{v} \) be generated using piecewise polynomial finite element approximations of degree \( k \) and \( r \), respectively (that is, \( M_h^G \subset P_k(\Omega) \) and \( M_L^H \subset P_r(\Omega) \)). Let \( S, T, \) and the data be such that \( \tilde{u} = \tilde{H}^m(\Omega) \), and \( \| S\tilde{v} \|_F \leq \| \tilde{v} \|_p \leq c \| S\tilde{v} \|_F, p \leq m \). Then

\[
\| e_u \|_m \leq c_0 h^{k+1-m} |u|_{k+1} + c_1 l^{r+1-p} |\tilde{v}|_{r+1} \tag{5.3}
\]

\[
\| e_v \|_p \leq c_1 r^{r+1-p} |\tilde{v}|_{r+1} + c_3 h^{k+1-m} |u|_{k+1} \tag{5.4}
\]

wherein \( c_i, i=0,1,2,3 \), are constants independent of \( h \) and \( l \).

**Proof:** The proof follows immediately from (5.1) and (5.2) upon introducing the interpolation error estimates (3.23).

Estimate (5.4) is generally pessimistic, and we can improve it for strongly stable schemes.

**Theorem 5.3.** Consider a mixed finite-element scheme which is consistent relative to \( \tilde{u} \) and strongly stable. Then (5.1) holds, but the error in \( \tilde{v} \) is such that

\[
\| e_v \|_{\tilde{v}} \leq \| E_v \|_{\tilde{v}} + \frac{1}{N_1} \| S E_v \|_F \tag{5.5}
\]
Proof: The fact that (5.1) still holds is obvious from the definition of strong stability and the first steps in the proof of Theorem 5.1. We now abandon (5.2) and use (4.25):

\[
\| \varepsilon_V \|_{\tilde{V}} \leq \| E_V \|_{\tilde{V}} + \| V - \pi_{\ell} V \|_{\tilde{V}} \\
\leq \| E_V \|_{\tilde{V}} + \frac{1}{N_1} \| \Pi_h S \|_{\ell} \| V - \pi_{\ell} V \|_{\tilde{F}} \\
\leq \| E_V \|_{\tilde{V}} + \frac{1}{N_1} \| \Pi_h S \|_{\ell} \| V - \pi_{\ell} V \|_{\tilde{F}} \\
\leq \| E_V \|_{\tilde{V}} + \frac{1}{N_1} \| S E_V \|_{\tilde{F}}
\]

The next result is obvious:

Corollary 5.3.1. Let $S$ map $\tilde{V}$ continuously into $F$. Then for strongly stable, consistent schemes, (5.1) holds and

\[
\| \varepsilon_V \|_{\tilde{V}} \leq (1 + \frac{N_0}{N_1}) \| E_V \|_{\tilde{V}}
\]

(5.6)

where $N_0$ and $N_1$ are constants.

Finally, we have:

Corollary 5.3.2. Consider a consistent, strongly stable mixed finite-element scheme for which (5.6) holds. Suppose that the polynomial subspaces $M_h^G$ and $M_\ell^H$ of Theorem 5.2 are used to construct the bases. Then
\[ ||e_u||_m \leq c_0 h^{k+1-m} |u|_{k+1} + c_1 \ell^{r+1-p} |v|_{r+1} \quad (5.7) \]
\[ ||e_v||_p \leq \hat{c}_2 \ell^{r+1-p} |v|_{r+1} \]

where \( c_0, c_1, \) and \( \hat{c}_2 \) are positive constants independent of \( h \) and \( \ell \).

Thus the error in \( v \) is, in a sense, independent of that in \( u \). We shall interpret these results in a later section.

6. Multiple Decompositions

The general procedure used to study mixed Galerkin approximations of the operator \( ST \) can be extended to multiple decompositions of the type described in (1.5). We shall outline the process below.

We begin by considering a general linear operator equation of the type

\[ Au = f \quad (6.1) \]

where \( A \) is the composition of \( N + 1 \) linear operators,

\[ A = A_N A_{N-1} \cdots A_2 A_1 A_0 = \sum_{i=0}^{N} A_{N-i} \quad (6.2) \]

The operator \( A \) maps a Hilbert space \( U \) continuously onto a space \( F \) containing the data \( f \), and each component \( A_i \) of \( A \) maps an intermediate space \( V_i \) continuously into \( V_{i+1} \); i.e.,

\[ A_0: U \to V_1; \quad A_1: V_1 \to V_2; \quad \cdots; \quad A_{N-1}: V_{N-1} \to V_N; \quad A_N: V_N \to F \quad (6.3) \]
and there exist constants $M_{i,j}$ such that

$$
||A_1 v_1||_{V_{i+1}} \leq M_{i+1,1} ||v_1||_{V_1}
$$

(6.4)

$i=0,1,\cdots,N$. This also leads to inequalities of the type

$$
||A_1 v_1||_{V_{i+1}} \leq M_{i+1,1} ||A_{i-1} v_{i-1}||_{V_1} \leq M_{i+1,1} M_{i-1,1-1} ||v_{i-1}||_{V_{i-1}}
$$

(6.5)

e tc., since the composition of continuous operators is also continuous. Examples of such situations are plentiful; if $u(x) \in H^4(0,L)$ and $Au = D^4u = \frac{d^4u}{dx^4}$ (as in a beam problem with a loading function $p(x)$ in $L^2(0,L)$), $A$ maps $H(0,L)$ continuously into $H^0(0,L)$ but $D = \frac{d}{dx}$ also maps $H(0,L)$ continuously into $H^3(0,L)$ and $H^3(0,L)$ continuously into $H^2(0,L)$, etc.

We next consider the decomposition of (6.1) into $N+1$ component equations,

$$
A_1 v_1 = v_{i+1}, \quad i=0,1,2,\cdots,N-1
$$

$$
A_N v_N = f, \quad v_0 = u
$$

(6.6)

The analysis of this collection of equations by finite-element Galerkin methods proceeds as follows:

1. Let $v_1 \in V_1$ in (6.6), and identify projection operators $\Pi_{h_1}$ of $V_1$ onto finite-dimensional subspaces $M_{h_1}$. The mixed Galerkin approximation (internal) of (6.6) is then
\[ \Pi_{h_1} A_{i-1} V_{i-1} = V, \quad i=0,1,2,\ldots,N \]
\[ \Pi_{h_0} A_N V_N = \Pi_{h_0} f \quad (6.7) \]

with \( V_0 = U \).

2. The approximation errors \( e_{v_1} \) and the interpolation errors \( E_{v_1} \), defined by
\[ e_{v_1} = v_1 - V_1 \quad \text{and} \quad E_{v_1} = v_1 - \Pi_{h_1} v_1 \quad (6.8) \]
are easily seen to obey the "orthogonality" conditions
\[ \Pi_{h_1} A_{i-1} e_{i-1} = e_{v_1} - E_{v_1}, \quad i=0,1,\ldots,N-1 \]
\[ \Pi_{h_0} A_N e_{v_N} = 0 \quad (6.9) \]

3. The approximation \( U \) of \( u \) is determined by the equation
\[ \Pi_{h_0} A_N \Pi_{h_1} A_{N-1} \cdots \Pi_{h_1} A_0 U = \Pi_{h_0} f \quad (6.10) \]

4. The mixed finite-element/Galerkin approximation (6.7) is termed consistent relative to \( \tilde{u}, V_1, \ldots, F \) whenever
\[ ||E_u||_{\tilde{u}}, \quad ||A_1 E_{v_1}||_{V_1}, \ldots, \quad ||A_N E_{v_N}||_F \quad (6.11) \]
vanish as \( h_0, h_1, \ldots, h_N \) tend to zero and when there exist constants \( M_0, M_{ij} \), independent of \( h_0, h_1, \ldots, h_N \) such that
5. The mixed scheme is stable relative to \( \tilde{u} \) if a constant \( C \), independent of \( h_0, h_1, \ldots, h_N \), exists such that

\[
\left\| \sum_{i=0}^{N-1} A_{N-1}^{i} h_{N-1}^{i} A_{0}^{i} v \right\|_{F} \leq M_{0} \| v \|_{\tilde{u}}, \forall v \in \tilde{u}
\]

\[
\left\| A_{i}^{1} h_{1}^{i} A_{i-1}^{i} w \right\|_{V_{i+1}} \leq M_{1,i+1} \left\| A_{i-1}^{i} w \right\|_{V_{i}}
\]

\[\forall w \in V_{i+1}\] (6.12)

6. For stable schemes, a priori error bounds for the error in \( u \) are obtained via (6.13) and the triangle inequality:

\[
\left\| e_{u} \right\|_{\tilde{u}} \leq \left\| E_{u} \right\|_{\tilde{u}} + \left\| U - \pi h_{0} u \right\|_{\tilde{u}}
\]

\[
\leq \left\| E_{u} \right\|_{\tilde{u}} + \frac{1}{c} \left\| \pi h_{0} A_{N}^{N} h_{N}^{N} A_{N-1}^{N-1} A_{0}^{N} h_{0}^{N} (U - \pi h_{0} u) \right\|_{F}
\]

\[
\leq \left\| E_{u} \right\|_{\tilde{u}} + \frac{1}{c} \left\| \pi h_{0} A_{N}^{N} \cdots A_{0}^{N} (u - \pi h_{0} u) + \pi h_{0} A_{N}^{N} E_{V_{N}} + \pi h_{0} A_{N}^{N} h_{N-1}^{N} A_{N-1}^{N-1} E_{V_{N-1}} + \cdots + \pi h_{0} A_{N}^{N} h_{N-1}^{N} \cdots h_{2}^{N} A_{1}^{N} E_{V_{1}} \right\|_{F}
\]

(6.14)

If the scheme is also consistent, we have
Using the conditions (6.8) and (6.12), we find the following error bounds for $V_i$, $i=1,2,\ldots,N$:

$$
\| e_{u} \|_{\tilde{u}} \leq (1 + \frac{M_0}{c}) \| E_{u} \|_{\tilde{u}} + \frac{1}{c} \sum_{i=1}^{N-1} M_{N,N-i} \| E_{v_{N-i}} \|_{V_{N-i}} 
$$

(6.15)

7. Using the conditions (6.8) and (6.12), we find the following error bounds for $V_i$, $i=1,2,\ldots,N$:

$$
\| A_{1} e_{v_{1}} \|_{V_{i+1}} \leq \| A_{1} e_{v_{1}} \|_{V_{i+1}} + \| A_{1} e_{v_{i}} \|_{V_{i+1}} 
$$

$$
\leq \| A_{1} e_{v_{1}} \|_{V_{i+1}} + M_{1,1-i} \| e_{v_{i-1}} \|_{V_{i-1}} 
$$

where the recurrence formula begins with the inequality

$$
\| A_{1} e_{v_{1}} \|_{V_{2}} \leq \| A_{1} e_{v_{1}} \|_{V_{2}} + M_{1,0} \| e_{u} \|_{u} 
$$

$$
M_{1,0}(1 + \frac{M_0}{c}) \| E_{u} \|_{u} + \| A_{1} e_{v_{1}} \|_{V_{2}} 
$$

$$
\frac{M_0}{c} \| A_{N} e_{v_{N}} \|_{F} + \sum_{i=1}^{N-1} M_{N,N-i} \| E_{v_{N-i}} \|_{V_{N-i}} 
$$

(6.17)

etc.

It is clear that stable and consistent schemes are, once again, convergent.

8. Estimates (6.16) and (6.17) are usually pessimistic, because they do not account for the fact that the
component operators $A_i$ have bounded inverses defined on their ranges. To take advantage of this, we shall term the mixed scheme strongly stable if constants $N_{i,j}$ exist

$$\| \Pi h_{N+1} A_{N-1} \Pi h_{N-1} v_{N-1} \| \hat{v}_{N+1} \leq N_{N+1-1,N-1} \| \Pi h_{N-1} v_{N-1} \| \hat{v}_{N-1}$$

$v_{N-1} \hat{v}_{N-1}, i=0,1,2,\ldots,N, \Pi h_{N+1} = \Pi h_0, \hat{v}_{N+1} = \hat{f}$. We then have

$$\| e_{v,N} \| \hat{v}_N \leq (1 + M_{N,N+1}) \| E_{v,N} \| \hat{v}_N \quad (6.19a)$$

$$\| e_{v,N-i} \| \hat{v}_{N-i} \leq (1 + \frac{M_{N+1-i,N-i}}{N_{N-i,N+1-i-1}}) \| E_{v,N-1} \| \hat{v}_{N-1}$$

$$+ \sum_{j=0}^{1-1} c_{ij} \frac{M_{N,N-i+j}}{N_{N-i+j,N}} \| E_{v,N-i+j} \| \hat{v}_{N-i+j} \quad (6.19b)$$

$i=1,2,\ldots,N$, where $c_{ij}, M_{ij}, N_{ij}$ are constants independent of the mesh parameters $h_i, h_0$.

**Theorem 6.1.** For strongly stable, consistent, mixed finite element schemes for multiple decompositions of the form (6.6) in which (6.4) and (6.18) hold, the error bounds (6.19) hold.

This is one of the principal results of this study. We shall interpret it in connection with some simple examples in the remaining sections.
7. Examples

Example 7.1. As a simple one-dimensional example, consider the third-order problem

\[
\frac{d^3u}{dx^3} = f(x) \quad 0 < x < a
\]  

(7.1)

\[
u(0) = u''(0) = u''(a) = 0
\]

We decompose the problem into three components,

\[
\frac{du}{dx} = v, \quad u(0) = 0
\]

(7.2)

\[
\frac{dv}{dx} = w, \quad w(0) = w(a) = 0
\]

Next, we construct finite-element approximations \((U,V,W)\) of \((u,v,w)\). If the interpolation functions used in the approximations \(U, V,\) and \(W\) are denoted \(\psi_i, \psi_j,\) and \(\eta_k\) respectively, the Galerkin approximations are determined from the system of equations

\[
\psi_i \left( \frac{dU}{dx} - V \right) dx = 0; \quad \eta_j \left( \frac{dV}{dx} - W \right) dx = 0; \quad \int \varphi_k \left( \frac{dU}{dx} - f \right) dx = 0
\]

\[
i = 1, 2, \ldots, K; \quad j = 1, 2, \ldots, L; \quad k = 1, 2, \ldots, H, \text{ where}
\]

\[
U = \sum_{k}^{N} a^k \psi_k, \quad V = \sum_{i}^{K} b^i \psi_i, \quad \text{and} \quad W = \sum_{j}^{L} c^j \eta_j.
\]
wherein \( h = a/H, \ k = a/K, \ \ell = a/L. \)

The internal discrete problem leads to the equation

\[
\Pi_h \mathbf{D} \Pi_k \mathbf{D} \Pi_{\ell} \mathbf{D} \Pi_h \mathbf{D} u = \Pi_h f \tag{7.3}
\]

Assuming \( f \in H^0(0,a), \) we have \( \Pi_h : H^3(0,a) \rightarrow M_h, \) \( \Pi_k : H^2(0,a) \rightarrow M_k, \) and \( \Pi_{\ell} : H^1(0,a) \rightarrow M. \) A pessimistic definition of stability for the discrete model is to require a constant \( c > 0 \) to exist such that

\[
||\Pi_h \mathbf{D} \Pi_k \mathbf{D} \Pi_{\ell} \mathbf{D} \Pi_h \mathbf{D} w||_0 \geq c ||\Pi_h w||_3, \quad \forall w \in H^3(0,a),
\]

where \( ||u||_m = ||u||_{H^m(0,a)}, \) and the approximation is stable whenever the interpolation errors \( E_u = u - \Pi_h u, \ E_v = v - \Pi_k v, \ E_w = w - \Pi_{\ell} w \) are such that \( ||E_u||_3, ||E_v||_2, \) and \( ||E_w||_1 \) vanish as \( h, k, \) and \( \ell \) tend to zero and if constants \( M_{ij} \) exist such that

\[
||\Pi_h \mathbf{D} \Pi_k \mathbf{D} \Pi_{\ell} \mathbf{D} \Pi_h \mathbf{D} u||_0 \leq M_{o3} ||u||_3, \quad \forall u \in H^3(0,a)
\]

\[
||\Pi_h \mathbf{D} \Pi_k \mathbf{D} \mathbf{D} \Pi_{\ell} \mathbf{D} v||_0 \leq M_{o2} ||v||_2, \quad \forall v \in H^2(0,a)
\]

\[
||\Pi_{\ell} \mathbf{D} \Pi_k \mathbf{D} \mathbf{D} \Pi_{\ell} \mathbf{D} u||_1 \leq M_{13} ||u||_3, \quad \forall u \in H^2(0,a)
\]

etc.

With all of these preliminaries, it is now a simple matter to obtain the following error bounds

\[
||e_u||_3 \leq (1 + \frac{M_{o3}}{c}) ||e_u||_3 + \frac{M_{o2}}{c} ||e_v||_2 + \frac{M_{o1}}{c} ||e_w||_1
\]

\[
||e_v||_2 \leq M_{23} (1 + \frac{M_{o3}}{c}) ||e_u||_3 + (1 + \frac{M_{o2}}{c}) ||e_v||_2 + \frac{M_{23} M_{o1}}{c} ||e_w||_1
\]
\[
|e_w|_1 \leq M_{13}(1 + \frac{M_{01}}{c})|E_u|_3 + \frac{M_{13}M_{02}}{c}|E_v|_2
\]
\[
+ (1 + \frac{M_{01}}{c})|E_w|_1
\]

(7.4)

The error estimates (7.4) portray a rather dismal picture of the mixed finite element model: the accuracy in each variable is governed by the accuracy of the approximation of \(u\). In turn, \(U\) must be at least a cubic to converge, and even then the accuracy is only of order \(h\). Moreover, piecewise linear approximations, it appears, will not lead to a convergent model at all!

What is wrong with this analysis is that in imposing our rather weak notion of stability we have guaranteed only global stability of the stiffness matrix while allowing its components to behave as wildly as they please. In other words, the approximation \(U\) depends continuously on \(f\) but \(\Pi_k DU\) may not depend continuously on \(V\). This is unrealistic, since it indicates that at least cubics are needed to get convergence in \(H^3(0,a)\). What we really want is convergence of \(U, V,\) and \(W\) to \(u, v,\) and \(w\) in \(H^1(0,a)\).

To achieve this, we redefine the notion of stability: the scheme (7.3) is strongly stable if it is stable and if there exist constants \(c_{0i} > 0, i = 1,2,3\) such that
\[
|\Pi_k D\Pi_h u|_O \geq c_{03}|\Pi_h u|_1, \quad \forall \ u \in H^3
\]
\[
|\Pi_k \Pi_h v|_O \geq c_{02}|\Pi_k v|_1, \quad \forall \ v \in H^2
\]
\[
|\Pi_h D\Pi_k w|_O \geq c_{01}|\Pi_k w|_1, \quad \forall \ w \in H^1
\]

(7.5)
If we denote by $\epsilon_u = U - \Pi_h u$ (and $\epsilon_v, \epsilon_w$) the projection errors, we see that for strongly stable schemes,

$$
||\epsilon_u||_1 \leq \frac{1}{c_{03}} ||\Pi_k D\Pi_h (U - \Pi_h u)||_o
= \frac{1}{c_{03}} ||V + \Pi_k D(u - \Pi_h u) - \Pi_k v||_o
\leq \frac{1}{c_{03}} ||\Pi_k D\epsilon_u||_o + \frac{1}{c_{03}} ||\epsilon_v||_o
$$

$$
||\epsilon_v||_o \leq ||\epsilon_v||_1 \leq \frac{1}{c_{02}} ||\Pi_k D\Pi_k (V - \Pi_k v)||_o
= \frac{1}{c_{02}} ||W + \Pi_k D(v - \Pi_k v) - \Pi_k w||_o
\leq \frac{1}{c_{02}} ||\Pi_k D\epsilon_v||_o + \frac{1}{c_{02}} ||\epsilon_w||_o
$$

$$
||\epsilon_w||_o \leq ||\epsilon_w||_1 \leq \frac{1}{c_{01}} ||\Pi_h D\Pi_k \epsilon_w||_o
\leq \frac{1}{c_{01}} ||\Pi_h D\epsilon_w||_o
$$

Now, using the triangle inequality ($||\epsilon_u|| \leq ||\epsilon_u|| + ||E_u||$)

$$
||\epsilon_u||_1 \leq (1 + \frac{1}{c_{03}})||E_u||_1 + \frac{1}{c_{03} c_{02}} ||E_v||_1 + \frac{1}{c_{03} c_{02} c_{01}} ||E_w||_1
$$

$$
||\epsilon_v||_1 \leq (1 + \frac{1}{c_{02}})||E_v||_1 + \frac{1}{c_{02} c_{01}} ||E_w||_1
$$

$$
||\epsilon_w||_1 \leq (1 + \frac{1}{c_{01}})||E_w||_1
$$

(7.6)

Suppose that we use polynomials of order $i$, $j$, and $k$ for $U$, $V$, and $W$ respectively. Then use of (7.6) together with (3.23) yields the error estimates
This is a remarkable result. It says that the order of error in $w$ is independent of the degree of polynomials used in approximating $v$ and $u$, but that the error in $u$ depends upon the approximations of all components, $u$, $v$, and $w$. If piecewise linear polynomials are used for $w$, it does us no good to use any higher degree polynomials for $u$ and $v$. However, if (say) cubics are used for $w$, we should at least use cubics for $u$ and $v$ to obtain the same order of accuracy for each. If cubics are used for approximating $w$ and linear polynomials for $u$ and $v$, the errors (in the $H^1$-norm) in $w$ are of order $h^3$ while those in $v$ and $u$ are only of order $h$ (assuming $h = k = \ell$).

**Example 7.2.** A special case is the self-adjoint problem

$$D^2u = f, \quad u(0) = u(a) = 0 \quad (7.7)$$

which leads to the pair,

$$\pi_h DU = V, \quad \pi_{\ell} DV = \pi_h f \quad (7.8)$$

For consistent strongly stable schemes, we have

$$||e_u||_1 \leq c_{00} ||E_u||_1 + c_{01} ||E_v||_1 \quad (7.9)$$

$$||e_v||_1 \leq c_{11} ||E_v||_1$$
Thus, for this problem we can draw the same sort of conclusions we derived in the first example.

8. Some Concluding Comments

We can now record a number of important conclusions of our investigation.

1. First, we make the general comment that our approach to a theory of convergence of mixed finite element models has made use of concepts of consistency and stability of Galerkin approximations. The idea of consistency amounts to requiring that the coordinate (or interpolation functions be complete in an appropriate norm, and that the Galerkin approximation of a continuous operator be also continuous for arbitrary choices of mesh parameters. The notion of stability is interpreted in the sense that the approximate solutions must depend continuously on the (discrete) data, and this implies that the matrices involved must have a bounded inverse for arbitrary choices of the mesh parameters. For mixed finite element models, this leads to alternate definitions of stability since, in some cases, component equations approximated in a certain decomposition may, themselves, have bounded inverses. When this happens, we say that the scheme is strongly stable.

2. This approach is quite general. We obtain as special cases the conventional convergence theory and error estimates for "displacement" formulations. However, the
method also applies to non-self-adjoint problems and includes the decomposition of self-adjoint operators as a special case.

3. For consistent and strongly stable schemes, the error in the N-th decomposition variable is independent of all the errors in the remaining approximations. For instance, in approximating the decomposition

\[
\begin{align*}
Du &= v \\
Dv &= w \quad \text{of} \quad D^3u &= f \\
Dw &= f
\end{align*}
\]

the error in \( w \) is independent of the degree of polynomials used in approximating \( u \) and \( v \). However, the error in \( u \) does depend on a linear combination of the interpolation errors in \( u \), \( v \), and \( w \). Consequently, if linear functions are used to approximate \( w \), it does no good to use any higher degree polynomials for \( u \) and \( v \). Likewise, the use of high degree polynomials for \( u \) does not improve the accuracy in \( w \). To obtain the same orders of accuracy for \( u \), \( v \), and \( w \) in the same norms, polynomials of the same degree and proportional mesh sizes should be used for each approximation.

4. For mixed approximations which are not strongly stable, the mixed method may often have little to offer. Here all approximation errors are linear combinations of all interpolation errors. This has advantages in problems in which only a single decomposition into two operators of the same order are used. Then equal orders of accuracy can be obtained in both dependent variables in the same norm.
5. If \( A_{h\ell} \) is any discrete approximation of an operator \( A \) obtained using mixed finite element approximations, the consistency and stability requirements are manifested in respective inequalities

\[
\|A_{h\ell} u_h\|_F \leq N \|u_h\|_u, \quad \|A_{h\ell} v_h\|_F \geq M \|u_h\|_u.
\]

The errors then involve terms like

\[
(1 + \frac{N}{M}) \|E_u\|_u. \quad \text{It is clear that } N \leq \lambda_{\max}, \quad M \geq \lambda_{\min}, \quad \text{where}
\]

\( \lambda's \) are eigenvalues of \( A_{h\ell} \); thus the coefficients in the error estimates depend upon the condition number of the stiffness matrix \( A_{h\ell} \). If these coefficients are to be independent of the mesh parameters \( h \) and \( \ell \), then \( h \) and \( \ell \) cannot be allowed to vanish independently. Generally, we must set

\[
\frac{h}{\ell} = \nu = \text{constant}
\]

or, in multiple decompositions, we set some combination of \( (h_i/h_j) \) or \( (h_i/h_j)^m \) equal to a constant. This done, the coefficients behave as amplification factors of the errors.

6. For consistent, strongly stable schemes such as (8.1), the first approximation \( u \) should be more sensitive to round-off errors than \( v \) or \( w \), since its error bound contains products of the amplification factors of \( v \) and \( w \). For schemes which are not strongly stable, the opposite situation occurs.
Acknowledgement. This work was done while I was a visiting professor at COPPE (Coordenacao dos Programas de Post-Graduacao de Engenharia) of the Universidade Federal do Rio de Janeiro, in Rio de Janeiro. My time there was made particularly pleasant by the hospitality of Messrs. Nelson Ebecken and Cid Gestieia. I am also indebted to my colleague, Dr. J. N. Reddy, with whom I have had many discussions of material related to this work. I also wish to acknowledge, with gratitude, that all of my recent studies on mixed finite elements have been supported by the U. S. Air Force Office of Scientific Research under Contract AFOSR-74-2660.

References


