SIAM J. NUMER. ANAL.
Vol. 21, No. 1, February 1984

© 1984 Society for Industrial and Applied Mathematics
0036-1429/84/2101-0004 $01.25/0

INTERPOLATION OF THE COEFFICIENTS IN NONLINEAR ELLIPTIC GALERKIN PROCEDURES*

J. M. SANZ-SERNA† AND L. ABIA†

Abstract. A continuity argument is employed to prove that the interpolation of the coefficients in nonlinear Galerkin procedures does not result in a reduction of the order of convergence.

1. Introduction. It is well known that the use of interpolation for the evaluation of integrals in nonlinear Galerkin methods can lead to important savings in computational effort. Douglas and Dupont [8] analyzed the parabolic case, while the treatment of the elliptic case goes back, at least, to Herbold, Schultz and Varga [11], [12]. These authors showed that the interpolation of the coefficients of *linear* elliptic problems does not result in a reduction of the order of convergence of the method. They did not consider, however, the nonlinear (elliptic) situation for which existence of the approximate solution does not follow from the variational argument they employed. Recently Christie et al. [4] have introduced the term product approximation to refer to finite-element techniques based on interpolation or related projections. In this paper we provide a proof of the existence and optimal convergence of the so-called product approximation for nonlinear elliptic problems. Although our results will be presented in a model, one-dimensional situation, the ideas and techniques possess a wider generality. In fact, the proofs are carried out in a fashion which renders them readily applicable to higher dimensional problems and some of the possible extensions are discussed in § 4. (Further results can be seen in [1].) Section 2 is devoted to the formulation of the problem to be solved, together with a discussion of the product approximation technique. The main theorem is stated and proved in § 3.

2. Formulation of the problem. We consider the nonlinear two-point boundary value problem

(2.1a) $$-u''+f(x,u)=0, \qquad 0<x<1,$$

(2.1b) $$u(0)=u(1)=0,$$

where a prime denotes differentiation with respect to x and f is a real function continuous and continuously differentiable with respect to u on the strip

$$\Omega=\{(x,v)\colon 0\leqq x\leqq 1, -\infty<v<\infty\}.$$

In addition we assume that there exists a constant $m>-\pi^2$ such that the partial derivative f_u satisfies on Ω the inequality $f_u\geqq m$. These hypotheses guarantee [6], [7], [9] that (2.1) possesses a unique weak solution, i.e. there exists a unique u in H_0^1 such that for v in H_0^1

(2.2) $$(u',v')+(f(\cdot,u),v)=0.$$

Furthermore it will be assumed that u is a classical solution and that f has bounded, continuous partial derivatives up to order $k+1$ on a strip

$$\Omega_\varepsilon=\{(x,v)\colon 0\leqq x\leqq 1, u(x)-\varepsilon<v<u(x)+\varepsilon\}.$$

(Here k is a positive integer and $\varepsilon>0$.) Thus u belongs to $C^{k+3}[0,1]$.

---

* Received by the editors March 5, 1982, and in revised form February 1, 1983.
† Departamento de Ecuaciones Funcionales, Facultad de Ciencias, Universidad de Valladolid, Valladolid, Spain.

For $0<h<1$ we consider a mesh

$$0=\bar{x}_0^h<\bar{x}_1^h<\cdots<\bar{x}_{N_h}^h=1$$

with $h=\max_i(\bar{x}_{i+1}^h-\bar{x}_i^h)$. It is required that the family of meshes is *quasiuniform*. We denote by S^h the space of continuous functions on $[0,1]$ which reduce to a polynomial of degree $\geqq k$ on every subinterval of the mesh. The approximations to u are sought in the subspace S_0^h consisting of those functions in S^h which satisfy the boundary conditions (2.1b). (It should be stressed that Hermite finite-element subspaces can be treated analogously.) As usual the Galerkin approximation $U_h\in S_0^h$ is defined by

$$(U_h', V')+(f(\cdot, U_h), V)=0, \qquad V\in S_0^h. \tag{2.3}$$

Next, every interval $(\bar{x}_i^h, \bar{x}_{i+1}^h)$, $i=0,1,\cdots,N_h-1$, is divided into k uniform subintervals $(\bar{x}_i^h, \bar{x}_{i+1/k}^h)$, $(\bar{x}_{i+1/k}^h, \bar{x}_{i+2/k}^h),\cdots$. In order to shorten the notation the subscript/superscript h will often be omitted in the sequel. Also expressions such as $f(\cdot,u)$, $f(\cdot,U)$ will be simply written as $f(u)$, $f(U)$. We relabel $\bar{x}_{i+j/k}^h$ as x_{ki+j}, $i=0,1,\cdots,N_h-1$ and set $M_h=N_hk-1$.

Let Q_h be the operator which associates with every continuous function g its interpolant $Q_hg\in S^h$ defined by $(Q_hg)(x_i)=g(x_i)$, $i=0,\cdots,M+1$. Then the so-called *product approximation* $W_h\in S_0^h$ is defined by

$$(W_h', V')+(Q_hf(W_h), V)=0, \qquad V\in S_0^h. \tag{2.4}$$

If $\psi_i(x)$, $i=0,1,\cdots,M+1$, is the usual interpolatory basis for S^h, then to compute W_h one has to solve the nonlinear system

$$\begin{aligned}\sum_{i=1}^{M} W_i(\psi_i', \psi_j')+\sum_{i=1}^{M} f(x_i, W_i)(\psi_i, \psi_j)+f(0,0)(\psi_0, \psi_j)&\\ +f(1,0)(\psi_{M+1}, \psi_j)=0, \qquad j=1,\cdots,M,&\end{aligned} \tag{2.5}$$

where W_i are the nodal values of W, $i=1,\cdots,M$. It is now clear that we can first find the inner products (ψ_i', ψ_j'), (ψ_i, ψ_j) and then iteratively solve (2.5). On the other hand, to obtain the standard Galerkin approximation the computation of inner products and the solution of the equations are not separate processes and at each step of the iteration numerical quadrature must be used in order to evaluate the contribution of the nonlinear terms.

By introducing the M-dimensional vectors $\mathbf{W}$, $\mathbf{f}(\mathbf{W})$, $\mathbf{b}$ with components W_i, $f(x_i, W_i)$, $f(0,0)(\psi_0,\psi_i)+f(1,0)(\psi_{M+1}, \psi_i)$ and the $M\times M$ matrices S and A with entries (ψ_i', ψ_j'), (ψ_i,ψ_j), the system (2.5) becomes

$$S\mathbf{W}+A\mathbf{f}(\mathbf{W})+\mathbf{b}=\mathbf{0}. \tag{2.6}$$

The Newton iteration for the solution of (2.6) is given by

$$S\mathbf{W}_{n+1}+A\mathbf{f}(\mathbf{W}_n)+\mathbf{b}+AD(\mathbf{W}_{n+1}-\mathbf{W}_n)=\mathbf{0}, \tag{2.7}$$

where D is the diagonal matrix with diagonal entries equal to $f_u(x_i, W_{ni})$. (We have denoted by W_{ni} the ith component of the iterant $\mathbf{W}_n$.) Thus the updating of the Jacobian can be carried out very cheaply. If we consider now the functions W_n in S_0^h defined by the vectors of nodal values $\mathbf{W}_n$, it is easy to see that (2.7) leads to

$$(W_{n+1}', V')+(Q_nf(W_n), V)+(Q_n[f_u(W_n)(W_{n+1}-W_n)], V)=0, \qquad V\in S_0^h, \tag{2.8}$$

which, in turn, can be interpreted as the result of the linearization of (2.4) around W_n.

Finally let us introduce some further notation. We shall consider the Sobolev space $W^{m,p}$ which, for any integer $m \geqq 0$ and any number p with $1 \leqq p \leqq \infty$, consists of those functions which have m distributional derivatives in $L^p(0, 1)$. The symbol $\|\cdot\|_{m,p}$ denotes the usual norm on $W^{m,p}$. If V belongs to S^h and v is an element of $W^{m,p}$, the norm $\|v - V\|_{m,p}$ must be interpreted elementwise; i.e.,

$$\|v - V\|_{m,\infty} = \max_i \max_{0 \leqq n \leqq m} \sup \left\{ \left| \frac{d^n}{dx^n}(v - V) \right| : \bar{x}_i < x < \bar{x}_{i+1} \right\},$$

with $\|\cdot\|_{m,p}$, $p < \infty$, defined in an analogous manner. Similar interpretations must be given to expressions like $\|f(V)\|_{m,p}$, $\|(d/dx)f(V)\|_{m,p}$, etc $\cdots$.

3. Main result. Under the above smoothness requirements on f and u, we can state the following

THEOREM. *There are positive constants C, h_0 such that for $h < h_0$ there exists a product approximation which verifies*

$$\|u - W_h\|_{0,2} + h\|u - W_h\|_{1,2} \leqq Ch^{k+1}. \tag{3.1}$$

Proof. We first prove the theorem under the additional hypothesis that f has *bounded* continuous partial derivatives up to order $k+1$ on Ω. We set $g(x) = f(x, u(x))$ and for $0 \leqq \lambda \leqq 1$ consider the problem given by (2.1b) and

$$-v'' + \lambda f(x, v) + (1 - \lambda)g = 0, \tag{3.2}$$

which obviously has u as its unique solution. We introduce the Galerkin approximation $U_{h\lambda} \in S_0^h$ given by

$$(U'_{h\lambda}, V') + \lambda(f(U_{h\lambda}), V) + (1 - \lambda)(g, V) = 0, \qquad V \in S_0^h, \tag{3.3}$$

and the approximation $W_{h\lambda} \in S_0^h$ defined by

$$(W'_{h\lambda}, V') + \lambda(Q_h f(W_{h\lambda}), V) + (1 - \lambda)(g, V) = 0, \qquad V \in S_0^h. \tag{3.4}$$

Note that, for $\lambda = 0$, (3.3) and (3.4) are identical, while for $\lambda = 1$, (3.4) reduces to (2.4).

In what follows C will denote a positive constant independent of h and λ and not necessarily the same at each occurrence. The proof, which is somewhat lengthy, is best presented by introducing some lemmas, whose proofs will be postponed until the end of the section.

LEMMA 1. (i) *For every h and λ the Galerkin approximation $U_{h\lambda}$ exists and is unique.*

(ii) *There exists a positive constant C, independent of h and λ such that*

$$\|u - U_{h\lambda}\|_{0,2} + h\|u - U_{h\lambda}\|_{1,2} \leqq Ch^{k+1}. \tag{3.5}$$

(iii) *The norms $\|U_{h\lambda}\|_{r,\infty}$, $r = 0, 1, \cdots, k$ can be bounded independently of h and λ.*

LEMMA 2. *There exists a positive constant, independent of h and λ such that if $W_{h\lambda}$ satisfies* (3.4) *then*

$$\|U_{h\lambda} - W_{h\lambda}\|_{1,2} \leqq C\|f(W_{h\lambda}) - Q_h f(W_{h\lambda})\|_{0,2}. \tag{3.6}$$

The theorem will follow from Lemmas 1 and 2, provided that a suitable bound can be found for the interpolation error in (3.6). This is achieved by the introduction of an a priori assumption, which will later be removed by means of a continuity argument.

We consider the sets

$$B_{h\lambda}=\{V\in S_0^h: \|U_{h\lambda}-V\|_{1,\infty}\leqq h^{k-1}\},$$
$$B_{h\lambda}^0=\{V\in S_0^h: \|U_{h\lambda}-V\|_{1,\infty}<h^{k-1}\}.$$

LEMMA 3. (i) *There exists a positive constant C, independent of h and λ, such that if* $V\in B_{h\lambda}$, *then* $\|V\|_{r,\infty}\leqq C$, $r=0,1,\cdots,k$.

(ii) *There exists a constant* $C>0$, *independent of h and λ, such that if* $W_{h\lambda}\in B_{h\lambda}$ *and* $W_{h\lambda}$ *satisfies* (3.4) *then*

$$\|u-W_{h\lambda}\|_{0,2}+h\|u-W_{h\lambda}\|_{1,2}\leqq Ch^{k+1}.$$

(iii) *There exists a constant* $h_1>0$ *such that, if* $h<h_1$, $0\leqq\lambda\leqq 1$, *and* $W_{h\lambda}\in B_{h\lambda}$, *then* $W_{h\lambda}\in B_{h\lambda}^0$.

(iv) *There exists a constant* $h_2>0$ *such that, if* $X\in B_{h\lambda}$, $T\in S_0^h$, *and* $0\leqq\lambda\leqq 1$ *satisfy*

$$(T',V')+\lambda(Q_h[f_u(X)T],V)=0,\qquad V\in S_0^h, \tag{3.7}$$

then $T=0$.

LEMMA 4. *There exists a constant* h_0 *such that, for any λ such that* $0\leqq\lambda\leqq 1$, *and any* $h<h_0$, *then* (3.4) *has a solution* $W_{h\lambda}\in B_{h\lambda}$.

It is clear that the theorem follows from Lemma 3(ii) and Lemma 4. In order to remove the hypothesis that f has bounded derivatives in Ω we note that (3.1) together with the inverse assumption on the spaces S_0^h show [5] that $\|u-W_h\|_{0,\infty}=o(1)$. Then it suffices to resort to a standard argument. (See, for instance, the proof of [2, Thm. 3.1].)

Proof of Lemma 1. For fixed λ, the existence, uniqueness and optimal rate of convergence of the Galerkin approximation are well known [6], [7], [9]. Inspection of the standard proofs of these facts shows that the constant C in (3.5) can be chosen independently of λ, $0\leqq\lambda\leqq 1$. Finally (3.5) and the inverse assumption imply that $\|\hat{u}_h-U_{h\lambda}\|_{0,\infty}=O(h^{k+1/2})$, where $\hat{u}_h$ is the interpolant of u, and (iii) follows easily. (In fact it is known [9] that $\|\hat{u}_{h\lambda}-U_{h\lambda}\|_{0,\infty}$ possesses an estimate of optimal order. This optimality does not always hold for two-dimensional problems [10] and therefore we have preferred not to resort to it.)

Proof of Lemma 2 (cf. [4]). From (3.3), (3.4) with $V=U_{h\lambda}-W_{h\lambda}$ we can write

$$(U_{h\lambda}'-W_{h\lambda}',U_{h\lambda}'-W_{h\lambda}')+\lambda(f(U_{h\lambda})-Q_hf(W_{h\lambda}),U_{h\lambda}-W_{h\lambda})=0,$$

whence

$$\begin{aligned}&\|U_{h\lambda}'-W_{h\lambda}'\|_{0,2}^2+\lambda(f(U_{h\lambda})-f(W_{h\lambda}),U_{h\lambda}-W_{h\lambda})\\&\qquad\leqq\|f(W_{h\lambda})-Q_hf(W_{h\lambda})\|_{0,2}\|U_{h\lambda}-W_{h\lambda}\|_{0,2}.\end{aligned}\tag{3.8}$$

The condition $f_u\geqq m>-\pi^2$ implies that the left-hand side of (3.8) is larger than $\alpha\|U_{h\lambda}-W_{h\lambda}\|_{1,2}^2$ where α is an absolute constant.

Proof of Lemma 3. The conclusion (i) follows from Lemma 1(iii), the inverse assumption and the triangle inequality. In order to establish (ii), we first write from (3.6)

$$\|U_{h\lambda}-W_{h\lambda}\|_{1,2}\leqq C\|f(W_{h\lambda})-Q_hf(W_{h\lambda})\|_{0,2}\leqq Ch^{k+1}\left\|\frac{d^{k+1}}{dx^{k+1}}f(W_{h\lambda})\right\|_{0,2}. \tag{3.9}$$

(Recall from § 2 that, if necessary, the Sobolev norms must be understood element-wise.) The derivative in (3.9), when expanded, yields a nonlinear expression involving the derivative of f up to order $k+1$ and the derivatives of $W_{h\lambda}$ up to order k. (Note that $(d^{k+1}/dx^{k+1})W_{h\lambda}$ is identically zero within each element.) Therefore, by the additional hypothesis on f and (i) of this lemma,

$$\|U_{h\lambda} - W_{h\lambda}\|_{1,2} \leqq Ch^{k+1}, \tag{3.10}$$

whence

$$\|U_{h\lambda} - W_{h\lambda}\|_{0,2} \leqq Ch^{k+1}. \tag{3.11}$$

Now (3.10), (3.11) lead to the required bound. A new application of the inverse assumption yields $\|U_{h\lambda} - W_{h\lambda}\|_{1,\infty} \leqq Kh^{k+1/2}$, with $K>0$ independent of h and λ, and thus h_1 can be chosen to be any positive number with $Kh^{k+1/2}<h^{k-1}$.

Let us now turn to the proof of (iv). With $V=T$ (3.7) gives

$$(T', T')+\lambda(f_u(X)T, T) \leqq \|(I-Q_h)[f_u(X)T]\|_{0,2}\|T\|_{0,2},$$

where I stands for the identity operator. As in the proof of Lemma 2, this leads to

$$\beta\|T\|_{1,2} \leqq \|(I-Q_h)[f_u(X)T]\|_{0,2}$$

with β an absolute constant. Upon bounding the interpolation error, we can write

$$\|T\|_{1,2} \leqq Ch^k\left\|\frac{d^k}{dx^k}(f_u(X)T)\right\|_{0,2}.$$

An argument similar to the one after (3.9) shows that for $X \in B_{h\lambda}$ the derivatives $(d^r/dx^r)(f_u(X))$, $r=0, 1, \cdots, k$ can be bounded independently of X, h, λ. Then the Leibnitz formula for the derivative of a product and the inverse assumption yield

$$\|T\|_{1,2} \leqq Ch^k\left(\sum_{r=0}^{k}\|T\|_{r,2}\right) \leqq Ch\|T\|_{1,2}$$

and the proof of the lemma is complete.

Proof of Lemma 4. Fix h with $h<\min(h_1, h_2)$ and consider the set $\Lambda \subset [0, 1]$ of all values of λ for which (3.4) has a solution in $B_{h\lambda}$. The value $\lambda=0$ belongs to Λ, because for $\lambda=0$ (3.4) reduces to (3.3). An easy continuity argument shows that Λ is closed. The proof will conclude if we prove that Λ is open in [0, 1]. Let $\lambda_0 \in \Lambda$ and consider the corresponding solution of (3.4) $W_{h\lambda_0} \in B_{h\lambda_0}$. It follows from Lemma 3(iii) that $W_{h\lambda_0} \in \overset{0}{B}_{h\lambda_0}$. According to Lemma 3(iv), with $X=W_{h\lambda_0}$ the system (3.4) has a nonsingular linearization around $W_{h\lambda_0}$, and thus the implicit function theorem forces the existence of solutions of (3.4) for λ in a neighborhood of λ_0. If λ is close enough to λ_0, then the corresponding solution lies by continuity in $B_{h\lambda}$ and therefore λ_0 is interior to Λ.

4. Remarks and extensions. (1) The fact that $\|u-U_k\|_{0,\infty}$ possesses an optimal estimate and the inequality (3.10) imply that the rate of convergence of the product approximation is also optimal in the L^∞ norm.

(2) Newton's method for the computation of the approximation W_h provided by the theorem is locally well defined and convergent. It suffices to note that Lemma 3(iv) with $\lambda=1$ forces the nonsingularity of the relevant Jacobian.

(3) The existence and uniqueness of the product approximation can be proved in some instances by analyzing the system (2.6). The reader is referred to [1] for a discussion of this point.

(4) So far, each element $(\bar{x}_i^h, \bar{x}_{i+1}^h)$ has been divided into *uniform* subintervals $(\bar{x}_i^h, \bar{x}_{i+1/k}^h)$, $(\bar{x}_{i+1/k}^h, \bar{x}_{i+2/k}^h)$, $\cdots$. If F, G are elementwise smooth functions, this leads to an estimate $|(F-Q_hF, G)| \leqq Ch^r$, where C depends on F and G and the

exponent r equals $k+1$ or $k+2$ according to whether r is odd or even respectively. If the nodes within each element $\bar{x}_{i+jk}$, $j=0, 1, \cdots, k$, are chosen to be the *Gauss-Lobatto* points and Q_h, $\psi_i(x)$ modified accordingly, then the form of the product approximation equations (2.4), (2.5) is not altered, while, in the above quadrature error estimate, the exponent r is raised to $2k$. For a linear problem (i.e. f independent of u), this implies readily [13, p. 107] *superconvergence* of order $2k$ at the knots. Recall however that for linear or quadratic elements the uniform subintervals would be identical with those associated with the Gauss–Lobatto abscissae.

(5) The results of § 3 can be extended to the problem

$$-\Delta u + f(x, y, u) = 0, \qquad (x, y) \in G,$$

$$u(x, y) = 0, \qquad (x, y) \in \partial G,$$

where G is a rectangle with sides parallel to the coordinate axes. The region G is divided uniformly into rectangular or triangular elements and S_h is taken to be one of the usual Lagrange type finite element spaces with *interpolatory* bases (linear, quadratic, Lagrange bicubic elements on rectangles, those elements which can be obtained from the previous ones by removing internal nodes . . .).

If f and u are smooth and k is the greatest integer such that all the polynomials of degree not larger than k lie in S_h, then the theorem in § 3 and its proof carry over to the two-dimensional situation with very minor modifications.

(a) The hypothesis $f_u \geqq m > -\pi^2$ must be replaced by $f_u \geqq m > \mu$, where μ is the smallest eigenvalue of the corresponding linear problem.

(b) The relation $\|V\|_{1,\infty} \leqq Ch^{-1/2}\|V\|_{1,2}$, $V \in S_0^h$, used in the proof and valid for one-dimensional problems should be replaced by $\|V\|_{1,\infty} \leqq Ch^{-1}\|V\|_{1,2}$.

(c) In the proof of Lemma 3 advantage was taken of the fact that the derivative $d^{k+1}\varphi/dx^{k+1}$ was elementwise identically zero if $\varphi \in S^h$. In two-dimensional situations not all derivatives of order $k+1$ of the trial functions are necessarily zero. It should be kept in mind, however, that in order to bound the interpolation error $\|f(U) - Qf(U)\|$ one must bound only those derivatives $D^\alpha f(U)$ (α a multi-index with $|\alpha| = k+1$) for which D^α restricted to S^h is identically zero [3]. Inspection of the expansion of $D^\alpha f(U)$ shows that only the derivatives $D^\beta U$, $|\beta| \leqq k$, feature in the bound for the interpolation error.

(6) The reader is referred to [1] for further extensions and numerical results.

Acknowledgment. The authors are very thankful to one of the editors (Prof. Jim Douglas Jr.) for pointing out the advantages to be gained by using the Gauss–Lobatto abscissae.

REFERENCES

[1] L. ABIA, *Interpolation of coefficients in nonlinear Galerkin methods*, Ph.D. Thesis, University of Valladolid, Spain, 1983.

[2] D. N. ARNOLD, J. DOUGLAS, JR. AND V. THOMÉE, *Superconvergence of a finite element approximation to the solution of a Sobolev equation in a single space variable*, Math. Comp., 36 (1981), pp. 53–63.

[3] J. H. BRAMBLE AND S. R. HILBERT, *Bounds for a class of linear functionals with applications to Hermite interpolation*, Numer. Math., 16 (1971), pp. 362–369.

[4] I. CHRISTIE, D. F. GRIFFITHS, A. R. MITCHELL AND J. M. SANZ-SERNA, *Product approximation for non-linear problems in the finite element method*, IMA J. Numer. Anal., 1 (1981), pp. 253–266.

[5] P. G. CIARLET, *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[6] P. G. CIARLET, M. H. SCHULTZ AND R. S. VARGA, *Numerical methods of high-order accuracy for nonlinear boundary value problems*, I. *One dimensional problems*, Numer. Math., 9 (1967), pp. 394–430.

[7] ———, V. *Monotone operator theory*, Numer. Math., 13 (1969), pp. 51–77.
[8] J. DOUGLAS, JR. AND T. DUPONT, *The effect of interpolating the coefficients in nonlinear parabolic Galerkin procedures*, Math. Comp., 29 (1975), pp. 360–389.
[9] G. FAIRWEATHER, *Finite Element Galerkin Methods for Differential Equations*, Marcel Dekker, New York, 1978.
[10] J. FREHSE AND R. RANNACHER, *Asymptotic L^∞-error estimates for linear finite element approximations of quasilinear boundary value problems*, this Journal, 15 (1978), pp. 418–431.
[11] R. J. HERBOLD, M. H. SCHULTZ AND R. S. VARGA, *The effect of quadrature errors in the numerical solution of boundary value problems by variational techniques*, Aeq. Math., 3 (1969), pp. 247–270.
[12] R. J. HERBOLD AND R. S. VARGA, *The effect of quadrature errors in the numerical solution of two-dimensional boundary value problems by variational techniques*, Aeq. Math., 7 (1972), pp. 36–58.
[13] J. DOUGLAS, JR. AND T. DUPONT, *Galerkin approximations for the two point boundary problem using continuous, piecewise polynomial spaces*, Numer. Math., 22 (1974), pp. 99–109.