

A UNIFIED THEORY OF SUPERCONVERGENCE FOR GALERKIN METHODS FOR TWO-POINT BOUNDARY PROBLEMS*

TODD DUPONT†

Abstract. Some superconvergence results of Douglas, Dupont, Rachford and Wheeler are extended and proved in a simple and unified way.

1. Introduction. In [1] Douglas and Dupont noted the fact that if solutions of two-point boundary problems are approximated using Galerkin methods in which the trial functions are merely continuous, then the values of the approximate solutions are much more accurate (superconvergent) at knots than is possible globally. Although derivatives of such approximations are not superconvergent at knots, M. F. Wheeler [8] showed that the method introduced by J. A. Wheeler [7], which involves very simple auxiliary computations on the Galerkin solutions, can be used to produce superconvergent approximations to the derivatives. Pointwise superconvergence at knots was noted by Douglas, Dupont and Wheeler [3] for the so-called H^1 -Galerkin method of Thomée and Wahlbin [6] in the cases in which the trial spaces are either C^1 or C^2 ; in the C^1 case, they noted superconvergence of the derivatives as well. Rachford and Wheeler [4] found that if the trial space is discontinuous, certain auxiliary computations can be made to yield superconvergent approximations for values and derivatives at knots in the H^{-1} -Galerkin method.

In this note, it is shown that performing simple auxiliary computations using the Galerkin solutions of a two-point boundary problem can give superconvergent approximations of the value and derivative (and hence, using the differential equation, *all* derivatives) of the true solution at any point in the interval. In the cases noted above [1], [3] in which the values of the approximate solutions are superconvergent, the auxiliary computations proposed here yield exactly those values as the superconvergent approximations; hence the theory presented here includes that given in those cases. In the cases noted above in which auxiliary computations were used to produce superconvergent approximations at knots, the same values are produced by the methods discussed here.

The theory presented here will all be for problems in which the boundary values are zero, but the extension to more general boundary conditions, such as nonzero boundary values or given derivatives at the endpoints, is straightforward.

2. The standard Galerkin method. Consider the problem

$$(1) \quad Ly = f \quad \text{on } I, \quad y(0) = y(1) = 0,$$

where $I = (0, 1)$, $f \in L^2(I)$, and

$$Lw = -(aw')' + bw' + cw.$$

Assume that a , b and c are in $C^\infty(\bar{I})$, that a is bounded below on \bar{I} by a positive constant a_0 , and that for any f in $L^2(I)$ there exists a unique solution y of (1) in the

* Received by the editors June 10, 1975.

† Department of Mathematics, University of Chicago, Chicago, Illinois, 60637.

Sobolev space $H^2(I)$. Note that it follows that for any f in $L^2(I)$ there exists a unique w in $H^2(I)$ such that

$$(2) \quad L^*w = f \quad \text{on } I, \quad w(0) = w(1) = 0,$$

where L^* is the formal adjoint of L :

$$L^*w = -(aw')' - (bw)' + cw.$$

If $\Delta = \{x_i\}_{i=1}^N$, $0 = x_0 < x_1 < \dots < x_N = 1$, is a partition of I , let $I_i = (x_{i-1}, x_i)$, $h_i = x_i - x_{i-1}$, $h = \max_i h_i$. For integers k and r with $-1 \leq k < r$ let

$$(3) \quad \mathcal{M}_k(r, \Delta) = \{V \in H^{k+1}(I) : V|_{I_i} \in P_r(I_i), 1 \leq i \leq N\},$$

where $H^s(I)$ is the Sobolev space of functions with derivatives through order s in $L^2(I)$ and $P_r(E)$ is the set of polynomials of degree at most r restricted to E . Note that for $0 \leq k$ we can replace $H^{k+1}(I)$ in (3) by $C^k(\bar{I})$ without changing the space $\mathcal{M}_k(r, \Delta)$. The subspace of $\mathcal{M}_k(r, \Delta)$ with zero boundary values will be denoted by $\mathcal{M}_k^0(r, \Delta)$:

$$\mathcal{M}_k^0(r, \Delta) = \{V \in \mathcal{M}_k(r, \Delta) : V(0) = V(1) = 0\}.$$

For given Δ , k and r with $0 \leq k < r$, the Galerkin solution of (1) is defined to be $Y \in \mathcal{M}_k^0(r, \Delta)$ such that

$$(4) \quad B(Y, V) = (f, V), \quad V \in \mathcal{M}_k^0(r, \Delta),$$

where (\cdot, \cdot) is the $L^2(I)$ -innerproduct and B is the bilinear form defined by

$$(5) \quad B(\phi, \psi) = (a\phi', \psi') + (b\phi' + c\phi, \psi).$$

It follows [5], [1] from the nonsingularity of (1) that there exists $h_0 = h_0(k) > 0$ such that, if $h = h(\Delta) < h_0$, then (4) has a unique solution in $\mathcal{M}_k^0(r, \Delta)$ for each $f \in L^2(I)$. It is, of course, possible to define the Galerkin approximation for f in more general classes than $L^2(I)$. However, we shall, throughout this paper, restrict attention to f 's in $L^2(I)$ since it simplifies the discussion at several points.

It is well known (see [1] for a proof) that there is a constant $C = C(r)$ such that

$$(6) \quad \|Y - y\|_{-l} \leq Ch^{s+l}\|y\|_s, \quad 1 \leq s \leq r+1, \quad -1 \leq l \leq r-1,$$

where, for $s \geq 0$, $\|\cdot\|_s$ is the norm on $H^s(I)$ and, for $s < 0$,

$$\|\phi\|_s = \sup \{(\phi, \psi) / \|\psi\|_{-s} : 0 \neq \psi \in H^{-s}(I)\}.$$

It is easily seen from the proof of (6) that the following slightly improved estimate holds: there exists $C = C(r)$ such that

$$(7) \quad (Y - y, \phi) \leq Ch^{s+l}\|y\|_s \left(\sum_{i=1}^N \|\phi\|_{H^l(I_i)}^2 \right)^{1/2}, \quad 1 \leq s \leq r+1, \quad 0 \leq l \leq r-1,$$

for all ϕ such that ϕ is in $H^l(I_i)$ for $1 \leq i \leq N$ and $\phi \in H^q(I)$ where $q = \max(0, \min(l, k-1))$.

In order to define the superconvergent approximations to the solution and its derivative we need two auxiliary functions. In the interest of clarity, a precise construction of these functions will be outlined, but there are many similar functions

that give either exactly the same or equivalent approximations. Let \bar{x} be a point in I ; let $I' = (0, \bar{x})$ and $I'' = (\bar{x}, 1)$. Define a sequence of functions G_n as follows:

$$G_1(x) = \begin{cases} 0 & \text{on } I', \\ x - \bar{x} & \text{on } I'', \end{cases}$$

and for $n > 1$,

$$G_n(x) = \begin{cases} 0 & \text{on } I', \\ G_{n-1}(x) + a_n(x - \bar{x})^n & \text{on } I'', \end{cases}$$

where a_n is chosen so that

$$\left(\frac{d}{dx}\right)^{n-2} L^*G_n(\bar{x}+) = -a(\bar{x})a_n n! + \left(\frac{d}{dx}\right)^{n-2} L^*G_{n-1}(\bar{x}+) = 0.$$

If $\bar{x} \notin \Delta$, let $m = r$; if $\bar{x} \in \Delta$, let $m = \max(k, 1)$. Define G by

$$(8) \quad G(x) = G_m(x) - xG_m(1).$$

Let ϕ be the $L^2(I)$ -function such that

$$(9) \quad \phi(x) = L^*G \quad \text{on } I' \cup I'';$$

note that as distributions $L^*G \neq \phi$. Since integration by parts shows that

$$(10) \quad a(\bar{x})y(\bar{x}) = (y, \phi) - (f, G),$$

it is reasonable to define an approximation $\sigma_{\bar{x}}$ of $y(\bar{x})$ by

$$(11) \quad a(\bar{x})\sigma_{\bar{x}} = (Y, \phi) - (f, G).$$

The definition of $\sigma_{\bar{x}}$ involves integrals over all of I since G is, in general, nonzero on all of I , but as we shall see later, $\sigma_{\bar{x}}$ can actually be computed using f and Y on a very small number (frequently one) of subintervals. Note that if $k = 0$ and $\bar{x} \in \Delta$, then $G \in \mathcal{M}_k^0(r, \Delta)$; thus, using (4) and integration by parts, we see that $\sigma_{\bar{x}} = Y(\bar{x})$ in this case.

In order to define the approximations of the derivatives, define a sequence of functions H_n on I as follows:

$$H_1(x) = \begin{cases} 0 & \text{on } I', \\ 1 & \text{on } I'', \end{cases}$$

and for $n > 1$,

$$H_n(x) = \begin{cases} 0 & \text{on } I', \\ H_{n-1}(x) + b_n(x - \bar{x})^n & \text{on } I'', \end{cases}$$

where b_n is chosen so that

$$\left(\frac{d}{dx}\right)^{n-2} L^*H_n(\bar{x}+) = -a(\bar{x})b_n n! + \left(\frac{d}{dx}\right)^{n-2} L^*H_{n-1}(\bar{x}+) = 0.$$

With m as above, define H by

$$(12) \quad H(x) = H_m(x) - xH_m(1).$$

Let η be the $L^2(I)$ -function such that

$$(13) \quad \eta = L^*H \quad \text{on } I' \cup I''.$$

Motivated by the fact that

$$(14) \quad a(\bar{x})y'(\bar{x}) = b(\bar{x})y(\bar{x}) - (y, \eta) + (f, H),$$

we define an approximation $v_{\bar{x}}$ of $y'(\bar{x})$ by

$$(15) \quad a(\bar{x})v_{\bar{x}} = b(\bar{x})\sigma_{\bar{x}} - (Y, \eta) + (f, H),$$

where $\sigma_{\bar{x}}$ is defined by (11). Just as in the case of $\sigma_{\bar{x}}$, the definition of $v_{\bar{x}}$ involves integration over all of I , but it can, in fact, be computed using Y and f on a very small number (frequently one) of subintervals.

If $\bar{x} = 0$ or 1, define $\sigma_{\bar{x}}$ to be zero; let $m = 1$, and define $v_{\bar{x}}$ by (15) where $H_m(1)$ in (12) is replaced by 1 if $\bar{x} = 1$. Note, in particular, that (14) continues to hold with H defined in this way.

For these approximations we have the following result.

THEOREM 1. *Let k and r be integers such that $0 \leq k < r$. There exists a constant C such that, if y and Y are the solutions of (1) and (4), respectively, if $y \in H^s(I)$ for some s satisfying $2 \leq s \leq r + 1$, and if $\bar{x} \in \bar{I}$, then $\sigma_{\bar{x}}$ and $v_{\bar{x}}$, defined above, satisfy*

$$(16) \quad |\sigma_{\bar{x}} - y(\bar{x})| + |v_{\bar{x}} - y'(\bar{x})| \leq Ch^{s+r-1} \|y\|_s.$$

Proof. From (10) and (11) we see that

$$a(\bar{x})|\sigma_{\bar{x}} - y(\bar{x})| = |(Y - y, \phi)|,$$

where ϕ is defined by (9). Since $\phi \in H^{r-1}(I_i)$ for each $i = 1, \dots, N$, and $\phi \in H^q(I)$, where $q = \max(0, \min(r-1, k-1)) = \max(0, k-1)$, it follows from (7) that

$$(17) \quad |\sigma_{\bar{x}} - y(\bar{x})| \leq Ch^{s+r-1} \|y\|_s \left(\sum_{i=1}^N \|\phi\|_{H^{r-1}(I_i)}^2 \right)^{1/2}.$$

For each n , the coefficients in the piecewise polynomials G_n are clearly bounded independently of $\bar{x} \in I$. Hence, since $1 \leq m \leq r$, we see that the coefficients in the piecewise polynomial G are bounded independently of $\bar{x} \in I$. Thus the sum in (17) is bounded by a constant. A very similar argument, using (17) in addition to (14) and (15), bounds $|v_{\bar{x}} - y'(\bar{x})|$.

If $k = 0$ and $\bar{x} \in \Delta$, then, as previously noted, $\sigma_{\bar{x}}$ is just the value of Y at \bar{x} ; if, on the other hand, $\bar{x} \notin \Delta$, then $\sigma_{\bar{x}}$ can be computed using f and Y only on I_{i_0} where $\bar{x} \in I_{i_0}$. In fact, it is easily checked that

$$(18) \quad a(\bar{x})\sigma_{\bar{x}} = \int_{I_{i_0}} Y\psi_1 - \int_{I_{i_0}} f\psi_2 - \bar{g}[aY(x_{i_0}) - aY(x_{i_0-1})] + aYG'_r(x_{i_0}),$$

where

$$\bar{g} = G_r(x_{i_0})/h_{i_0},$$

$$\psi_2(x) = G_r(x) - (x - x_{i_0-1})\bar{g},$$

$$\psi_1(x) = L^*\psi_2 \quad \text{on } I_{i_0} \cap (I' \cup I'').$$

If $k \geq 1$, then (4) shows that $\sigma_{\bar{x}}$ is unchanged if G is replaced by $G - \chi$ where χ is anything in $\mathcal{M}_k^0(r, \Delta)$. Using the interpolation process in Lemma 2.1 of [2], we see that we need only use q subintervals where q is any integer not less than $(k + 1)/(r - k)$. For any $k \geq 0$ and $r > k$, in the important special case in which $\bar{x} \in \partial I$, the value of $v_{\bar{x}}$ can be found using a single subinterval since there is a function in $\mathcal{M}_k^0(r, \Delta)$ which agrees with H on all but the subinterval next to \bar{x} . Specifically, with $\mu_0 = 1 - (x/h_1)^{k+1}$, $\mu_1 = ((x - x_{N-1})/h_N)^{k+1}$, $J_0 = I_1$ and $J_1 = I_N$,

$$a(i)v_i = \int_{J_i} (f\mu_i - YL^*\mu_i) dx, \quad i = 0 \text{ and } 1.$$

3. H^1 -Galerkin method. For $1 \leq k < r$ and $r \geq 3$ (superconvergence does not occur for $r = 2$) the H^1 -Galerkin approximation of the solution y of (1) is the function $W \in \mathcal{M}_k^0(r, \Delta)$ satisfying

$$(19) \quad (LW, V'') = (f, V''), \quad V \in \mathcal{M}_k^0(r, \Delta).$$

It follows easily from arguments analogous to those in [1] for the standard Galerkin method that there exists $h_0 = h_0(k) > 0$ such that, if $h = h(\Delta) < h_0$, then there exists a unique $W \in \mathcal{M}_k^0(r, \Delta)$ satisfying (19). It is shown in [3] that there exists a constant C such that

$$(20) \quad \|W - y\|_{-l} \leq Ch^{s+l}\|y\|_s, \quad 2 \leq s \leq r+1, \quad -2 \leq l \leq r-3.$$

The proof in [3] is only done in the case $Lw = w''$, but the extension to the more general result is indicated and is easily done. The proof of (20) shows that

$$(21) \quad (W - y, \phi) \leq Ch^{s+l}\|y\|_s \left[\sum_{i=1}^N \|\phi\|_{H^l(I_i)}^2 \right]^{1/2}, \quad 2 \leq s \leq r+1, \quad 0 \leq l \leq r-3,$$

for all ϕ such that $\phi \in H^l(I_i)$ for $i = 1, \dots, N$ and such that $\phi \in H^q(I)$ where $q = \max(0, \min(l, k-3))$.

Take $\bar{x} \in I$. If $\bar{x} \notin \Delta$, let $m = r - 2$, and, if $\bar{x} \in \Delta$, let $m = \max(1, k - 2)$. Define functions G and H by

$$(22) \quad \begin{aligned} G(x) &= G_m(x) - xG_m(1), \\ H(x) &= H_m(x) - xH_m(1), \end{aligned}$$

where G_m and H_m are defined in § 2. Define $\sigma_{\bar{x}}$ and $v_{\bar{x}}$ by

$$(23) \quad \begin{aligned} a(\bar{x})\sigma_{\bar{x}} &= (W, \phi) - (f, G), \\ a(\bar{x})v_{\bar{x}} &= b(\bar{x})\sigma_{\bar{x}} - (W, \eta) + (f, H), \end{aligned}$$

where ϕ and η are $L^2(I)$ -functions defined by

$$(24) \quad \begin{aligned} \phi &= L^*G \quad \text{on } I' \cup I'', \\ \eta &= L^*H \quad \text{on } I' \cup I''. \end{aligned}$$

If $\bar{x} = 0$ or 1, use $\sigma_{\bar{x}} = 0$ and

$$(25) \quad H(x) = H_1(x) - x.$$

Note that y satisfies

$$(26) \quad \begin{aligned} a(\bar{x})y(\bar{x}) &= (y, \phi) - (f, G), \\ a(\bar{x})y'(\bar{x}) &= b(\bar{x})y(\bar{x}) - (y, \eta) + (f, H). \end{aligned}$$

Also note that, if $k = 1$ or 2 and $\bar{x} \in \Delta$, then, since $r \geq 3$, there exists $\tilde{G} \in \mathcal{M}_k^0(r, \Delta)$ such that $\tilde{G}'' \equiv G$ on I ; thus in this case, $\sigma_{\bar{x}} = W(\bar{x})$. Similarly, if k is arbitrary and $x \in \partial I$ or if $k = 1$ and $\bar{x} \in \Delta$, then $v_{\bar{x}} = W'(\bar{x})$.

It follows easily from (23), (24), (26), (21) and the definitions of G and H that the following result holds.

THEOREM 2. *Let k and r be integers such that $1 \leq k < r$ and $3 \leq r$. Then there exists a constant C such that if y and W are the solutions of (1) and (19), respectively, if $y \in H^s(I)$ for some s satisfying $2 \leq s \leq r + 1$, and if $\bar{x} \in \bar{I}$, then $\sigma_{\bar{x}}$ and $v_{\bar{x}}$ defined above satisfy*

$$(27) \quad |\sigma_{\bar{x}} - y(\bar{x})| + |v_{\bar{x}} - y'(\bar{x})| \leq Ch^{s+r-3} \|y\|_s.$$

Just as in the standard Galerkin case, the equation that defines the approximate solution can be used to reduce the number of subintervals that need be used in computing $\sigma_{\bar{x}}$ and $v_{\bar{x}}$. If $k = 1$ or 2 and $\bar{x} \in I_{i_0}$, then a very close analogue of (18) can be used to compute $\sigma_{\bar{x}}$; the only changes are that Y is replaced by W and that G_r is replaced by G_{r-2} . To verify such a formula, one uses (19) (with V chosen so that $V'' \in \mathcal{M}_0(r-2, \Delta)$, $V'' \equiv G$ on \bar{I}/I_{i_0} and V'' is linear on I_{i_0}) and (23). A similar formula for $v_{\bar{x}}$ can be derived in this manner with H replacing G . If $k \geq 3$, then we can compute $\sigma_{\bar{x}}$ and $v_{\bar{x}}$ using q subintervals where $q \geq (k-1)/(r-k)$; interpolate G into $\mathcal{M}_{k-2}(r-2, \Delta)$ using the previously mentioned interpolation process of [2].

4. H^{-1} -Galerkin method. For $-1 \leq k < r$ the H^{-1} -Galerkin approximation of the solution y of (1) is the function $Z \in \mathcal{M}_k(r, \Delta)$ such that

$$(28) \quad (Z, L^*V) = (f, V), \quad V \in \mathcal{M}_{k+2}^0(r+2, \Delta).$$

It is shown in [4] that there exists $h_0 = h_0(k) > 0$ such that, if $h = h(\Delta) < h_0$, then there exists a unique solution of (28); actually, in [4] it is assumed that the mesh is quasi-uniform, but this is only needed in their proofs for max-norm estimates. It also follows from [4] that there exists a constant C such that

$$(29) \quad \|Z - y\|_{-l} \leq Ch^{s+l} \|y\|_s, \quad 0 \leq s \leq r+1, \quad 0 \leq l \leq r+1,$$

The proof of (29) can be modified to show that

$$(30) \quad (Z - y, \phi) \leq Ch^{s+l} \|y\|_s \left[\sum_{i=1}^N \|\phi\|_{H^1(I_i)}^2 \right]^{1/2}, \quad 0 \leq s \leq r+1, \quad 0 \leq l \leq r+1,$$

for all ϕ such that $\phi \in H^1(I_i)$ for $i = 1, \dots, N$ and such that $\phi \in H^q(I)$, where $q = \min(l, k+1)$.

Take $\bar{x} \in I$. If $\bar{x} \notin \Delta$, let $m = r+2$, and, if $\bar{x} \in \Delta$, let $m = k+2$. Define functions G and H by

$$(31) \quad \begin{aligned} G(x) &= G_m(x) - xG_m(1), \\ H(x) &= H_m(x) - xH_m(1), \end{aligned}$$

where G_m and H_m are defined in § 2. Define $\sigma_{\bar{x}}$ and $v_{\bar{x}}$ by

$$(32) \quad \begin{aligned} a(\bar{x})\sigma_{\bar{x}} &= (Z, \phi) - (f, G), \\ a(\bar{x})v_{\bar{x}} &= b(\bar{x})\sigma_{\bar{x}} - (Z, \eta) + (f, H), \end{aligned}$$

where ϕ and η are $L^2(I)$ -functions defined by

$$(33) \quad \phi = L^*G, \quad \eta = L^*H \quad \text{on } I' \cup I''.$$

If $\bar{x} = 0$ or 1, use $\sigma_{\bar{x}} = 0$ and

$$(34) \quad H(x) = H_1(x) - x.$$

Since

$$(35) \quad \begin{aligned} a(\bar{x})y(\bar{x}) &= (y, \phi) - (f, G), \\ a(\bar{x})y'(\bar{x}) &= b(\bar{x})y(\bar{x}) - (y, \eta) + (f, H), \end{aligned}$$

the following result follows easily from (32), (35) and (30).

THEOREM 3. *Let k and r be integers such that $-1 \leq k \leq r$ and $r \geq 1$. Then there exists C such that if y and Z are the solutions of (1) and (29), respectively, if $y \in H^s(I)$ for some s satisfying $2 \leq s \leq r+1$, and if $\bar{x} \in \bar{I}$, then $\sigma_{\bar{x}}$ and $v_{\bar{x}}$ defined above satisfy*

$$(36) \quad |\sigma_{\bar{x}} - y(\bar{x})| + |v_{\bar{x}} - y'(\bar{x})| \leq Ch^{r+s+1} \|y\|_s.$$

Note that (28) implies that the values of $\sigma_{\bar{x}}$ and $v_{\bar{x}}$ are not changed if G and H are modified (in (32) and (33)) by subtracting any elements of $\mathcal{M}_{k+2}^0(r+2, \Delta)$. Thus, by using the interpolation process of [2], we see that we need only use q subintervals to compute $\sigma_{\bar{x}}$ and $v_{\bar{x}}$, where $q \geq (k+3)/(r-k)$. In the case $r=1$ and $k=-1$, for example, we can subtract from G and H their Hermite cubic interpolants to get formulas for $\sigma_{\bar{x}}$ and $v_{\bar{x}}$ that involve integrals over only one subinterval.

REFERENCES

- [1] J. DOUGLAS, JR., AND T. DUPONT, *Galerkin approximations for the two point boundary value problem using continuous, piece-wise polynomial spaces*, Numer. Math., 22 (1974), pp. 99–109.
- [2] J. DOUGLAS, JR., T. DUPONT AND L. WAHLBIN, *Optimal L_∞ error estimates for Galerkin approximations to solutions of two point boundary value problems*, Math. Comp., 29 (1975), pp. 475–483.
- [3] J. DOUGLAS, JR., T. DUPONT AND M. F. WHEELER, *Some superconvergence results for an H^1 -Galerkin procedure for the heat equation*, Springer Lecture Notes in Computer Science, No. 10, Springer-Verlag, Berlin, 1974, pp. 288–309.
- [4] H. H. RACHFORD, JR., AND M. F. WHEELER, *An H^{-1} -Galerkin procedure for the two-point boundary value problem*, Mathematical Aspects of Finite Elements in Partial Differential Equations, C. deBoor, ed., Academic Press, New York, 1974, pp. 353–382.
- [5] A. H. SCHATZ, *An observation concerning Ritz-Galerkin methods with indefinite bilinear forms*, Math. Comp., 28 (1975), pp. 959–962.
- [6] V. THOMÉE AND L. WAHLBIN, *On Galerkin methods in semilinear parabolic problems*, this Journal, 12 (1975), pp. 378–389.
- [7] J. A. WHEELER, *Simulation of heat transfer from a warm pipeline buried in permafrost*, 74th Nat. Meeting, Amer. Inst. of Chem. Eng., New Orleans, March 1973.
- [8] M. F. WHEELER, *A Galerkin procedure for estimating the flux for two-point boundary problems*, this Journal, 11 (1974), pp. 764–768.