MATHEMATICS OF COMPUTATION
VOLUME 36, NUMBER 153
JANUARY 1981

Local Piecewise Polynomial Projection Methods for an O.D.E. Which Give High-Order Convergence at Knots

By Carl de Boor and Blair Swartz*

Abstract. Local projection methods which yield $C^{(m-1)}$ piecewise polynomials of order $m + k$ as approximate solutions of a boundary value problem for an mth order ordinary differential equation are determined by the k linear functionals at which the residual error in each partition interval is required to vanish on. We develop a condition on these k functionals which implies breakpoint superconvergence (of derivatives of order less than m) for the approximating piecewise polynomials. The same order of superconvergence is associated with eigenvalue problems. A discrete connection between two particular projectors yielding $\mathcal{O}(|\Delta|^{2k})$ superconvergence, namely (a) collocation at the k Gauss-Legendre points in each partition interval and (b) "essential least-squares" (i.e., local moment methods), is made by asking that this same order of superconvergence result when using collocation at $k - r$ points per interval and simultaneous local orthogonality of the residual to polynomials of order r; the $k - r$ points then necessarily form a subset of the k Gauss-Legendre points.

Introduction. This is the last in a triple (see [2], [3]) of papers concerned with high-order approximation to eigenvalues of an O.D.E. using collocation at Gauss points. Correspondingly, its two sections are labelled 9 and 10, but it can be read without reference to [3], i.e., to Sections 5–8. Items labelled x.y or (x.y) are to be found in Section x, e.g., in [2] in case x is less than 5.

When writing [2], we were forced to go through the arguments in [1] once again and ended up improving upon them somewhat; see the proof of Theorem 9.2 below. In the process, we considered more general local piecewise polynomial projection methods in an effort to discover just what produces the superconvergence at breakpoints in Gauss-point collocation. This led us to a simple set of conditions on the local projector used which, so we found, had been formulated much earlier by Pruess [4] in another context. In addition to updating our earlier results in [1] and [2] to cover this wider class of projection methods, we give a detailed analysis of these special local projectors and establish a simple link between the two best known among these, viz. Interpolation at Gauss points and Least-squares approximation.

9. Some Projectors Which Yield Superconvergence. As de Boor and Swartz [1] describe it, local projection methods which involve sufficiently rough piecewise

Received April 11, 1980.

1980 *Mathematics Subject Classification.* Primary 65L15, 65Jxx.

Key words and phrases. Piecewise polynomial, projection methods, ordinary differential equations, boundary value problems, eigenvalue problems, superconvergence, ssuper projectors.

*Sponsored by the United States Army under Contract No. DAAG29-75-C-0024 and by the United States Department of Energy under Contract No. W-7405-Eng.36.

0025-5718/81/0000-0002/$04.25

polynomials are basically determined by a bounded linear projector Q which carries $C[-1, 1]$ onto $\mathbf{P}_k$ (polynomials of order k, i.e., of degree $< k$), and hence satisfies

(1) $$\|f - Qf\| \leqslant \mathrm{const}_Q \|D^k f\|_\infty, \quad \text{all } f \in C^{(k)}[-1, 1],$$

for some constant const_Q. Then, given a partition $\Delta := (t_i)_0^l$ of $[0, 1]$ with

$$0 = t_0 < \cdots < t_l = 1, \qquad |\Delta| := \max_i \Delta t_i,$$

Q determines a map Q_Δ projecting $\times_{i=1}^l C[t_{i-1}, t_i] =: C_\Delta$ onto $\mathbf{P}_{k,\Delta}$ (the space of piecewise polynomials of order k with breakpoints in Δ) by translating the procedure for $C[-1, 1]$ to each partition interval; i.e., by requiring that, on each $[t_i, t_{i+1}]$ and for $y \in C_\Delta$,

(2) $$Q_\Delta y = S_i^{-1} Q S_i y, \quad \text{with } (S_i g)(s) := g(t_{i+1/2} + s\Delta t_i/2) \text{ for } s \in [-1, 1].$$

Then, from (1),

(1_Δ) $$\|f - Q_\Delta f\|_{(i)} \leqslant \mathrm{const}_Q |\Delta t_i|^k \|D^k f\|_{(i)},$$

with

$$\|g\|_{(i)} := \sup\{|g(t)|: t_i < t < t_{i+1}\}.$$

Finally, the projection method for the mth order differential equation $Mx = y$, $\boldsymbol{\beta} x = \mathbf{0}$, is determined by requiring that $x_\Delta \in \mathbf{P}_{m+k,\Delta}^m := \mathbf{P}_{m+k,\Delta} \cap C^{(m-1)}[0, 1]$ satisfy

(3) $$Q_\Delta M x_\Delta = Q_\Delta y, \qquad \boldsymbol{\beta} x_\Delta = \mathbf{0}.$$

We consider a set of constraints upon Q which permits proof of $\mathcal{O}(|\Delta|^{k+n})$ breakpoint superconvergence for this projection method. These constraints, constructed by Pruess in another context [4, pp. 553–554, esp. p. 554, line 5], can be stated as follows:

For some positive integer $n \leqslant k$ (and in terms of $\mathbf{L}_2[-1, 1]$),

(4) $$\mathbf{P}_i \perp (1 - Q)[\mathbf{P}_{k+n+1-i}], \qquad i = 1, \ldots, n.$$

This condition is equivalent to the following: For some sequence $(f_i)_1^{k+n}$ with

(5a) $$\mathbf{P}_j = \mathrm{span}(f_i)_1^j, \quad \text{all } j,$$

we have

(5b) $$Qf_j = 0 \quad \text{for } j > k,$$

and

(5c) $$\int_{-1}^1 f_i f_j = 0 \quad \text{for } i \leqslant k < j \leqslant k + n + 1 - i.$$

Indeed, by (5a), (4) is equivalent to having

$$\int f_r (1 - Q) f_j = 0 \quad \text{for } r \leqslant i, \ \ j \leqslant k + n + 1 - i, \text{ and } i = 1, \ldots, n.$$

In fact, since $(1 - Q) f_j = 0$ for $j \leqslant k$, (4) is equivalent to having

$$\int f_r (1 - Q) f_j = 0 \quad \text{for } r \leqslant i, \ \ k < j \leqslant k + n + 1 - i, \text{ and } i = 1, \ldots, n,$$

i.e., for $j > k$ and $r \leqslant k + n + 1 - j$, and so, by (5b), (4) is equivalent to (5c). This shows that (5a–c) implies (4). On the other hand, for any linear projector Q onto $\mathbf{P}_k$, we can find $(f_i)_1^{k+n}$ satisfying (5a–b) by taking

$$f_i := \begin{cases} g_i, & i \leqslant k, \\ (1 - Q)g_i, & i > k, \end{cases} \quad \text{with } (g_i) \text{ s.t. } \mathbf{P}_j = \operatorname{span}(g_i)_1^j, \text{ all } j,$$

hence the argument also shows that (4) implies (5a–c). Finally, this last statement shows (with $g_i(t) = t^{i-1}$, all i) that (4) is also equivalent to

$$\int_{-1}^{1} t^r(1 - Q)t^s = 0 \quad \text{for } r < n, \quad r + s < k + n. \tag{6}$$

Since Pruess was the first to consider projectors satisfying (4) (i.e., (6), see [5]), we call any linear projector Q onto $\mathbf{P}_k$ and satisfying (4) a *ssuper projector of order* (k, n).

Example 1. *Collocation*. Taking, in particular, $f_i(t) = \Pi_{j<i}(t - \rho_j)$, with $\rho_1, \ldots, \rho_k$ the collocation pattern and $\rho_{k+1}, \ldots, \rho_{k+n}$ arbitrary, we find, from (5), the condition

$$\int_{-1}^{1} p(t) \prod_{j=1}^{k} (t - \rho_j)\, dt = 0 \quad \text{for all } p \in \mathbf{P}_n$$

(used in [1]) to imply that Q, given by polynomial interpolation at $\rho_1, \ldots, \rho_k$, is a ssuper projector of order (k, n).

Example 2. *Essential least squares* (*method of moments, or of iterated integrals*). Taking, in particular, $f_i = P_{i-1}$:= the Legendre polynomial of degree $i - 1$, all i, we find that Q, given as least squares approximation from $\mathbf{P}_k$, is a ssuper projector, of order (k, k). We have called the corresponding process "*essential* least squares" because the associated projection method (3) requires that the residual error, $Mx_\Delta - y$, be orthogonal to $\mathbf{P}_{k,\Delta} = D^m[\mathbf{P}^m_{m+k,\Delta} \cap \ker \boldsymbol{\beta}]$ (assuming $\boldsymbol{\beta} = (\beta_i)_1^m$ to be linearly independent on $\mathbf{P}_m$); while ordinary least squares asks that this residual be orthogonal to $M[\mathbf{P}^m_{m+k,\Delta} \cap \ker \boldsymbol{\beta}]$. This process has also been called a "local moment method" for an mth order equation. In this connection, recall that Wittenbrink [6, Example 3c] shows this to be equivalent to asking that the iterated integrals of order j, $1 \leqslant j \leqslant k$, of the residual error vanish at all the breakpoints. We have chosen, however, to emphasize its connection to least squares.

The validity of (4) suffices for proof of the following result from which we shall conclude $\mathcal{O}(|\Delta|^{k+n})$ breakpoint superconvergence. The lemma (and its proof) are a variant of Pruess' result [4, Section 3] and [5, Lemma 2].

LEMMA 9.1. *Let Q be a ssuper projector of order (k, n). Then*

$$\left| \int_{-1}^{1} f(1 - Q)g \right| \leqslant \text{const}_Q \sum_{j<n} \|D^j f\| \, \|D^{k+n-j} g\|. \tag{7}$$

Proof. Let $(T_j f)(t) := \Sigma_{i<j} D^i f(0)\, t^i/i!$. Then

$$\int_{-1}^{1} f(1 - Q)g = \int_{-1}^{1} f(1 - Q)T_{k+n}g + \mathcal{O}(\|f\| \, \|D^{k+n}g\|),$$

while

$$\int_{-1}^{1} f(1 - Q)T_{k+n}g = \sum_{r=k}^{k+n-1} D^r g(0)/r! \int_{-1}^{1} f(t)(1 - Q)t^r,$$

since Q reproduces $\mathbf{P}_k$. On the other hand, since Q is a ssuper projector of order (k, n),

$$\int_{-1}^{1}(T_{k+n-r}f)(1-Q)t^r = 0$$

(using (6)), and so

$$\int_{-1}^{1}f(1-Q)t^r = \int_{-1}^{1}(f - T_{k+n-r}f)(1-Q)t^r = \mathcal{O}(\|D^{k+n-r}f\|).$$

Consequently,

$$\int_{-1}^{1}f(1-Q)g = \mathcal{O}\left(\sum_{r=k}^{k+n}\|D^r g\|\ \|D^{k+n-r}f\|\right),$$

and the substitution $j := k + n - r$ brings this into the form (7). □

With the definition (2) of Q_Δ, it follows that

$$\left|\int_0^1 f(1-Q_\Delta)g\right| = \left|\sum_{i=0}^{l-1}(\Delta t_i/2)\int_{-1}^{1}(S_i f)(1-Q)(S_i g)\right|$$
$$\leqslant \operatorname{const}_Q \sum_{i=0}^{l-1}(\Delta t_i/2)\sum_{j<n}\|D^j S_i f\|\ \|D^{k+n-j}S_i g\|$$

while, e.g.,

$$\|D^j S_i f\| = (\Delta t_i/2)^j\|D^j f\|_{(i)}.$$

Consequently, we have

COROLLARY 1. *If Q is a ssuper projector of order (k, n), then there exists* const_Q *so that for $f \in \times_0^{l-1}\mathbf{L}_\infty^{(n)}[t_i, t_{i+1}]$ and $g \in \times_0^{l-1}\mathbf{L}_\infty^{(k+n)}[t_i, t_{i+1}]$,*

$$\left|\int_0^1 f(1-Q_\Delta)g\right| \leqslant \operatorname{const}_Q \sum_{i=0}^{l-1}(\Delta t_i)^{1+k+n}\|f\|_{n,(i)}\|g\|_{k+n,(i)}$$

with

$$\|f\|_{r,(i)} := \max_{j<r}\|D^j f\|_{(i)}.$$

If now f and/or g in Lemma 9.1 are not as smooth as required, say, $f \in \mathbf{L}_\infty^{(n_f)}$, $g \in \mathbf{L}_\infty^{(k+n_g)}$, with n_f, $n_g \leqslant n$, then we are only entitled to consider $T_{k+n_g}g$ and $T_j f$ for $j \leqslant n_f$, hence, instead of (7), we get

$$\left|\int_{-1}^{1}f(1-Q)g\right| \leqslant \operatorname{const}_Q\sum_{j<n_f}\|D^j f\|\ \|D^{k+(n_g-j)_+}g\|. \tag{8}$$

Correspondingly, we get

COROLLARY 2. *If Q is a ssuper projector of order (k, n), then there exists* const *so that for $f \in \times_0^{l-1}\mathbf{L}_\infty^{(n_{f,i})}[t_i, t_{i+1}]$ and $g \in \times_0^{l-1}\mathbf{L}_\infty^{(k+n_{g,i})}[t_i, t_{i+1}]$ with $n_{f,i}$, $n_{g,i} \leqslant n$, all i,*

$$\left|\int_0^1 f(1-Q_\Delta)g\right| \leqslant \operatorname{const}\sum_{i=0}^{l-1}(\Delta t_i)^{1+k+\min\{n_{f,i},n_{g,i}\}}\|f\|_{n_{f,i},(i)}\|g\|_{k+n_{g,i},(i)}. \tag{9}$$

We now sketch proofs concerning the convergence of x_Δ satisfying (3) to x. From the proof of Theorem 3.1 in de Boor and Swartz [1], we find that x_Δ exists uniquely for $|\Delta|$ sufficiently small; that if $Mx \in C_\Delta[0, 1]$, then

$$\|D^r(x - x_\Delta)\|_\infty \leqslant \text{const}\, \omega_\Delta(Mx), \qquad 0 \leqslant r \leqslant m,$$

with $\omega_\Delta(f) := \sup_i \sup\{|f(t) - f(s)|: t_i < s, t < t_{i+1}\}$; and that if $x \in C^{(m-1)}[0, 1] \cap C_\Delta^{(m+k)}[0, 1]$, then

$$\|D^r(x - x_\Delta)\|_\infty \leqslant \text{const}|\Delta|^k \|x\|_{m+k,\Delta}, \qquad 0 \leqslant r \leqslant m; \qquad \|x\|_{j,\Delta} := \max_i \|x\|_{j,(i)}.$$

The proof of Lemma 4.1 in that paper, which uses this convergence of $x_\Delta =: Rx$ together with the Markov inequality for polynomials, yields additionally that

$$\|x_\Delta\|_{r,(i)} \leqslant \text{const}(|\Delta|/\Delta t_i)^k \|x\|_{m+k,\Delta}, \qquad r \geqslant 0.$$

The proof of superconvergence then goes as follows: for fixed $s \in [t_i, t_{i+1}]$ and for fixed $r < m$,

$$D^r(x - x_\Delta)(s) = \int_0^1 v(t)[M(x - x_\Delta)(t)]\, dt, \quad \text{where}$$

$$v(t) := (\partial^r G/\partial s^r)(s, t) \in \mathbf{L}_\infty^{(m-1-r)}[0, 1] \cap \big(C^{(n)}[0, s] \times C^{(n)}[s, 1]\big),$$

$G :=$ Green's function for M under suitable homogeneous side conditions $\boldsymbol{\beta}$.

Since

$$M(x - x_\Delta) = (1 - Q_\Delta)Mx + Q_\Delta M(x - x_\Delta) + (Q_\Delta - 1)Mx_\Delta,$$

with the second term vanishing by (3), the corollaries to Lemma 9.1 yield (uniformly in s)

$$|D^r(x - x_\Delta)(s)| \leqslant \text{const}\left[\sum_{i=0}^{l-1} (\Delta t_i)^{k+1+n(i)} \|v\|_{n(i),(i)} (\|Mx\|_{k+n,(i)} + \|Mx_\Delta\|_{k+n,(i)})\right],$$

where

$$n(i) := \begin{cases} \min\{m - 1 - r, n\} & \text{if } s \in (t_i, t_{i+1}), \\ n & \text{otherwise.} \end{cases}$$

Thus, we conclude the superconvergence rates of the following theorem which generalizes the collocation conclusions of [1, Theorem 4.1]:

THEOREM 9.2. *Let Q be a ssuper projector of order (k, n). Then, for sufficiently small $|\Delta|$, there exists $x_\Delta \in \mathbf{P}_{m+k,\Delta}^m$ satisfying* (3), *hence, then, the linear projector P_Δ given by the rule*

$$(9) \qquad Q_\Delta P_\Delta f = Q_\Delta f, \qquad P_\Delta f \in \{Mz: z \in \mathbf{P}_{m+k,\Delta}^m, \boldsymbol{\beta} z = 0\}$$

is well defined. Further, consider $x_\Delta \in \mathbf{P}_{m+k,\Delta}^m$ satisfying (3) *as an approximate solution to $Mx = y$, $\boldsymbol{\beta} x = \mathbf{0}$, where the coefficients of M lie in $C^{(n+k)}[0, 1]$ and the side conditions $\boldsymbol{\beta}$ are suitable. Then, uniformly in the maximum mesh size $|\Delta|$, we have the global estimates*

$$\|D^r(x - x_\Delta)\|_\infty \leqslant \text{const}\, \omega_\Delta(y), \qquad 0 \leqslant r \leqslant m;$$

$$\|D^r(x - x_\Delta)\|_\infty \leqslant \text{const}|\Delta|^{k+\min\{m-r,n\}} \|x\|_{m+k+\min\{m-r,n\},\Delta}, \qquad 0 \leqslant r \leqslant m;$$

while, uniformly over the breakpoints $(t_i)_0^l$ of Δ,

$$|D^r(x - x_\Delta)(t_i)| \leqslant \text{const}|\Delta|^{k+n} \|x\|_{m+k+n,\Delta}, \qquad 0 \leqslant r < m.$$

Remarks. Isolated solutions in $C^{(m+k+n)}[0, 1]$ to nonlinear problems can be handled as in [1, Theorem 3.1], where the question of superconvergence is reduced to the superconvergence associated with a linearized problem (which we have just settled).

We have left open so far the question of when the side conditions $\boldsymbol{\beta}$ are "suitable". Simply put, the side conditions are "suitable" if Green's function resulting from them allows the earlier argument to be made. If, for example, $\boldsymbol{\beta}$ consists of multipoint conditions, then one fixes a partition $\Delta_0 = (t_i^{(0)})_0^L$ of $[0, 1]$ whose partition points contain all the points involved in $\boldsymbol{\beta}$, and insists that all partitions Δ under consideration are refinements of Δ_0. Green's function for $(M, \boldsymbol{\beta})$ then satisfies

$$(\partial/\partial s)^r G(s, \cdot) \in C_{\Delta_0}^{(n)}[0, s] \times C_{\Delta_0}^{(n)}[s, 1], \tag{10}$$

and this is enough to complete the argument for x_Δ correspondingly in $\times_{i=1}^{L} \mathbf{P}_{m+k,\Delta}^m[t_{i-1}^{(0)}, t_i^{(0)}]$. In fact, it is easy to see now how to handle the more general situation in which we have differential operators of possibly different orders on the different intervals given by the partition Δ_0, with appropriate side conditions at the points of Δ_0 tying the pieces together.

Turning now to the eigenvalue problem, Corollary 1 of Lemma 9.1 is the general version promised in [2] of Lemma 3.1 there. It therefore permits the following generalization of Theorem 3.1 there.

THEOREM 9.3. *Let $T = NM^{-1}$ be the compact map on $\mathbf{L}_p[0, 1]$, $1 \leqslant p \leqslant \infty$, associated with the sufficiently smooth operators M, N, and $\boldsymbol{\beta}$ of* (0.2). *Let μ be a nonzero eigenvalue of T with corresponding invariant subspace S, and let J be a matrix representation for $T|_S$. Let $T_\Delta = P_\Delta T$, where P_Δ is the projector given by* (9) *associated with a ssuper projector Q of order (k, n). Then, for all small $|\Delta|$, T_Δ has an invariant subspace S_Δ, and $T_\Delta|_{S_\Delta}$ has a matrix representation J_Δ for which*

$$\|J - J_\Delta\| \leqslant \text{const}|\Delta|^{k+n}.$$

10. Ssuper Projectors of Order (k, k) Associated With Point Evaluations. We now look in more detail at the possible ssuper projectors of order (k, k). To begin with, we only consider their action on $\mathbf{P}_{2k}$, and this we can describe fully by specifying their action on the elements of some basis for $\mathbf{P}_{2k}$. We found it particularly convenient to work with the basis $(P_i)_0^{2k-1}$ consisting of the Legendre polynomials. Then, for any linear projector Q onto $\mathbf{P}_k$,

$$\left.\begin{aligned} QP_j &= P_j \\ QP_{k+j} &= \sum_{r=1}^{k} a_{jr} P_{k-r} \end{aligned}\right\}, \qquad j = 0, \ldots, k-1, \tag{1}$$

and *different projectors Q correspond to different matrices (a_{ij})*. Further, *two such projectors agree on $\mathbf{P}_{k+r}$ if and only if the corresponding matrices agree in rows* $0, \ldots, r-1$.

Let

$$f_j := \begin{cases} P_{j-1}, & j \leqslant k, \\ P_{j-1} - QP_{j-1}, & j > k. \end{cases}$$

Then (f_j) satisfies (9.5a–b), hence, with (9.5c), Q is ssuper of order (k, k) if and only if

$$\int_{-1}^{1} P_{i-1} \sum_{r=1}^{k} a_{j-k-1,r} P_{k-r} = 0 \quad \text{for } i \leqslant k < j \leqslant 2k + 1 - i.$$

This holds if and only if

$$a_{j-k-1,r} = 0 \quad \text{for} \quad i - 1 = k - r \quad \text{and} \quad i \leqslant k < j \leqslant 2k + 1 - i,$$

i.e., if and only if

$$a_{qr} = 0 \quad \text{for } 0 \leqslant q \leqslant r - 1.$$

We have proved

LEMMA 10.1. *The conditions*

$$Qf_{k+j} = 0 \quad \textit{with}$$

$$f_{k+1} := P_k \quad \textit{and} \quad f_{k+j} := P_{k+j-1} - \sum_{r=1}^{j-1} a_{j-1,r} P_{k-r}, \qquad j = 2, \ldots, k, \tag{2}$$

establish a one-to-one correspondence between ssuper projectors Q on $\mathbf{P}_{2k}$ and lower triangular matrices $(a_{ir})_{i,r=1}^{k-1}$.

Of course, any such ssuper projector Q on $\mathbf{P}_{2k}$ can be extended to infinitely many such on $C[-1, 1]$; and any such can be obtained in the form QP, with P an arbitrary linear projector on $C[-1, 1]$ onto $\mathbf{P}_{2k}$. We choose to ignore this aspect, though, since the property of being ssuper of order (k, k) depends only on the action on $\mathbf{P}_{2k}$.

Lemma 10.1 gives rise to several observations.

The first interesting basis function, viz. f_{k+1}, is simply the kth Legendre polynomial, P_k. Hence, if we think now of Q as being given by the rule

$$Qf \in \mathbf{P}_k \quad \text{and} \quad q_i^* Qf = q_i^* f, \qquad i = 1, \ldots, k, \tag{3}$$

for suitably chosen linear functionals $q_1^*, \ldots, q_k^*$, then we must have

$$q_i^* P_k = 0, \qquad i = 1, \ldots, k. \tag{4}$$

Now, in Example 1 (Collocation), we had

$$q_i^* f = f(\rho_i), \qquad i = 1, \ldots, k,$$

and so (4) is satisfied (for $n = k$ in Example 1) since then $(\rho_i)_1^k$ is simply the sequence of zeros of P_k. In Example 2 (Least squares),

$$q_i^* f = \int_{-1}^{1} P_{i-1} f, \qquad i = 1, \ldots, k,$$

and, again, (4) is satisfied since P_k is orthogonal to $\mathbf{P}_k = \operatorname{span}(P_{i-1})_1^k$. Suppose now that, in an attempt to bridge the gap between these two particular ssuper projectors, we look for ssuper projectors for which *some* of the interpolation conditions are point evaluations, say

$$q_i^* f = f(\sigma_i), \qquad i = 1, \ldots, r,$$

for some r. Then we conclude from (4) that $\{\sigma_1, \ldots, \sigma_r\}$ *must be a subset of* $\{\rho_1, \ldots, \rho_k\}$:= zeros of P_k. This leads us to consider

Example 3. *Ssuper projectors using point evaluations*. Let $(\rho_i)_1^k$ be the zeros of P_k in some order. Then, for $r = 0, \ldots, k$, the conditions

(5) $\quad Q_r f \in \mathbf{P}_k, \quad f - Q_r f \perp \mathbf{P}_r, \quad Q_r f(\rho_i) = f(\rho_i) \quad \text{for } i = r+1, \ldots, k,$

define a ssuper projector Q_r of order (k, k).

This provides us with a sequence of ssuper projectors of order (k, k), with Q_0, i.e., interpolation at the Gauss-Legendre points, at one end and Q_k, i.e., Least-squares approximation, at the other, and so demonstrates a perhaps surprisingly simple connection between the two.

We now verify Example 3. In order to confirm that (5) defines a linear projector Q_r, we note that the conditions mentioned are equivalent to demanding that

$$Q_r f \in \mathbf{P}_k, \quad q_i^* Q_r f = q_i^* f, \qquad i = 1, \ldots, k,$$

with

$$q_i^* f = \begin{cases} \int_{-1}^{1} P_{i-1} f, & i = 1, \ldots, r, \\ f(\rho_i), & i = r+1, \ldots, k. \end{cases} \tag{6}$$

Thus it suffices to show that the matrix

$$(q_i^* P_{j-1})_{i,j=1}^k \tag{7}$$

is invertible. For this, assume that this matrix maps $\mathbf{a} = (a_i)_1^k$ to $\mathbf{0}$, i.e.,

$$q_i^* p = 0, \quad i = 1, \ldots, k \quad \text{with } p := \sum_{j=1}^{k} a_j P_{j-1}. \tag{8}$$

Let (w_i) be the weight vector (known to be strictly positive) for the corresponding quadrature rule

$$\int_{-1}^{1} f = \sum_{j=1}^{k} w_j f(\rho_j), \quad \text{all } f \in \mathbf{P}_{2k}.$$

Then (6) and (8) imply that

$$0 = \int_{-1}^{1} p P_{i-1} = \sum_{j=1}^{r} w_j p(\rho_j) P_{i-1}(\rho_j), \qquad i = 1, \ldots, r,$$

which shows that the invertible matrix $(P_{i-1}(\rho_j))_{i,j=1}^r$ maps the vector $(w_j p(\rho_j))_1^r$ to $\mathbf{0}$, and consequently p not only vanishes at $\rho_{r+1}, \ldots, \rho_k$ (by (6) and (8)) but also $p(\rho_j) = 0$ for $j = 1, \ldots, r$. Thus $p = 0$, and so $\mathbf{a} = \mathbf{0}$.

Note that the invertibility of (7) just proven implies the invertibility of

$$(P_{j-1}(\rho_i))_{i,j=r+1}^k \tag{9}$$

since $q_i^* P_{j-1} = \int P_{i-1} P_{j-1} = 0$ for $i \leqslant r < j$.

To verify that Q_r is ssuper (a fact not immediately obvious to us), we now show that Q_r can be obtained from Q_0 by a suitable modification. For this, we need to consider these projectors on $\mathbf{P}_{2k+1}$ (on which the ssuper projectors of order (k, k) form a $k(k+1)/2$-dimensional hyperplane). Let

$$f_j^{[r]} := \begin{cases} P_{j-1}, & j = 1, \ldots, k, \\ (1 - Q_r) P_{j-1}, & j = k+1, \ldots, 2k+1, \end{cases}$$

hence,

$$f^{[r]}_{k+j} = P_{k+j-1} - \sum_{s=1}^{k} a^{[r]}_{j-1,s} P_{k-s}, \qquad j = 1, \ldots, k+1,$$

for some matrix $(a^{[r]}_{j-1,s})^{k+1,k}_{j,s=1}$. With this notation, we get, for $r = 0$,

$$f^{[0]}_{k+j}(\rho_i) = 0, \qquad i = 1, \ldots, k, \qquad j = 2, \ldots, k+1,$$

and so

$$\begin{bmatrix} a^{[0]}_{11} & & 0 \\ \vdots & \ddots & \\ a^{[0]}_{k1} & \cdots & a^{[0]}_{kk} \end{bmatrix} \begin{bmatrix} P_{k-1}(\rho_k) & \cdots & P_{k-1}(\rho_1) \\ \vdots & & \vdots \\ P_{k-k}(\rho_k) & \cdots & P_{k-k}(\rho_1) \end{bmatrix} = \begin{bmatrix} P_{k+1}(\rho_k) & \cdots & P_{k+1}(\rho_1) \\ \vdots & & \vdots \\ P_{k+k}(\rho_k) & \cdots & P_{k+k}(\rho_1) \end{bmatrix} \tag{10}$$

with $(a^{[0]}_{ij})$ a lower triangular matrix (by Lemma 10.1) since Q_0 is ssuper. Now write (10) in terms of partitioned matrices as

$$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \tag{10$'$}$$

with A_{11}, B_{11}, and C_{11} all of order $k - r$. Our intent is to replace A_{22} by 0 and to modify A_{21} correspondingly in such a way that the equality in (10) or (10)$'$ is preserved at least in the first $k - r$ columns. Explicitly,

$$C_{21} = A_{21}B_{11} + A_{22}B_{21} = \left[A_{21} + A_{22}B_{21}B_{11}^{-1}\right]B_{11}$$

(and B_{11} is indeed invertible since it is just a permutation of the matrix (9)). Thus

$$\begin{bmatrix} A_{11} & 0 \\ \tilde{A}_{21} & 0 \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & \tilde{C}_{22} \end{bmatrix}, \tag{11}$$

with

$$\tilde{A}_{21} := A_{21} + A_{22}B_{21}B_{11}^{-1}, \qquad \tilde{C}_{22} := \tilde{A}_{21}B_{12}.$$

Now consider the linear projector Q on $\mathbf{P}_{2k+1}$ given by

$$Qf_i = \begin{cases} f_i, & i \leqslant k, \\ 0, & i > k, \end{cases} \tag{12}$$

with

$$\begin{aligned} f_i &:= P_{i-1}, \qquad i \leqslant k, \\ f_{k+1} &:= P_k, \qquad\qquad \text{and } (a_{ij}) := \begin{bmatrix} A_{11} & 0 \\ \tilde{A}_{21} & 0 \end{bmatrix}. \\ f_{k+i} &:= P_{k+i-1} - \sum_{j=1}^{k} a_{i-1,j} P_{k-j}, \qquad i > 1, \end{aligned} \tag{13}$$

Then (a_{ij}) is lower triangular. Hence Q is a ssuper projector of order (k, k) by Lemma 10.1. Further, on comparing (10)–(10)′ with (11), we see that, for the linear functionals q_i^* of (6),

$$q_j^* f_{k+i} = f_{k+i}(\rho_j) = 0 \quad \text{for } i = 2, \dots, k+1, \ \ j = r+1, \dots, k,$$

while the fact that the last r columns of (a_{ij}) are zero implies that $f_{k+i} \perp P_r$, $i = 2, \dots, k+1$, i.e., also

$$q_j^* f_{k+i} = 0 \quad \text{for } i = 2, \dots, k+1, \ \ j = 1, \dots, r.$$

In addition, trivially, $q_j^* f_{k+1} = 0$, all j. We conclude that $\ker Q|_{\mathbf{P}_{2k+1}} = \operatorname{span}(f_{k+j})_1^{k+1}$ is contained in $\bigcap \ker q_j^*|_{\mathbf{P}_{2k+1}}$ and thus must equal it since both are of dimension $k+1$. This shows that $Q = Q_r$ (on $\mathbf{P}_{2k+1}$), i.e.,

$$(a_{ij}^{[r]}) = \begin{bmatrix} A_{11} & 0 \\ A_{21} + A_{22}B_{21}B_{11}^{-1} & 0 \end{bmatrix}. \tag{14}$$

We have established, in particular, that Q_r is ssuper of order (k, k). In addition, comparing again (10)–(10)′ with (11), we see that Q_r agrees with Q_0 on $\operatorname{span}(f_i)_{i=1}^{2k+1-r} = \mathbf{P}_{2k+1-r}$. Thus, we could not tell Q_1 and Q_0 apart on $\mathbf{P}_{2k}$.

We close our discussion of Example 3 with the following four observations.

(i) *For $r \neq s$, Q_r differs from Q_s somewhere on* $\mathbf{P}_{2k+2-\max\{r,s\}}$ (while, as we just noted, the two agree with Q_0, hence with each other, on $\mathbf{P}_{2k+1-\max\{r,s\}}$). For the proof, apply both sides of (10) to the matrix

$$\left(w_{k+1-i} P_{k-j}(\rho_{k+1-i})\right)_{i,j=1}^{k}.$$

Then, assuming that the Legendre polynomials have all been normalized to

$$1 = \int_{-1}^{1} P_j^2 = \sum w_i P_j^2(\rho_i),$$

we find that

$$(a_{ij}^{[0]}) = \left(\sum_{s=1}^{k} w_s P_{k+i}(\rho_s) P_{k-j}(\rho_s) \right) \tag{15}$$

and, in particular,

$$\begin{aligned} a_{ii}^{[0]} &= \sum_{s=1}^{k} w_s P_{k+i}(\rho_s) P_{k-i}(\rho_s) \\ &= \int_{-1}^{1} P_{k+i} P_{k-i} + \operatorname{const}_k D^{2k}(P_{k+i} P_{k-i}) \neq 0. \end{aligned} \tag{16}$$

This shows with (14) that $\operatorname{rank}(a_{ij}^{[r]}) = k - r$, all r, and so proves our assertion.

(ii) The agreement of Q_r with Q_0 on $\mathbf{P}_{2k+1-r}$ is not restricted to the particular ssuper projector Q_r. If Q is *any* ssuper projector of order (k, k) which enforces agreement at $k - r$ points, then, not only must the $k - r$ points all be zeros of P_k, say the points $\rho_{r+1}, \dots, \rho_k$ (in some suitable ordering), but *such Q then necessarily agrees with Q_0 on* $\mathbf{P}_{2k+1-r}$. For, by Lemma 10.1, Q satisfies (2) for some lower triangular matrix. The matching of function values at $\rho_{r+1}, \dots, \rho_k$ then forces the equality

$$\begin{bmatrix} a_{11} & & 0 \\ \vdots & \ddots & \\ a_{k-r,1} & \cdots & a_{k-r,k-r} \end{bmatrix} \begin{bmatrix} P_{k-1}(\rho_k) & \cdots & P_{k-1}(\rho_{r+1}) \\ \vdots & & \vdots \\ P_{k-r}(\rho_k) & \cdots & P_{k-r}(\rho_{r+1}) \end{bmatrix}$$
$$= \begin{bmatrix} P_{k+1}(\rho_k) & \cdots & P_{k+1}(\rho_{r+1}) \\ \vdots & & \vdots \\ P_{2k-r}(\rho_k) & \cdots & P_{2k-r}(\rho_{r+1}) \end{bmatrix}.$$

In the terms of (10)′, this reads

$$(a_{ij})_{i,j=1}^{k-r} B_{11} = C_{11}$$

and the invertibility of B_{11} used earlier now proves that therefore

$$(a_{ij})_{i,j=1}^{k-r} = A_{11} = \left(a_{ij}^{[0]}\right)_{i,j=1}^{k-r}.$$

(iii) The sequence $Q_1, \ldots, Q_{k-1}$ connecting Q_0 to Q_k constructed in Example 3 depends on the particular order $\rho_1, \ldots, \rho_k$ in which we have chosen to write down the k zeros of P_k. *If* (Q_r') *is the sequence corresponding to the ordering* $\rho_1', \ldots, \rho_k'$, *then* $Q_r = Q_r'$ *on* $\mathbf{P}_{2k+1}$ *if and only if the two sets* $\{\rho_{r+1}, \ldots, \rho_k\}$ *and* $\{\rho_{r+1}', \ldots, \rho_k'\}$ *coincide*. Indeed, from (14), $Q_r = Q_r'$ on $\mathbf{P}_{2k+1}$ if and only if

$$A_{21} + A_{22}B_{21}B_{11}^{-1} = A_{21}' + A_{22}'B_{21}'(B_{11}')^{-1}.$$

Now, obviously, $A_{ij} = A_{ij}'$, –(15) makes this quite explicit, but it is clear anyway since Q_0 does not depend on the order in which we write down the interpolation points,– and A_{22} is invertible, e.g., by (16). Thus, $Q_r' = Q_r$ on $\mathbf{P}_{2k+1}$ if and only if $D := B_{21}'(B_{11}')^{-1} - B_{21}B_{11}^{-1} = 0$. Let Π be the permutation matrix for which $B' = B\Pi$. Then, on partitioning Π as B is in (10)′, we find

$$B_{11}' = B_{11}\Pi_{11} + B_{12}\Pi_{21}, \qquad B_{21}' = B_{21}\Pi_{11} + B_{22}\Pi_{21},$$

or

$$\begin{aligned} DB_{11}' &= B_{21}' - B_{21}B_{11}^{-1}B_{11}' \\ &= B_{21}\Pi_{11} + B_{22}\Pi_{21} - B_{21}B_{11}^{-1}(B_{11}\Pi_{11} + B_{12}\Pi_{21}) \\ &= \left(B_{22} - B_{21}B_{11}^{-1}B_{12}\right)\Pi_{21}. \end{aligned}$$

Now note that $B_{22} - B_{21}B_{11}^{-1}B_{12}$ is the lower right diagonal block obtained by block Gauss elimination applied to B, and hence is invertible (since B is). We conclude that $D = 0$ if and only if $\Pi_{21} = 0$, and that says that Π permutes the first $k - r$ columns of B among themselves.

(iv) Finally, we observed earlier that the collection of all ssuper projectors of order (k, k) on $\mathbf{P}_{2k+1}$ forms a linear manifold or hyperplane of dimension $k(k + 1)/2$. We now show that this linear manifold is spanned by the particular ssuper projectors Q_r introduced here. Precisely, we show that *the collection of all ssuper projectors of order* (k, k) *on* $\mathbf{P}_{2k+1}$ *is the affine hull of the* $1 + k(k + 1)/2$ *particular projectors*

$$Q_0, Q_{11}, \ldots, Q_{1k}, Q_{22}, \ldots, Q_{2k}, Q_{33}, \ldots, Q_{k-1,k}, Q_{kk},$$

with Q_0 interpolation at all the zeros $(\rho_i)_1^k$ of P_k, while, for $1 \leqslant r \leqslant s \leqslant k$, Q_{rs} is given by orthogonality to $\mathbf{P}_r$ and matching of function values at the $k - r$ points $\rho_r, \ldots, \rho_{s-1}, \rho_{s+1}, \ldots, \rho_k$. Here, (ρ_i) are in any particular order; in fact, in the definition of Q_{rs}, this order could even change with r (though not with s). To prove the assertion, it is sufficient to show that the $k(k+1)/2$ linear maps $\{Q_{rs} - Q_0: 1 \leqslant r \leqslant s \leqslant k\}$ are linearly independent (as points in the linear space of all linear maps on $\mathbf{P}_{2k+1}$) and for this, it is sufficient to exhibit points x_{ij} in $\mathbf{P}_{2k+1}$ and linear functionals μ_{ij} on $\mathbf{P}_{2k+1}$ for which

$$\begin{aligned} \mu_{ij}(Q_{rs} - Q_0)x_{ij} &\neq 0 \quad \text{for } (i,j) = (r,s), \\ &= 0 \quad \text{for } i > r, \text{ and for } i = r \text{ and } j > s, \end{aligned}$$

since this insures that the matrix $(\mu_{ij}(Q_{rs} - Q_0)x_{ij})$ is upper triangular with nonzero diagonal entries (using the ordering $11, 12, \ldots, 1k, 22, \ldots, 2k, 33, \ldots, kk$), hence invertible. (We are using here the standard argument whereby the sequence (y_s) in a linear space is linearly independent if and only if there exists a corresponding sequence (ν_r) of linear functionals on that space for which the matrix $(\nu_r y_s)$ is invertible.) First, pick $x_{ij} = P_{2k+1-i}$, all i, j. Then, since Q_{rs} forces agreement at $k - r$ points, it agrees, by (ii), with Q_0 on $\mathbf{P}_{2k+1-r}$ and so $\mu_{ij}(Q_{rs} - Q_0)x_{ij} = 0$ for $2k + 1 - i < 2k + 1 - r$, i.e., for $i > r$ no matter how we pick μ_{ij}. Further, pick $\mu_{ij}: f \mapsto f(\rho_j)$, all i, j. Then, as both Q_0 and Q_{rs} match the value at ρ_j when $r \leqslant j \neq s$, we conclude that $\mu_{ij}(Q_{rs} - Q_0)x_{ij} = 0$ also for $i = r$ and $j > s$ $(\geqslant r)$. Finally, we claim that $\mu_{ij}(Q_{ij} - Q_0)x_{ij} \neq 0$. For, otherwise, $Q_{ij}P_{2k+1-i}$ would agree with P_{2k+1-i} at $\rho_i, \ldots, \rho_k$ as well as at the linear functionals $f \mapsto \int P_{r-1} f$, $r = 1, \ldots, i-1$, i.e., Q_{ij} would agree with $Q_{i-1,i-1}$ at P_{2k+1-i} and this would contradict (i).

Finally, up to this point, this section has been concerned with ssuper projectors of order (k, k). But we think it worth recording a version of Lemma 10.1 for ssuper projectors Q of order (k, n), $1 \leqslant n \leqslant k$, along with a corresponding corollary concerning the k linear functionals $(q_i^*)_1^k$ associated with Q.

LEMMA 10.2. *The conditions*

$$Qf_{k+j} = 0 \quad \text{with } f_{k+j} := P_{k+j-1} - \sum_{r=1}^{k-n+j-1} a_{j-1,r}P_{k-r}, \quad j = 1, \ldots, N, \tag{17}$$

establish a one-to-one correspondence between ssuper projectors Q of order (k, n) on $\mathbf{P}_{k+N}$, $N \geqslant n$, and lower trapezoidal matrices

$$(a_{qr})_{q=0}^{N-1}\,{}_{r=1}^{k}, \quad \text{with } a_{qr} = 0 \text{ for } k - n + q < r \leqslant k. \tag{18}$$

In terms of linear functionals $(q_i^*)_1^k$ associated with Q *via* (3), the ssuper projector criterion (9.4) may be expressed as

$$P_i \perp \left(P_{k+n+1-i} \cap \bigcap_{r=1}^{k} \ker q_r^* \right), \qquad i = 1, \ldots, n;$$

and Lemma 10.2 may be restated as

COROLLARY. *The linear functionals* $(q_i^*)_1^k$ *define a ssuper projector* Q *of order* (k, n) *on* $\mathbf{P}_{k+N}$, $N \geqslant n$, *if and only if the two blocks of the* $(k + N) \times k$ *matrix*

$$\begin{bmatrix} B_{k\times k} \\ C_{N\times k} \end{bmatrix} \quad \text{with } b_{ij} := q_j^* P_{k-1-i}, \quad 0 \leqslant i < k, \quad \text{and} \quad c_{ij} := q_j^* P_{k+i}, \quad 0 \leqslant i < N,$$

(whose transpose describes the action of the functionals with respect to the basis $(P_{k-1}, \ldots, P_0, P_k, \ldots, P_{k+N-1})$ of $\mathbf{P}_{k+N}$) *satisfy*

$$CB^{-1} = A,$$

where the matrix $A = (a_{ij})_{i=0}^{N-1}\ {}_{j=1}^{k}$ *has the lower trapezoidal form* (18). *Two sets of linear functionals define distinct ssuper projectors precisely to the extent that the corresponding matrices* A *are distinct.*

Proof. Q is a projector onto $\mathbf{P}_k$ if and only if B is nonsingular. Q is ssuper of order (k, n) on $\mathbf{P}_{k+N}$ if and only if the matrix A of (18) is connected with Q *via* (17). The $(j-1, i)$ element of the assertion $C - AB = 0$ is found by applying q_i^* to (17).

Mathematics Research Center
University of Wisconsin-Madison
610 Walnut Street
Madison, Wisconsin 53706

Theoretical Division
University of California
Los Alamos Scientific Laboratory
Los Alamos, New Mexico 87545

1. C. DE BOOR & B. SWARTZ, "Collocation at Gaussian points," *SIAM J. Numer. Anal.*, v. 10, 1973, pp. 582–606.

2. C. DE BOOR & B. SWARTZ, "Collocation approximation to eigenvalues of an ordinary differential equation: The principle of the thing," *Math. Comp.*, v. 35, 1980, pp. 679–694.

3. C. DE BOOR & B. SWARTZ, "Collocation approximation to eigenvalues of an ordinary differential equation: Numerical illustrations," *Math. Comp.*, v. 36, 1981, pp.

4. S. A. PRUESS, "Solving linear boundary value problems by approximating the coefficients," *Math. Comp.*, v. 27, 1973, pp. 551–561.

S. A. PRUESS, "High order approximation to Sturm-Liouville eigenvalues," *Numer. Math.*, v. 24, 1975, pp. 241–247.

6. K. A. WITTENBRINK, "High order projection methods of moment and collocation type for nonlinear boundary value problems," *Computing*, v. 11, 1973, pp. 255–274.