Numer. Math. 21, 109–129 (1973)

Transfinite Element Methods: Blending-Function Interpolation over Arbitrary Curved Element Domains

William J. Gordon

Mathematics Department, Research Laboratories, General Motors Corporation
Warren, Michigan

Charles A. Hall

Department of Mathematics, University of Pittsburgh, Pittsburgh, Pennsylvania

Received April 17, 1972

Abstract. In order to better conform to curved boundaries and material interfaces, curved finite elements have been widely applied in recent years by practicing engineering analysts. The most well known of such elements are the "isoparametric elements." As Zienkiewicz points out in [18, p. 132] there has been a certain parallel between the development of "element types" as used in finite element analyses and the independent development of methods for the mathematical description of general free-form surfaces. One of the purposes of this paper is to show that the relationship between these two areas of recent mathematical activity is indeed quite intimate. In order to establish this relationship, we introduce the notion of a "transfinite element" which, in brief, is an invertible mapping $\vec{T}$ from a square parameter domain $\mathscr{S}$ onto a closed, bounded and simply connected region $\mathscr{R}$ in the xy-plane together with a "transfinite" blending-function type interpolant to the dependent variable f defined over $\mathscr{R}$. The "subparametric," "isoparametric" and "superparametric" element types discussed by Zienkiewicz in [18, pp. 137–138] can all be shown to be special cases obtainable by various discretizations of transfinite elements Actual error bounds are derived for a wide class of semi-discretized transfinite elements (with the nature of the mapping $\vec{T}: \mathscr{S} \to \mathscr{R}$ remaining unspecified) as applied to two types of boundary value problems. These bounds for semi-discretized elements are then specialized to obtain bounds for the familiar isoparametric elements. While we consider only two dimensional elements, extensions to higher dimensions is straightforward.

1. Introduction

The conventional finite element method (FEM) involves the partitioning of a *polygonal* domain $\mathscr{R}$ into rectangular and/or triangular elements. Quite often, however, a structural engineer is faced with a boundary value problem over a *nonpolygonal* domain $\mathscr{R}$. The first step in a finite element solution then requires that the boundary, $\partial\mathscr{R}$, of $\mathscr{R}$ be *approximated* by a polygonal arc. Obviously, the accuracy of the FEM is limited by the accuracy of the polygonal approximation to $\partial\mathscr{R}$. At worst, the convergence of the FEM may be destroyed if boundary conditions are not handled properly. For example, see [2] for a discussion of what has been termed the "Babuska paradox."

Ergatoudis, Irons and Zienkiewicz [6, 17] have developed various curved parametric "finite elements" which circumvent the above dilemma. In [9] and below, the authors extend the notion of parametric elements and remove the "element of chance" referred to in [17, p. 382] from the derivation of the bivariate (and multivariate) interpolation formulae which underlie the finite dimensional parametric techniques.

In Section 2, we introduce the concept of *transfinite interpolation* by which we mean interpolation methods which match a given function on a nondenumerable set of points. We then use such transfinite formulas in Section 4 to generate certain *finite parameter* interpolation formulae as cases of special interest. These serve as the basis for the Ritz spaces involved in the numerical implementation of the *transfinite element method*. Discretization error bounds for these numerical methods and for two specific boundary value problems are developed in Section 5. As special cases of these results for the Lagrange transfinite elements, we obtain error bounds for isoparametric elements. Finally, in Section 6 we explore *isoparametric transfinite elements* in which the same transfinite interpolation formula is used *both* for the mapping of the independent variables and the approximation of the dependent variable. The finite parameter isoparametric elements of [17, 18] precipitate as special cases.

In another paper [9], we consider the application of transfinite interpolation methods to domain mapping and mesh generation problems associated with the geometric aspects of the finite element method.

2. Transfinite Interpolation Schemes

In this section we review a class of interpolation schemes which will serve later as the basis for subsequent analyses. We refer to this class of interpolation formulae as being transfinite, since their *precision sets* (i.e., the set of points in the domain of the independent variables s, t on which the interpolant matches the original function) are nondenumerable. In particular, the transfinite schemes which we consider are of the type referred to as "blending-function methods" [8]. Approximation error bounds are also given for several special but important cases; namely, the *transfinite Lagrange interpolation schemes.*

Let f be a continuous function of two independent variables with domain $\mathcal{S}$: $[0, h] \times [0, h]$ in the st-plane. We seek approximations $\tilde{f} \approx f$ which interpolate f on certain (denumerable or nondenumerable) point sets contained in $\mathcal{S}$. To accomplish this, we will rely on the algebraic theory of multivariate approximation developed in [7].

By a *projector*, $\mathcal{P}$, we mean an idempotent linear operator from the linear space $\mathcal{F}$ of all continuous bivariate functions f, with domain $\mathcal{S}$, onto a subspace of functions. For example, if the operator $\mathcal{P}_s$ is defined by the formula

$$\mathcal{P}_s[f] = (1 - s/h)\, f(0, t) + (s/h)\, f(h, t), \tag{1}$$

then $\mathcal{P}_s$ is both linear:

$$\mathcal{P}_s[f + g] = \mathcal{P}_s[f] + \mathcal{P}_s[g]$$

and idempotent:

$$\begin{aligned}\mathscr{P}_s \mathscr{P}_s [f] &= (1 - s/h)\, \mathscr{P}_s [f]|_{s=0} + (s/h)\, \mathscr{P}_s [f]|_{s=h} \\ &= (1 - s/h)\, f(0, t) + (s/h)\, f(h, t) \\ &= \mathscr{P}_s [f].\end{aligned}$$

In the expression for the projector $\mathscr{P}_s$, s is the operational variable and t is essentially a parameter. It is easy to verify that the projection $\mathscr{P}_s [f]$ interpolates f along the lines $s=0$ and $s=h$.

Throughout this paper, we restrict our attention to the class of projectors of the following type:

$$\mathscr{P}_s [f] = \sum_{i=0}^{m} f(s_i, t)\, \varphi_i(s) \tag{2}$$

where $0=s_0<s_1<\cdots<s_m=h$ and

$$\varphi_i(s) = \prod_{j\neq i} (s - s_j) / \prod_{j\neq i} (s_i - s_j), \qquad 0 \leq i \leq m \tag{3}$$

are the fundamental (cardinal) functions for *Lagrange polynomial interpolation* [5]. In the context of transfinite interpolation, the functions $\{\varphi_i(s)\}_{i=0}^{m}$ are called the *blending functions* [8]. The projection $\mathscr{P}_s [f]$ interpolates to f along the m lines $s=s_i$, $0\leq i\leq m$, in the s t-plane. Since $\mathscr{P}_s [f]$ coincides with f at a nondenumerable number of points, this is a simple example of a transfinite interpolation scheme. However, we shall reserve the term "transfinite bivariate Lagrange interpolation" for the more general class of methods of Theorem 1.

For completeness and later reference, we display the analogous formula for $\mathscr{P}_t [f]$:

$$\mathscr{P}_t [f] = \sum_{j=0}^{n} f(s, t_j)\, \psi_j(t), \tag{4}$$

where $0=t_0<t_1<\cdots<t_n=h$ and

$$\psi_j(t) = \prod_{i\neq j} (t - t_i) / \prod_{i\neq j} (t_j - t_i), \qquad 0 \leq j \leq n. \tag{5}$$

Probably the most well-known class of formulae for bivariate interpolation-approximation are the (tensor product) *bipolynomial Lagrange interpolation formulae*. This class of formulae is obtained as the *product* of the above defined projectors $\mathscr{P}_s$ and $\mathscr{P}_t$ in (2) and (4), respectively:

$$\begin{aligned}\mathscr{P}_s \mathscr{P}_t [f] &\equiv \mathscr{P}_s [\mathscr{P}_t [f]] \\ &= \sum_{i=0}^{m} \sum_{j=0}^{n} f(s_i, t_j)\, \varphi_i(s)\, \psi_j(t).\end{aligned} \tag{6}$$

The *product operator* $\mathscr{P}_s \mathscr{P}_t$ is itself a projector and $\mathscr{P}_s \mathscr{P}_t [f]$ interpolates f at the $(m+1)(n+1)$ points (s_i, t_j), $0\leq i\leq m$, $0\leq j\leq n$.

Nota bene: Although the precision sets of $\mathscr{P}_s$ and $\mathscr{P}_t$ themselves are transfinite, the precision set of $\mathscr{P}_s \mathscr{P}_t$ is a finite point set.

There is a second and stronger way to compound the projectors $\mathscr{P}_s$ and $\mathscr{P}_t$. The *Boolean sum* [7]

$$\mathscr{P}_s \oplus \mathscr{P}_t \equiv \mathscr{P}_s + \mathscr{P}_t - \mathscr{P}_s \mathscr{P}_t \tag{7}$$

is also a projector and serves as a basis for the following result.

Theorem 1. *Transfinite Bivariate Lagrange Interpolation.*

Let $\mathscr{P}_s$ and $\mathscr{P}_t$ be defined as above, then

$$\mathscr{P}_s \oplus \mathscr{P}_t[f] = \mathscr{P}_s[f] + \mathscr{P}_t[f] - \mathscr{P}_s \mathscr{P}_t[f] \tag{8}$$

interpolates f along the lines $s = s_i$, $0 \leqq i \leqq m$ and $t = t_j$, $0 \leqq j \leqq n$.

Proof. Use the expressions (2), (4) and (6) to verify that

$$\mathscr{P}_s \oplus \mathscr{P}_t[f](s_i, t) = f(s_i, t), \quad 0 \leqq i \leqq m$$

$$\mathscr{P}_s \oplus \mathscr{P}_t[f](s, t_j) = f(s, t_j), \quad 0 \leqq j \leqq n. \quad \text{Q.E.D.}$$

Various extensions and generalizations of Theorem 1 are immediate. For example, the theorem remains valid if the projector $\mathscr{P}_s$ is taken to be the *cubic spline interpolation projector* in the operational variable s and $\mathscr{P}_t$ is taken to be a *trigonometric polynomial interpolation* projector [7]. All that is really essential is that the functions $\varphi_i(s)$ and $\psi_j(t)$ in (2) and (4) satisfy the cardinality conditions $\varphi_i(s_k) = \delta_{ik}$ $(i, k = 0, 1, \ldots, m)$ and $\psi_j(t_l) = \delta_{jl}$ $(j, l = 0, 1, \ldots, n)$. Also, one need not restrict the domain of f to be a square. In [1] for instance, transfinite interpolation formulae over triangular domains were derived.

We now develop bounds on the approximation error associated with *transfinite bivariate Lagrange* interpolation schemes. Let $\mathscr{R}_s \equiv I - \mathscr{P}_s$ and $\mathscr{R}_t \equiv I - \mathscr{P}_t$ (where I is the identity operator) be the remainder projectors [7] associated with $\mathscr{P}_s$ and $\mathscr{P}_t$ respectively. The remainder associated with the *product operator* $\mathscr{P}_s \mathscr{P}_t$ is

$$\begin{aligned} I - \mathscr{P}_s \mathscr{P}_t &= I - (I - \mathscr{R}_s)(I - \mathscr{R}_t) \\ &= \mathscr{R}_s + \mathscr{R}_t - \mathscr{R}_s \mathscr{R}_t. \end{aligned} \tag{9}$$

Similarly, the remainder associated with the transfinite *Boolean sum* operator $\mathscr{P}_s \oplus \mathscr{P}_t$ is

$$\begin{aligned} I - \mathscr{P}_s \oplus \mathscr{P}_t &= I - (I - \mathscr{R}_s) \oplus (I - \mathscr{R}_t) \\ &= \mathscr{R}_s \mathscr{R}_t. \end{aligned} \tag{10}$$

We now seek bounds on $\mathscr{R}_s \mathscr{R}_t[f]$, or equivalently $f - \mathscr{P}_s \oplus \mathscr{P}_t[f]$, and its various derivatives. The analogous results for $f - \mathscr{P}_s \mathscr{P}_t[f]$ will follow directly from our development in Section 4.

Theorem 2. *Let $f \in C^{(m+1, n+1)}(\mathscr{S})$ and $\mathscr{P}_s \oplus \mathscr{P}_t[f]$ be the transfinite bivariate Lagrange interpolant defined by* (8). *There exist constants ε_{mk} and ε_{nl} such that*

$$\begin{gathered} \|(f - \mathscr{P}_s \oplus \mathscr{P}_t[f])^{(k,l)}\|_\infty \leqq \varepsilon_{mk}\, \varepsilon_{nl} \|f^{(m+1, n+1)}\|_\infty\, h^{m+n+2-k-l} \\ 0 \leqq k \leqq m \quad \text{and} \quad 0 \leqq l \leqq n. \end{gathered} \tag{11}$$

[Here and below $g^{(k,l)} \equiv \partial^{k+l} g / \partial s^k \partial t^l$.]

Proof. Let $D^{(p,q)}$ be the operator defined by

$$D^{(p,q)}[g] = g^{(p,q)}.$$

Then, from (2) and the above hypothesis on f, it follows that $\mathscr{P}_s$ and $D^{(0,l)}$ commute

$$D^{(0,l)} \mathscr{P}_s = \mathscr{P}_s D^{(0,l)}, \qquad 0 \leqq l \leqq n+1. \tag{12}$$

That is, fix t (arbitrarily), then for all s

$$(\mathscr{P}_s[f])^{(0,l)}(s,t) = \mathscr{P}_s[f^{(0,l)}](s,t).$$

With $\mathscr{R}_s \equiv I - \mathscr{P}_s$ we also have that

$$\begin{aligned} D^{(0,l)}\mathscr{R}_s &= D^{(0,l)} - D^{(0,l)}\mathscr{P}_s \\ &= D^{(0,l)} - \mathscr{P}_s D^{(0,l)} = \mathscr{R}_s D^{(0,l)}. \end{aligned} \tag{13}$$

Hence $D^{(0,l)}$ and $\mathscr{P}_s$ commute also, for $0 \leqq l \leqq n+1$.

From (13) and from [11, p. 289], it follows that for $0 \leqq k \leqq m$

$$\begin{aligned} \|D^{(k,l)}\mathscr{R}_s[f]\|_\infty &= \|D^{(k,0)}\mathscr{R}_s[f^{(0,l)}]\|_\infty \\ &\leqq \varepsilon_{mk}\|f^{(m+1,l)}\|_\infty h^{m+1-k} \end{aligned} \tag{14}$$

where $\varepsilon_{mk} \equiv 1/(m-k+1)!$.

Now $\mathscr{P}_t - \mathscr{P}_s\mathscr{P}_t = \mathscr{P}_t - \mathscr{P}_t\mathscr{P}_s = \mathscr{P}_t\mathscr{R}_s$ yields

$$\begin{aligned} (f - \mathscr{P}_s \oplus \mathscr{P}_t[f])^{(k,l)} &= (f - \mathscr{P}_s[f])^{(k,l)} - (\mathscr{P}_t[f] - \mathscr{P}_s\mathscr{P}_t[f])^{(k,l)} \\ &= D^{(0,l)}\{D^{(k,0)}\mathscr{R}_s[f] - D^{(k,0)}\mathscr{P}_t\mathscr{R}_s[f]\} \\ &= D^{(0,l)}\{D^{(k,0)}\mathscr{R}_s[f] - \mathscr{P}_t D^{(k,0)}\mathscr{R}_s[f]\} \end{aligned} \tag{15}$$

where we have used the fact that $\mathscr{P}_t D^{(k,0)} = D^{(k,0)}\mathscr{P}_t$.

But the last display line in (15) is the l-th derivative with respect to t of the error in Lagrange interpolation in the t variable to the function $D^{(k,0)}\mathscr{R}_s[f]$. That is,

$$(f - \mathscr{P}_s \oplus \mathscr{P}_t[f])^{(k,l)} = D^{(0,l)}\mathscr{R}_t[D^{(k,0)}\mathscr{R}_s[f]] \tag{15'}$$

and from [11, p. 289] we have for $0 \leqq l \leqq n$

$$\begin{aligned} \|(f - \mathscr{P}_s \oplus \mathscr{P}_t[f])^{(k,l)}\|_\infty &\leqq \varepsilon_{nl}\|D^{(0,n+1)}D^{(k,0)}\mathscr{R}_s[f]\|_\infty h^{n+1-l} \\ &= \varepsilon_{nl}\|D^{(k,n+1)}\mathscr{R}_s[f]\|_\infty h^{n+1-l} \end{aligned} \tag{16}$$

where $\varepsilon_{nl} \equiv 1/(n-l+1)!$.

Eqs. (16) and (14) then yield (11).

3. Transfinite Elements

In the previous section, we established a class of transfinite interpolation formulae based upon the use of Lagrange polynomial blending functions. The domain of the projections discussed in Section 2 was the square $\mathscr{S}$: $[0,h]\times[0,h]$ in the st-plane. Now, we describe how to construct transfinite interpolants over arbitrary curvilinear quadrilateral regions in the plane.

Let $\mathscr{E}$ be a closed, bounded and simply connected region in the xy-plane whose boundary, $\partial\mathscr{E}$ is subdivided into four parametric curve segments. Let $\vec{T}(s,t)$ be a univalent (one-to-one) mapping of the square $\mathscr{S}$: $[0,h]\times[0,h]$ in the st-plane onto the region $\mathscr{E}$ in the xy-plane:

$$\vec{T}(s,t) = \begin{pmatrix} x(s,t) \\ y(s,t) \end{pmatrix}.$$

Since $\vec{T}$ is univalent and onto, the boundary of $\mathscr{S}$ maps onto the boundary of $\mathscr{E}$ and the mapping is invertible. In [9], the authors consider in detail the construction of such mappings for arbitrary closed, bounded and simply connected regions $\mathscr{E}$.

Let the four parametric curves $\vec{F}(0, t)$, $\vec{F}(h, t)$, $\vec{F}(s, 0)$ and $\vec{F}(s, h)$ constitute the boundary of $\mathscr{E}$. The simplest, but very useful, mapping discussed in [9] is the *vector-valued bilinearly blended map*

$$\begin{aligned}\vec{T}(s,t) = \begin{bmatrix} x(s,t) \\ y(s,t) \end{bmatrix} &= (1-s/h)\,\vec{F}(0,t) + (s/h)\,\vec{F}(h,t) \\ &\quad + (1-t/h)\,\vec{F}(s,0) + (t/h)\,\vec{F}(s,h) \\ &\quad - (1-s/h)(1-t/h)\,\vec{F}(0,0) - (1-s/h)(t/h)\,\vec{F}(0,h) \\ &\quad - (1-t/h)(s/h)\,\vec{F}(h,0) - (t/h)(s/h)\,\vec{F}(h,h).\end{aligned} \tag{17}$$

By construction $\vec{T}$ maps $\partial\mathscr{S}$ onto $\partial\mathscr{E}$. If we could also establish that the Jacobian is non-zero,

$$\det J \equiv \det \begin{vmatrix} x_s & x_t \\ y_s & y_t \end{vmatrix} \neq 0, \tag{18}$$

then we could conclude that T was at least locally invertible. However the determination of the validity of (18) for the map $\vec{T}$ in (17), or any other $\vec{T}$, is a separate problem and has been considered in [9]. The approach adopted there is heuristic: It is guided by experience, geometric intuition and analysis, and is accomplished with the aid of computer graphics, i.e. actual visual inspection.

To fix ideas, consider as a simple example of a univalent map, the region $\mathscr{E}$ in Fig. 1. The boundary of $\mathscr{E}$ is described by the four curves:

$$\vec{F}(0,t) = \begin{bmatrix} 1+(t/h)\sqrt{2} \\ 0 \end{bmatrix} \qquad 0 \leqq t \leqq h,$$

$$\vec{F}(h,t) = \begin{bmatrix} 0 \\ 1+(t/h)\sqrt{2} \end{bmatrix} \qquad 0 \leqq t \leqq h,$$

$$\vec{F}(s,0) = \begin{bmatrix} \operatorname{Cos}(s\pi/2h) \\ \operatorname{Sin}(s\pi/2h) \end{bmatrix} \qquad 0 \leqq s \leqq h,$$

$$\vec{F}(s,h) = \begin{bmatrix} (1+\sqrt{2})\operatorname{Cos}(s\pi/2h) \\ (1+\sqrt{2})\operatorname{Sin}(s\pi/2h) \end{bmatrix} \qquad 0 \leqq s \leqq h.$$

The bilinearly blended map (17) for this very simple case reduces to

$$\vec{T}(s,t) = \begin{bmatrix} \left(1+(t/h)\sqrt{2}\right)\operatorname{Cos}(s\pi/2h) \\ \left(1+(t/h)\sqrt{2}\right)\operatorname{Sin}(s\pi/2h) \end{bmatrix}.$$

It is easy to confirm that $\vec{T}$ is univalent and that the Jacobian is non-zero. In fact, it is easy to see that the *flow lines* of this map (i.e., the image in the xy-plane of the lines $s=\text{const}$ and $t=\text{const}$) are just radial lines and circular

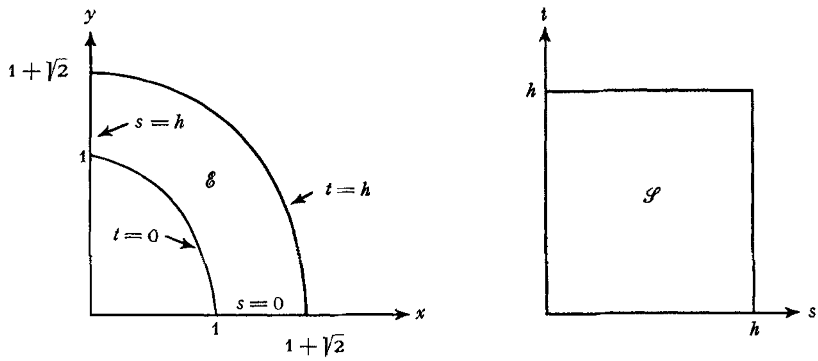


Fig. 1

arcs joining corresponding points on opposite boundaries of the region $\mathscr{E}$. In Section 6, we will see how (17) can be used to generate most of the isoparametric elements considered in [17].

Any univalent map $\vec{T}$: $\mathscr{S}\to\mathscr{E}$ provides a unique correspondence between a point $(s,t)\in\mathscr{S}$ and its image $(x,y)\in\mathscr{E}$. Hence, if a function $f(s,t)$ is defined for $(s,t)\in\mathscr{S}$, then via an *a priori* map $\vec{T}$ we obtain a function, say $f^*(x,y)$ defined for all $(x,y)\in\mathscr{E}$ by the identification

$$f^*\big(x(s,t),y(s,t)\big)\equiv f(s,t),\qquad (s,t)\in\mathscr{S}.$$

Now, let f be defined for all $(s,t)\in\mathscr{S}$; let $\mathscr{P}_s\oplus\mathscr{P}_t[f]$ be the (m,n)-degree transfinite Lagrange interpolant given by (8); and let $\vec{T}$: $\mathscr{S}\to\mathscr{E}$ be *any* univalent map of $\mathscr{S}$ onto $\mathscr{E}$.

Definition. The function $\mathscr{P}_s\oplus\mathscr{P}_t[f]^*$ defined for all $(x,y)\in\mathscr{E}$ by the identification

$$\mathscr{P}_s\oplus\mathscr{P}_t[f]^*\big(x(s,t),y(s,t)\big)\equiv\mathscr{P}_s\oplus\mathscr{P}_t[f](s,t)$$

together with the mapping $\vec{T}$: $\mathscr{S}\to\mathscr{E}$ is a *transfinite element* with domain $\mathscr{E}$.

If the primitive function f^* is defined, as above, for all $(x,y)\in\mathscr{E}$ by the relation $f^*(x,y)\equiv f(s,t)$, then $\mathscr{P}_s\oplus\mathscr{P}_t[f]^*$ interpolates f^* along the two sets of curves in $\mathscr{E}$ which are the images under $\vec{T}$ of the two families of lines $s=s_i$ and $t=t_j$ in $\mathscr{S}$, i.e.,

$$\mathscr{P}_s\oplus\mathscr{P}_t[f]^*\big(x(s_i,t),y(s_i,t)\big)=f^*\big(x(s_i,t),y(s_i,t)\big),\qquad i=0,1,\ldots,m$$

$$\mathscr{P}_s\oplus\mathscr{P}_t[f]^*\big(x(s,t_j),y(s,t_j)\big)=f^*\big(x(s,t_j),y(s,t_j)\big),\qquad j=0,1,\ldots,n.$$

As we shall subsequently see, essentially all of the conventional finite elements are obtainable as specializations of the transfinite elements just defined. In the very special case when $\vec{T}$ is the identity map $(x\equiv s, y\equiv t)$, one obtains a class of transfinite elements with the square domain $[0,h]\times[0,h]$. Or, if we take $\vec{T}$ to be the map $x=s$, $y=st$ (which maps the square $\mathscr{S}=[0,h]\times[0,h]$ onto the

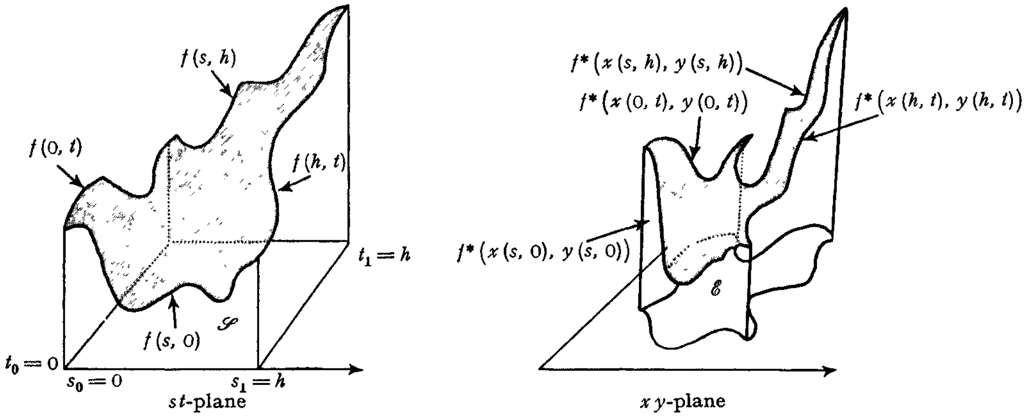


Fig. 2. Bilinear blending over $\mathscr{S}$ together the map $\mathscr{S} \to \mathscr{E}$ induces the transfinite surface element $\mathscr{P}_s \oplus \mathscr{P}_t[f]^*(x, y)$ over $\mathscr{E}$ with the properties that $\mathscr{P}_s \oplus \mathscr{P}_t[f]^*$ matches f^* on the perimeter of $\mathscr{E}$. The domain $\mathscr{E}$ is an arbitrary closed, bounded and simply connected region in the xy-plane

triangle with vertices $(x, y) = (0, 0), (h, 0), (h, h^2))$, then we obtain a class of transfinite elements defined over this triangle with the property that $\mathscr{P}_s \oplus \mathscr{P}_t[f]^*$ coincides with f^* along the set of $m+1$ parallel lines $x = x_i \equiv s_i$ and the set of $n+1$ radial lines defined by $y = s\, t_j$.

Fig. 2 graphically illustrates the construction of a simple transfinite element. Under the invertible mapping $\vec{T}$ (whose form is unspecified), we have a one-to-one correspondence $(s, t) \leftrightarrow (x, y)$. Suppose that the function f^* defined over $\mathscr{E}$ is given and that we wish to interpolate to f^* on $\partial \mathscr{E}$. First, define the function f over $\mathscr{S}$ by the identification of values: $f(s, t) \equiv f^*(x(s, t), y(s, t))$. Then, construct the bilinearly blended interpolant $\mathscr{P}_s \oplus \mathscr{P}_t[f]^*(x, y) \equiv \mathscr{P}_s \oplus \mathscr{P}_t[f](s, t)$ such that

$$\mathscr{P}_s \oplus \mathscr{P}_t[f]^*|_{\partial \mathscr{E}} = f^*|_{\partial \mathscr{E}},$$

which is the desired transfinite element with domain $\mathscr{E}$.

4. Discrete Approximations to Transfinite Interpolants

In order to develop practical interpolation schemes for use in the Ritz method and other applications, the arbitrary univariate functions $f(s_i, t)$, $0 \leqq i \leqq m$ and $f(s, t_j)$, $0 \leqq j \leqq n$ in (2) and (4) must be approximated in terms of some finite number of scalar parameters. In this section, therefore, we derive several finite parameter interpolation formulae which precipitate as approximations to the *transfinite bivariate Lagrange interpolants*.

Nota bene: The concept of a transfinite element involves two notions: (a) the transfinite interpolation formula $\mathscr{P}_s \oplus \mathscr{P}_t[f]$, and (b) the mapping $\vec{T}$: $\mathscr{S} \to \mathscr{E}$ which provides the correspondence between $f^*(x, y)$ and $f(s, t)$. In this section and the next, we consider methods for approximating the transfinite interpolant by finite parameter interpolants; i.e., discrete approximations to transfinite Lagrange interpolants. We do not, however, assume anything about the nature of the map $\vec{T}\colon \mathscr{S}\to\mathscr{E}$ since it is, in general, totally unrelated to the interpolation scheme. In particular, the map $\vec{T}$ itself must generally be regarded as being transfinite, since it is not possible to obtain a finite parameter interpolatory map of $\mathscr{S}$ onto $\mathscr{E}$ (more pointedly, $\partial\mathscr{S}$ onto $\partial\mathscr{E}$) for an arbitrary closed, bounded and simply connected domain $\mathscr{E}$ (cf. [9]). Thus, although we discretize the interpolation formula used in defining an element, this element is still to be regarded as transfinite because of the, in general, transfinite nature of the mapping $\vec{T}$ (cf. Section 6). This idea is closely related to Zienkiewicz's remarks [18, pp. 137–138] concerning sub-, super-, and isoparametric elements.

The derivation of finite parameter interpolation formulas can be viewed as a two-stage decomposition: We first project f, via $\mathscr{P}_s\oplus\mathscr{P}_t$, to obtain the transfinite interpolant $\mathscr{P}_s\oplus\mathscr{P}_t[f]$. The two components $\mathscr{P}_s[f]$ and $\mathscr{P}_t[f]$ are then approximated by finite parameter tensor product interpolants $\overline{\mathscr{P}}_t\mathscr{P}_s[f]$ and $\overline{\mathscr{P}}_s\mathscr{P}_t[f]$ obtained by projecting with two new projectors $\overline{\mathscr{P}}_t$ and $\overline{\mathscr{P}}_s$. Hence

$$\begin{aligned}\mathscr{P}_s\oplus\mathscr{P}_t &= \mathscr{P}_s+\mathscr{P}_t-\mathscr{P}_s\mathscr{P}_t\\ &\equiv\{\overline{\mathscr{P}}_t\mathscr{P}_s+\overline{\mathscr{P}}_s\mathscr{P}_t-\mathscr{P}_s\mathscr{P}_t\}+\overline{\mathscr{R}}_t\mathscr{P}_s+\overline{\mathscr{R}}_s\mathscr{P}_t\end{aligned}\tag{19}$$

where $\overline{\mathscr{R}}_s=I-\overline{\mathscr{P}}_s$ and $\overline{\mathscr{R}}_t=I-\overline{\mathscr{P}}_t$.

Consider expressions (2) and (4) for $\mathscr{P}_s[f]$ and $\mathscr{P}_t[f]$. If $\overline{\mathscr{P}}_s$ and $\overline{\mathscr{P}}_t$ are chosen as projectors of the same class as $\mathscr{P}_s$ and $\mathscr{P}_t$, respectively, then for example

$$\overline{\mathscr{P}}_t\mathscr{P}_s[f]=\sum_{i=0}^{m}\overline{\mathscr{P}}_t[f(s_i,t)]\,\varphi_i(s)=\sum_{i=0}^{m}\sum_{j=0}^{\bar{n}}f(s_i,\bar{t}_j)\,\varphi_i(s)\,\overline{\psi}_j(t)$$

is an approximation to $\mathscr{P}_s[f]$.

More specifically, suppose that $\overline{\mathscr{P}}_s$ is determined by Lagrange interpolation at the points $0=\bar{s}_0<\bar{s}_1<\cdots<\bar{s}_{\overline{m}}=h$ and $\overline{\mathscr{P}}_t$ by Lagrange interpolation at the points $0=\bar{t}_0<\bar{t}_1<\cdots<\bar{t}_{\bar{n}}=h$. Thus $\overline{\mathscr{P}}_s[f]$ can be obtained from (2) and (3) by putting bars over $\mathscr{P}$, m, s_i, and φ_i. However, one is not free to choose the $\bar{s}_i$ and $\bar{t}_j$ arbitrarily if the projection $(\overline{\mathscr{P}}_t\mathscr{P}_s+\overline{\mathscr{P}}_s\mathscr{P}_t-\mathscr{P}_s\mathscr{P}_t)[f]$ is to interpolate to f at a point set in $\mathscr{S}$. To obtain an interpolatory formula, it is natural to require the containment relations $\{s_i\}\subseteq\{\bar{s}_i\}$ and $\{t_j\}\subseteq\{\bar{t}_j\}$. Even if this is not the case, however, one still obtains an approximation (but not, in general, an interpolation) to f. In the former case, we deduce

Theorem 3. *Discrete Approximation of Transfinite Bivariate Lagrange Interpolation.*

Let $\mathscr{P}_s$, $\overline{\mathscr{P}}_s$, $\mathscr{P}_t$ and $\overline{\mathscr{P}}_t$ be defined as above with $\{s_i\}\subseteq\{\bar{s}_i\}$ and $\{t_j\}\subseteq\{\bar{t}_j\}$. Then,

$$\tilde{f}(m,n,\overline{m},\bar{n})\equiv(\overline{\mathscr{P}}_t\mathscr{P}_s+\overline{\mathscr{P}}_s\mathscr{P}_t-\mathscr{P}_s\mathscr{P}_t)\,[f]\tag{20}$$

interpolates to f at the $(\overline{m}+1)(\bar{n}+1)-(\overline{m}-m)(\bar{n}-n)$ points

$$\begin{aligned}&(s_i,\bar{t}_j), && 0\leqq i\leqq m, && 0\leqq j\leqq\bar{n}\\ &(\bar{s}_i,t_j), && i\in\{k\,|\,\bar{s}_k\notin\{s_i\}\}, && 0\leqq j\leqq n.\end{aligned}\tag{20'}$$

If $m=\overline{m}$ and $n=\overline{n}$, then $\tilde{f}(m, n, \overline{m}, \overline{n})$ is the tensor product projection $\mathscr{P}_s \mathscr{P}_t[f]$. If $m=\overline{m}=n=\overline{n}$, then $\tilde{f}(m, m, m, m)$ in the *bivariate Lagrange polynomial interpolant* of f which interpolates f at the $(m+1)^2$ points (s_i, t_j), $0 \leqq i, j \leqq m$. In general, however, the point stencil described by Eqs. (20′) is *not* a Cartesian product stencil.

From the remainder formula

$$\begin{aligned} f-\tilde{f} &= \left(I-(\overline{\mathscr{P}}_t \mathscr{P}_s+\overline{\mathscr{P}}_s \mathscr{P}_t-\mathscr{P}_s \mathscr{P}_t)\right)[f] \\ &= (I-\mathscr{P}_s \oplus \mathscr{P}_t)[f]+\overline{\mathscr{R}}_t \mathscr{P}_s[f]+\overline{\mathscr{R}}_s \mathscr{P}_t[f] \\ &= \mathscr{R}_s \mathscr{R}_t[f]+\overline{\mathscr{R}}_t[f]-\overline{\mathscr{R}}_t \mathscr{R}_s[f]+\overline{\mathscr{R}}_s[f]-\overline{\mathscr{R}}_s \mathscr{R}_t[f] \\ &= \overline{\mathscr{R}}_s[f]+\overline{\mathscr{R}}_t[f]+\mathscr{R}_s \mathscr{R}_t[f]-\overline{\mathscr{R}}_s \mathscr{R}_t[f]-\mathscr{R}_s \overline{\mathscr{R}}_t[f], \end{aligned} \tag{21}$$

which uses (19) and (10), we can easily deduce the following

Theorem 4. *Let* $f \in C^{(p,q)}(\mathscr{S})$, $p=\max(m+1, \overline{m}+1)$, $q=\max(n+1, \overline{n}+1)$ *and* $\tilde{f}=\tilde{f}(m, n, \overline{m}, \overline{n})$ *as in* (20). *Then*

$$\begin{aligned} \|(f-\tilde{f})^{(k,l)}\|_\infty \leqq\, & \varepsilon_{\overline{m}k}\|f^{(\overline{m}+1,l)}\|_\infty h^{\overline{m}+1-k} \\ & +\varepsilon_{\overline{n}l}\|f^{(k,\overline{n}+1)}\|_\infty h^{\overline{n}+1-l} \\ & +\varepsilon_{mk}\,\varepsilon_{nl}\|f^{(m+1,n+1)}\|_\infty h^{m+n+2-k-l} \\ & +\varepsilon_{\overline{m}k}\,\varepsilon_{nl}\|f^{(\overline{m}+1,n+1)}\|_\infty h^{\overline{m}+n+2-k-l} \\ & +\varepsilon_{mk}\,\varepsilon_{\overline{n}l}\|f^{(m+1,\overline{n}+1)}\|_\infty h^{m+\overline{n}+2-k-l} \end{aligned} \tag{22}$$

$0 \leqq k \leqq \min(m, \overline{m})$, $0 \leqq l \leqq \min(n, \overline{n})$.

Remarks. 1. Fix m and n. Then for all $\overline{m}, \overline{n} \geqq m+n+1$

$$\|(f-\tilde{f})^{(k,l)}\|_\infty = \mathcal{O}(h^{m+n+2-k-l}) \qquad \text{as } h \to 0. \tag{23}$$

Hence, there is a limit to the improvement (in the asymptotic order) that one can realize at the second stage of decomposition by increasing $\overline{m}$ and $\overline{n}$, and this limit is determined by the accuracy of the original transfinite approximation to f.

2. Fix m and n. Then the "optimal" order in (23) is obtained from (22) by taking $\overline{m}=\overline{n}=m+n+1$ and no smaller integer.

3. Let $\overline{m}=m$ and $\overline{n}=n$. Then, $\tilde{f}=\mathscr{P}_t \mathscr{P}_s[f]$ is the product projection as in (6) and we have from (21) that

$$\begin{aligned} \|(f-\mathscr{P}_t \mathscr{P}_s[f])^{(k,l)}\|_\infty \leqq\, & \varepsilon_{mk}\|f^{(m+1,l)}\|_\infty h^{m+1-k} \\ & +\varepsilon_{nl}\|f^{(k,n+1)}\|_\infty h^{n+1-l} \\ & +\varepsilon_{mk}\,\varepsilon_{nl}\|f^{(m+1,n+1)}\|_\infty h^{m+n+2-k-l} \end{aligned} \tag{24}$$

$0 \leqq k \leqq m$, $0 \leqq l \leqq n$. Comparing (24) with (16) we observe the increased accuracy of the Boolean sum projection $\mathscr{P}_s \oplus \mathscr{P}_t[f]$ over the tensor product projection $\mathscr{P}_t \mathscr{P}_s[f]$. Indeed, since the precision set of the operator $\mathscr{P}_s \oplus \mathscr{P}_t$ is a transfinite point set in $\mathscr{S}$, it is expected to be more accurate than any finite parameter approximation in the class of schemes given by Theorem 3.

Table 1

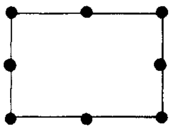
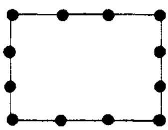
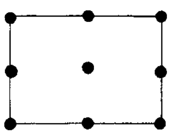
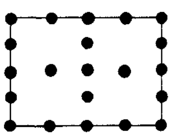
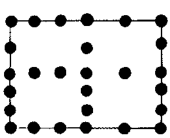
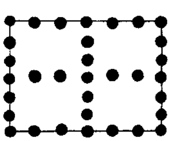
$p_1 \ni \|f - \mathscr{P}_s \oplus \mathscr{P}_t[f]\|_\infty = \mathcal{O}(h^{p_1})$ as $h \to 0$

$p_2 \ni \|f - \mathscr{P}_s \mathscr{P}_t[f]\|_\infty = \mathcal{O}(h^{p_2})$ as $h \to 0$

$p_3 \ni \|f - \tilde{f}(m, n, \overline{m}, \overline{n})\|_\infty = \mathcal{O}(h^{p_3})$ as $h \to 0$

N = number of points in stencil

$\tilde{f}(m, n, \overline{m}, \overline{n})$ = Function defined by Eq. (20)

$\tilde{f}(1, 1, 1, 1)$ 	$p_1 = 4$,	$p_2 = 2$,	$p_3 = 2$,	$N = 4$
$\tilde{f}(1, 1, 2, 2)$ 	$p_1 = 4$,	$p_2 = 2$,	$p_3 = 3$,	$N = 8$
$\tilde{f}(1, 1, 3, 3)$ 	$p_1 = 4$,	$p_2 = 2$,	$p_3 = 4$,	$N = 12$
$\tilde{f}(2, 2, 2, 2)$ 	$p_1 = 6$,	$p_2 = 3$,	$p_3 = 3$,	$N = 9$
$\tilde{f}(2, 2, 4, 4)$ 	$p_1 = 6$,	$p_2 = 3$,	$p_3 = 5$,	$N = 21$
$\tilde{f}(2, 2, 5, 5)$ 	$p_1 = 6$,	$p_2 = 3$,	$p_3 = 6$,	$N = 27$
$\tilde{f}(2, 2, 6, 6)$ 	$p_1 = 6$,	$p_2 = 3$,	$p_3 = 6$,	$N = 33$

4. See Table 1. This table shows seven simple stencils which are obtained from the two-stage decomposition scheme described by Theorem 3. The first stencil corresponds to ordinary bilinear interpolation, $m=n=1$ in (6), or $m=n=\overline{m}=\overline{n}=1$ in (20). The second stencil results from first using the bilinearly blended transfinite formula

$$\begin{aligned}\mathscr{P}_s\oplus\mathscr{P}_t[f] = {} & (1-s/h)\,f(0,t)+(s/h)\,f(1,t)\\ & +(1-t/h)\,f(s,0)+(t/h)\,f(s,1)\\ & -(1-s/h)(1-t/h)\,f(0,0)-(1-s/h)(t/h)\,f(0,h)\\ & -(s/h)(1-t/h)\,f(h,0)-(s/h)(t/h)\,f(h,h)\end{aligned}$$

obtained by taking $m=n=1$ in (8), and then approximating each of the four univariate functions (i.e., $f(0,t)$, $f(h,t)$, $f(s,0)$ and $f(s,h)$) in this expression by parabolas of interpolation at the argument values 0, $h/2$ and h. The third stencil results from approximating the univariate functions in the bilinearly blended interpolant by interpolating cubic polynomials. The last four stencils arise from first using transfinite biquadratic interpolation ($m=n=2$ in Theorem 1) followed by quadratic, quartic, quintic and sixth degree polynomial interpolations, respectively, to the six univariate functions (i.e., $f(0,t)$, $f(h/2,t)$, $f(h,t)$, $f(s,0)$, $f(s,h/2)$ and $f(s,h)$) contained in the transfinite expression $\mathscr{P}_s\oplus\mathscr{P}_t[f]$. Note that some of these stencils belong to the so-called "serendipity class" derived in [17] by different methods. (See also Section 6 below.)

5. Discretization Error for Boundary Value Problems

We now describe the application of transfinite elements to the approximate solution of boundary value problems via the Ritz method. The resulting numerical techniques will be termed *transfinite element methods*. These new methods include apparently all of the standard finite element and isoparametric finite element methods found in the literature for quadrilateral and curvilinear quadrilateral element domains $\mathscr{E}_j$.[1] However, the transfinite element methods are substantially more general since, for example, even though the local *interpolant* is a finite parameter function, the domain $\mathscr{E}_j$ over which it is defined is generally to be considered arbitrary. The relation of (parametric) transfinite elements to isoparametric finite elements is explored in the following section.

Consider the second-order linear partial differential equation

$$L[v]\equiv-\begin{bmatrix}\frac{\partial}{\partial x} & \frac{\partial}{\partial y}\end{bmatrix}\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}\begin{bmatrix}\partial v/\partial x\\ \partial v/\partial y\end{bmatrix}+cv=f \tag{25}$$

for $(x,y)\in\mathscr{R}$, a simply connected, bounded domain in the xy-plane. Here a_{ij}, c and f are continuous functions of (x,y) in $\overline{\mathscr{R}}$, the closure of $\mathscr{R}$; and

$$\begin{gathered}a_{ij}(x,y)=a_{ji}(x,y),\qquad \sum_{i,j=1}^{2}a_{ij}\zeta_i\zeta_j\geqq\mu\sum_{i=1}^{2}\zeta_i^2,\quad \zeta_i \text{ real}\\ \mu\equiv\text{constant}>0,\qquad c(x,y)\geqq 0.\end{gathered} \tag{25'}$$

1 Analogous transfinite methods for triangles can be developed on the basis of the interpolation schemes in [1].

For simplicity, we assume

$$v \equiv 0 \quad \text{on } \partial\mathscr{R}. \tag{26}$$

The true solution v to (25)–(26) minimizes the functional

$$F_1[w] \equiv \int_{\mathscr{R}} \{(w_x, w_y)\, A\, (w_x, w_y)^t + c\, w^2 - 2 f w\}\, dx\, dy \tag{27}$$

over the Sobolev space $\overset{\circ}{W}{}_2^1(\mathscr{R})$, [12, 21], where the 2 by 2 matrix A is given in (25) and superscript t denotes transpose.

Suppose $\mathscr{R}$ has been partitioned into "curvilinear quadrilateral elements" $\mathscr{E}_j$ $(1 \leqq j \leqq N)$. Assume the partitioning is chosen such that for each $\mathscr{E}_j$ there exists a univalent map $\vec{T}_j$ from $\mathscr{S}_j$: $[0, h_j] \times [0, h_j]$ onto $\mathscr{E}_j$,[2]

$$\vec{T}_j : \mathscr{S}_j \to \mathscr{E}_j, \quad 1 \leqq j \leqq N$$

where

$$h_j \equiv \max\{[(x_2 - x_1)^2 + (y_2 - y_1)^2]^{\frac{1}{2}} \mid (x_k, y_k) \in \mathscr{E}_j,\ k = 1, 2\}$$

is the diameter of $\mathscr{E}_j$. Of course other choices of h_j are possible, as long as $h_j \to 0$ if and only if the diameter of $\mathscr{E}_j \to 0$. Further, assume that for each j,

$$[\vec{T}_j]_i \in C^{(p,q)}(\mathscr{S}_j), \quad i = 1, 2, \tag{28}$$

where p and q are determined in Theorem 5 below. If, for example, T_j is the vector-valued bilinearly blended map given in (17), then (28) simply requires that the boundary segments $\vec{F}(0, t)$, $\vec{F}(h, t)$, $\vec{F}(s, 0)$ and $\vec{F}(s, h)$ be "smooth".

If J_j is the Jacobian matrix associated with $\vec{T}_j$, assume that

$$\begin{aligned} &J_j^{-1} \text{ exists for each } j, \text{ and that} \\ &\gamma_h \equiv \max_j \max_{\mathscr{E}_j} \|J_j^{-1}\|_2 \leqq \gamma < \infty \quad \text{as } h = \max_j h_j \to 0. \end{aligned} \tag{29}$$

Using the discretized interpolants developed in Theorem 3, we define a finite dimensional space of functions with domain $\mathscr{R}$ by

$$\begin{aligned} S(m, n, \overline{m}, \overline{n}) = \{f^* \mid\ & f^* \text{ satisfies (26) and for each } \mathscr{E}_j \text{ there exists} \\ & \text{an } \tilde{f}(m, n, \overline{m}, \overline{n}) \text{ such that } f^*\big(x(s,t), y(s,t)\big) = \tilde{f}(m, n, \overline{m}, \overline{n})\,(s, t)\}. \end{aligned} \tag{30}$$

Note the $S(m, n, \overline{m}, \overline{n})$ is a finite dimensional function space, the *Ritz space*, [12], associated with the N elements determined by the choice of $m, n, \overline{m}, \overline{n}$.

---

2 For example, in applications of isoparametric methods [18], the maps $\vec{T}_j$ are often constructed element by element. However, a more practical means of obtaining the $\vec{T}_j$ is described in § 6 and in the papers [9] and [19]. In these papers, Zienkiewicz and Phillips [19] apply isoparametric element methods to large sections (zones) of a given problem domain in order to introduce local curvilinear coordinates; whereas in [9], Gordon and Hall apply transfinite mapping techniques for the same purpose, i.e., to introduce a curvilinear coordinate system onto the problem domain $\mathscr{R}$. Such curvilinear coordinate systems obviously can be used to decompose $\mathscr{R}$ into curvilinear quadrilateral element domains $\mathscr{E}_j$.

The transfinite element Ritz approximation to (25)–(26) is then defined as the function $V(m, n, \bar{m}, \bar{n})$ in $S(m, n, \bar{m}, \bar{n})$ such that

$$F_1[V] = \min_{w \in S} F_1[w] \tag{31}$$

where $V \equiv V(m, n, \bar{m}, \bar{n})$ and $S = S(m, n, \bar{m}, \bar{n})$, for fixed $m, \bar{m}, n, \bar{n}$.

Recalling that the Sobolev (energy) norm is

$$\|w\|_{W_2^1(\mathscr{R})} \equiv \left[\sum_{0 \leq i+j \leq 1} \int_{\mathscr{R}} [w^{(i,j)}]^2 \, dx \, dy\right]^{\frac{1}{2}},$$

we seek asymptotic (as $h \to 0$) bounds on $\|v - V\|_{W_2^1(\mathscr{R})}$. The main result of this section is

Theorem 5. *Let v be the true solution to* (25)–(26); i.e., *the solution determined by*

$$F_1[v] = \min_{w \in \overset{\circ}{W}{}_2^1(\mathscr{R})} F_1[w].$$

Let V be the Ritz approximation to v from $S(m, n, \bar{m}, \bar{n})$ and determined by (31). *If $v \in C^{(p,q)}(\mathscr{R})$ where $p = \max(m+1, \bar{m}+1)$ and $q = \max(n+1, \bar{n}+1)$, then*

$$\|v - V\|_{W_2^1(\mathscr{R})} = \mathcal{O}(\gamma_h h^d) \quad \text{as } h \to 0 \tag{32}$$

where $d = \min\{\bar{m}, \bar{n}, m + n + 1\}$.

Proof. As in [20, p. 396] define

$$D(\varphi, \psi) \equiv \int_{\mathscr{R}} \{(\varphi_x, \varphi_y) A (\psi_x, \psi_y)^t + c\,\varphi\,\psi\} \, dx \, dy$$

for φ and ψ in $\overset{\circ}{W}{}_2^1(\mathscr{R})$. From [20, p. 396], with $D(z) \equiv D(z, z)$, there exists a constant C such that

$$\|v - V\|^2_{W_2^1(\mathscr{R})} \leq C\,D(v - V) \leq C\,D(v - \tilde{V}) \tag{33}$$

where $\tilde{V}$ is any function in S. In particular, we choose $\tilde{V}(m, n, \bar{m}, \bar{n})$ to be the element in $S(m, n, \bar{m}, \bar{n})$ determined by the function v; i.e., $\tilde{V}$ interpolates v at the specified mesh points. Note that by the chain rule and (28), if $v \in C^{(p,q)}(\mathscr{E}_j)$, as a function of x and y, then $v \in C^{(p,q)}(\mathscr{S}_j)$, as a function of s and t.

Let $\varepsilon = v - \tilde{V}$. Then

$$\begin{aligned} D(\varepsilon) &= \int_{\mathscr{R}} \{(\varepsilon_x, \varepsilon_y) A (\varepsilon_x, \varepsilon_y)^t + c\,\varepsilon^2\} (x, y) \, dx \, dy \\ &= \sum_j \int_{\mathscr{S}_j} \{(\varepsilon_s, \varepsilon_t) [J_j^{-1} A (J_j^{-1})^t] (\varepsilon_s, \varepsilon_t)^t + c\,\varepsilon^2\} (s, t) \, |\det J_j| \, ds \, dt, \end{aligned} \tag{34}$$

where ε_s means the partial derivative $\partial\varepsilon/\partial s$ regarded as a function of s and t.

But, A is symmetric and positive definite by (25′), hence it has a symmetric positive definite square root [4, p. 169]. Let $Y = (\varepsilon_s, \varepsilon_t)$, then

$$\begin{aligned} Y(J_j^{-1}) A (J_j^{-1})^t Y^t &= Y(J_j^{-1} A^{\frac{1}{2}}) (J_j^{-1} A^{\frac{1}{2}})^t Y^t \\ &\leq \|J_j^{-1} A^{\frac{1}{2}}\|_2^2 \, Y Y^t \\ &\leq \|A\|_2 \cdot \|J_j^{-1}\|_2^2 \, Y Y^t. \end{aligned}$$

Substituting into (34)

$$D(\varepsilon) \leqq \sum_j \int_{\mathscr{S}_j} \{\|A\|_2 \|J_j^{-1}\|_2^2 \{\varepsilon_s^2 + \varepsilon_t^2\} + c\,\varepsilon^2\} |\det J_j|\, ds\, dt$$

$$\leqq \left[\max_{\mathscr{R}} \|A\|_2 \max_j \left\{\max_{\mathscr{E}_j} \|J_j^{-1}\|_2^2\right\} \|\varepsilon_s^2 + \varepsilon_t^2\|_\infty + \|c\|_\infty \|\varepsilon^2\|_\infty\right] \cdot \sum_j \int_{\mathscr{S}_j} |\det J_j|\, ds\, dt. \tag{35}$$

But $\sum_j \int_{\mathscr{S}_j} |\det J_j|\, ds\, dt = \sum_j \int_{\mathscr{E}_j} dx\, dy = |\mathscr{R}| \equiv$ area of $\mathscr{R}$.

From (33) and (35), we have

$$\|v - V\|_{W_2^1(\mathscr{R})} \leqq M(\gamma_h) \{\|\varepsilon_s^2\|_\infty + \|\varepsilon_t^2\|_\infty + \|\varepsilon^2\|_\infty\}^{\frac{1}{2}} \tag{36}$$

where $M(\gamma_h) = \{C |\mathscr{R}| \max[\|c\|_\infty, \max_{\mathscr{R}} \|A\|_2 \cdot \gamma_h^2]\}^{\frac{1}{2}}$.

The orders of convergence in (32) then follow from (36) and Theorem 4. Q.E.D.

Let us next consider the *plane strain elasticity problem* [3, 10, 14, 16, 21]: Find $u, v \in \overset{\circ}{W}{}_2^1(\mathscr{R})$ such that

$$F_2[u, v] = \min_{w_i \in \overset{\circ}{W}{}_2^1(\mathscr{R})} F_2[w_1, w_2] \tag{37}$$

where the potential energy is

$$F_2[w_1, w_2] = \tfrac{1}{2} \int_{\mathscr{R}} (\bar{\varepsilon} - \alpha \bar{T})^t \bar{\sigma}\, dx\, dy - \int_{\mathscr{R}} \bar{\delta}^t \bar{p}\, dx\, dy, \tag{38}$$

$\bar{\delta} = (w_1, w_2)^t$ is the vector of displacements, $\bar{T} = [T, T, 0, T]^t$,

$$\bar{\varepsilon}(w_1, w_2) \equiv \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{xy} \\ \varepsilon_z \end{bmatrix} = \begin{bmatrix} w_1^{(1,0)} \\ w_2^{(0,1)} \\ w_1^{(0,1)} + w_2^{(1,0)} \\ 0 \end{bmatrix}$$

is the strain vector,

$$\bar{\sigma}(w_1, w_2) \equiv \begin{bmatrix} \sigma_x \\ \sigma_y \\ \sigma_{xy} \\ \sigma_z \end{bmatrix} = \frac{2G}{(1-2\nu)} \begin{bmatrix} (1-\nu) & \nu & 0 & 0 \\ \nu & (1-\nu) & 0 & 0 \\ 0 & 0 & \frac{(1-2\nu)}{2} & 0 \\ \nu & \nu & 0 & 0 \end{bmatrix} \bar{\varepsilon} - \frac{2G(1+\nu)}{(1-2\nu)} \alpha \bar{T}$$

is the stress vector, ν is Poisson's ratio, G is the shear modulus, $T(x, y)$ is the temperature distribution, α is the coefficient of thermal expansion and $\bar{p} = (X, Y)^t$ is the vector of body forces. For simplicity, let

$$u = v = 0 \quad \text{on } \partial\mathscr{R}.$$

We now establish the following convergence results

Theorem 6. *Let u and v satisfy* (37) *and U, V in the space $S(m, n, \bar{m}, \bar{n})$ satisfy*

$$F_2(U, V) = \min_{w_i \in S} F_2(w_1, w_2).$$

If u and v belong to $C^{(p,q)}(\mathscr{R})$ where $p=\max(m+1, \bar{m}+1)$ and $q=\max(n+1, \bar{n}+1)$, then

$$\max\{\|(U-u)_x\|_{L_2}, \|(V-v)_y\|_{L_2}, \|(U-u)_y+(V-v)_x\|_{L_2}\}=\mathcal{O}(\gamma_h h^d) \quad \text{as } h\to 0 \tag{39}$$

where $d=\min\{\bar{m}, \bar{n}, m+n+1\}$.

Remark. The quantities bounded in (39) are the errors in the components of strain. Obviously, by Hooke's law [16] the components of stress satisfy the same relation. Also, as in [3, Corollary to Theorem 1], the errors in displacement components

$$\|U-u\|_{L_2} \quad \text{and} \quad \|V-v\|_{L_2}$$

also are $\mathcal{O}(\gamma_h h^d)$ as $h\to 0$.

Proof of Theorem 6. From [3, Eq. (16)], let

$$D(\varphi,\psi)\equiv\frac{1}{2}\int_{\mathscr{R}}\bar{\varepsilon}^t\bar{\sigma}\,dx\,dy=\int_{\mathscr{R}}\bar{\varepsilon}^t H\,\varepsilon\,dx\,dy$$

where

$$H=\frac{G}{(1-2\nu)}\begin{bmatrix}(1-\nu) & \nu & 0 & 0\\ \nu & (1-\nu) & 0 & 0\\ 0 & 0 & \frac{(1-2\nu)}{2} & 0\\ \nu & \nu & 0 & 0\end{bmatrix}$$

is a positive semi-definite matrix.

The squares of the quantities bounded in (39) are bounded by a constant times $D(\tilde{U}-u, \tilde{V}-v)$ where $\tilde{U}(\tilde{V})$ is the element in $S(m,n,\bar{m},\bar{n})$ determined by $u(v)$. This follows from [3, Eq. (8)].

The proof now parallels that of Theorem 5 and we only sketch its completion. Write $\xi=\tilde{U}-u$ and $\eta=\tilde{V}-v$, then

$$\bar{\varepsilon}(\xi,\eta)H\bar{\varepsilon}(\xi,\eta)^t=(\xi_x,\xi_y,\eta_x,\eta_y)A(\xi_x,\xi_y,\eta_x,\eta_y)^t$$

where

$$A=\begin{bmatrix}(1-\nu) & 0 & 0 & \nu\\ 0 & \frac{(1-2\nu)}{2} & \frac{(1-2\nu)}{2} & 0\\ 0 & \frac{(1-2\nu)}{2} & \frac{(1-2\nu)}{2} & 0\\ \nu & 0 & 0 & (1-\nu)\end{bmatrix}$$

is positive semi-definite (with a positive semi-definite square root).

Let $B=\begin{bmatrix}J^{-1} & 0\\ 0 & J^{-1}\end{bmatrix}^t$, then

$$\begin{aligned}D(\tilde{U}-u,\tilde{V}-v)&=\int_{\mathscr{R}}\bar{\varepsilon}(\xi,\eta)H\bar{\varepsilon}(\xi,\eta)^t dx\,dy\\ &=\sum_j\int_{\mathscr{S}_j}(\xi_s,\xi_t,\eta_s,\eta_t)BAB^t(\xi_s,\xi_t,\eta_s,\eta_t)^t\cdot|\det J_j|\,ds\,dt\\ &\leqq|\mathscr{R}|\max_{\mathscr{R}}\|A\|_2\max_j\|J_j^{-1}\|_2^2\{\|\xi_s\|_\infty^2+\|\xi_t^2\|_\infty+\|\eta_s^2\|_\infty+\|\eta_t^2\|_\infty\}.\end{aligned}$$

The orders of convergence in (39) then follow from Theorem 4. Q.E.D.

We wish to emphasize that Theorems 5 and 6 are *independent* of the nature of the maps $\bar{T}_j$: $\mathcal{S}_j \to \mathcal{E}_j$. We have assumed nothing about these maps save that they are invertible. In particular, note that the definition (30) of the Ritz space $S(m, n, \bar{m}, \bar{n})$ is independent of the form of the mappings. In (32), the asymptotic behavior of the discretization error depends on γ_h, which reflects the choice of the mappings $\{\vec{T}_j\}$, and d which reflects the choice of Ritz space S. Note that if each $\mathcal{E}_j$ is a square, then T_j is merely the identity map $x \equiv s$, $y \equiv t$ and we have $T_j = \frac{1}{\sqrt{2}} I$ for all j and $\gamma_h = \sqrt{2}$. Finally, we reiterate that Theorems 5 and 6 establish orders of convergence for isoparametric elements [18] when the mappings T_j are chosen appropriately.

The authors would like to thank one of the referees for bringing to our attention the paper [23]. In [23], as well as [24], Ciarlet and Raviart have independently developed discretization error estimates for various Lagrange and Hermite type isoparametric finite elements. In essence, their results indicate that conditions (28) and (29) are reasonable assumptions. For example, in [24, p. 433] sufficient conditions are given for curved finite elements of *type* $(2, Q)$ to be *2-regular*, which (cf. [24, p. 426]) implies, in turn, that the associated map $\vec{T}_j$ satisfies (28) and (29) above.

For the convenience of the interested reader, we now indicate the connection between our generically designated map $\vec{T}_j$ and Ciarlet and Raviart's map $\vec{F}_h$. The "curved element" $\mathcal{E}_j$ is designated K in [23, 24] and their $\hat{K}$ is normally chosen to be the unit square [23, p. 245, Fig. 6]. The map $\vec{F}_h$ is constructed so that

$$\vec{F}_h: \hat{K} \mapsto \mathcal{E}_j.$$

Defining $\vec{G}_h$ to be the affine (expansive) map

$$\vec{G}_h: \mathcal{S}_j: [0, h_j] \times [0, h_j] \mapsto \hat{K}: [0, 1] \times [0, 1]$$

we note that

$$\vec{T}_j = \vec{F}_h \vec{G}_h.$$

Also, if the curved elements under consideration are k-regular, then $\vec{F}_h$ is a C^{k+1}-diffeomorphism and hence the Jacobian matrix

$$J_j \equiv D\vec{T}_j = D\vec{F}_h \cdot D\vec{G}_h$$

where, as in [23, 24], $D\vec{T}_j$ is the Fréchet derivative. By construction

$$J_j^{-1} = \begin{bmatrix} h & 0 \\ 0 & h \end{bmatrix} \cdot D\vec{F}_h^{-1}$$

from which we conclude that $\|J_j^{-1}\|_2 \leq \sqrt{2}\, h \|D\vec{F}_h^{-1}\|_\infty \leq 2\sqrt{2} c_{-1}$, where c_{-1} is the constant in [24, p. 427, Eq. (2.16)]. Finally, we note that the other assumptions in [24, Eqs. (2.17)–(2.18)] are necessary in their development to guarantee that if a scalar function $v \in C^{(p,q)}(\mathcal{E}_j)$, then $v \in C^{(p,q)}(\hat{K})$, and such conditions are implicit in requiring that our map $\vec{T}_j$ satisfy (28).

6. Isoparametric Elements

The notion of an isoparametric finite element introduced by Ergatoudis, Irons and Zienkiewicz in [6] is based upon the use of a single finite parameter interpolation formula for both the mapping $\vec{T}$: $\mathcal{S} \to \mathcal{E}$ and the definition of the dependent variable defined over $\mathcal{E}$. Thus far, in this paper we have persisted in keeping these two aspects separate. That is, we have assumed that an invertible map of $\mathcal{S}$ onto $\mathcal{E}$ is given and have concentrated attention on the problem of interpolating to function values $f^*(x, y)$ given on certain point sets in $\mathcal{E}$. We now briefly describe a specialization of (parametric) transfinite elements which *couples* the mapping problem and the interpolation problem. These special formulas will be termed *isoparametric transfinite* elements, since they are the exact transfinite *analogs* of the familiar isoparametric *finite* elements [18].

Since isoparametric (finite) elements are discussed in detail in [17, 18] we shall herin restrict the exposition to a very simple, but illustrative transfinite example based upon bilinearly blended interpolation. Suppose that the boundary of the region $\mathcal{E}$ onto which we seek to map the square $\mathcal{S} = [0, h] \times [0, h]$ is described by the four vector-valued functions $\vec{F}(0, t)$, $\vec{F}(h, t)$, $\vec{F}(s, 0)$ and $\vec{F}(s, h)$. Whatever the map $\vec{T}$: $\mathcal{S} \to \mathcal{E}$ may be, it is obvious that the four boundary segments of $\mathcal{E}$ must be the images of the four sides of the square $[0, h] \times [0, h]$. The simplest map which satisfies these constraints is the bilinearly blended transfinite map of expression (17). The reader may easily verify that $\vec{T}(s, t)$ does, in fact, reduce to the four specified curves when the variables s and t range along the perimeter of $\mathcal{S}$.

As we have repeatedly emphasized throughout this paper, the interpolation method used to approximate the function f^* over $\mathcal{E}$ has, in general, no relation to the mapping formula $\vec{T}(s, t)$. But, by definition, in the special class of *isoparametric* methods, precisely the *same* formula is used both to execute the map and to construct the interpolating approximation to the dependent variable $z = f^*(x, y) \equiv f(s, t)$. Thus, the simplest isoparametric transfinite element for curvilinear quadrilateral elements is given by the mapping formula (17) together with the bilinearly blended transfinite interpolation formula

$$\begin{aligned} z(s, t) = {} & (1 - s/h)\, f(0, t) + s/h\, f(h, t) \\ & + (1 - t/h)\, f(s, 0) + t/h\, f(s, h) \\ & - (1 - s/h)(1 - t/h)\, f(0, 0) - (1 - s/h)(t/h)\, f(0, h) \\ & - (s/h)(1 - t/h)\, f(h, 0) - (s/h)(t/h)\, f(h, h) \end{aligned}$$

which serves to define the dependent variable z over the region $\mathcal{E}$. If the map of $\mathcal{S} \to \mathcal{E}$ in Fig. 2 is imagined to be of the form (17), then the element depicted there can be viewed as an example of a bilinearly blended isoparametric transfinite element.

More succinctly, if we let $X_1 = x$, $X_2 = y$ and $X_3 = z$, then the general class of isoparametric transfinite Lagrange elements can be described by the generic formula

$$X_i(s, t) = (\mathcal{P}_s \oplus \mathcal{P}_t)[f_i](s, t) \qquad (i = 1, 2, 3),$$

where the f_i represent the given functional information for the i-th variable and the projectors $\mathcal{P}_s$ and $\mathcal{P}_t$ are Lagrange-type projectors of the form described in Section 2 (see Theorem 1).

Having specialized to isoparametric transfinite schemes and now recalling the discretization techniques of Section 4, it is easy to see how to obtain the finite parameter isoparametric methods of Ergatoudis, Irons and Zienkiewicz. Specifically, in the notation of this paper all of the (quadrilateral) isoparametric *finite* element schemes described in [17, 18] can be represented as

$$X_i(s, t) = (\bar{\mathcal{P}}_t \mathcal{P}_s + \bar{\mathcal{P}}_s \mathcal{P}_t - \mathcal{P}_s \mathcal{P}_t)[f_i](s, t) \qquad (i = 1, 2, 3).$$

In particular, recall the stencils in Table 1 of Section 4.

Nota bene. Although isoparametric elements are capable of handling very complex geometries (i.e., oddly shaped regions $\mathcal{R}$ and, after decomposition, subdomains $\mathcal{E}_j$ such that $\bigcup_{j=1}^{N} \mathcal{E}_j = \mathcal{R}$) and are therefore of great practical value, they have one very serious defect: Namely, the mapping $\vec{T}$: $\mathcal{S} \to \mathcal{E}$ defined (implicitly) by the isoparametric formulas may not be one-to-one. That is, there may be two (or more) points (s_1, t_1) and (s_2, t_2) in $\mathcal{S} = [0, h] \times [0, h]$ which map onto the same point (x^*, y^*) in $\mathcal{E}$. Consequently, the dependent variable z will be double valued at the point (x^*, y^*)—a physically meaningless situation for variables such as stress, strain, temperature, etc. Zienkiewicz *et al.* have pointed out this deficiency for finite parameter isoparametric elements in [18], and the same deficiency is also inherent in the isoparametric transfinite schemes above. Indeed, it is precisely this difficulty which motivated us to decouple the two aspects of parametric element definition; i.e., the mapping of the independent variables and the interpolation of the dependent variable. This paper is an adjunct to another paper [9] in which we examine mapping problems per se in much greater detail.

However, [9] is quite different in its outlook than the present paper, for there we are interested in *global maps* by which we mean a single univalent mapping which maps the canonical domain $\mathcal{S} = [0, 1] \times [0, 1]$ onto the entire problem domain $\mathcal{R}$. Given such a map, the sub-domains $\mathcal{E}_j$ needed in the subsequent transfinite (or finite) element solution of a boundary value problem are defined as the images under $\vec{T}$: $\mathcal{S} \mapsto \mathcal{R}$ of rectangular sub-domains of $\mathcal{S}$. Or, more practically, the boundary segments of the element domains $\mathcal{E}_j$ which need to be curved in order to conform to the exact geometry of the problem domain $\mathcal{R}$ are taken to be the images of constant s or t lines in $\mathcal{S}$; whereas, those $\mathcal{E}_j$ which are not immediately proximate to $\partial\mathcal{R}$ or to curved material interfaces interior to $\mathcal{R}$ (and hence need to be curved) are defined as quadrilaterals formed by joining the images of the vertices of a subrectangle in $\mathcal{S}$ by straight line segments. (See also the last paragraph on p. 217 of [23], and pp. 777–781 of [25].) In other words, the transfinite mapping function $\vec{T}$ is used as a generalized "mesh-generator" to define the subdomains $\mathcal{E}_j$. Since the subdomains arise from a transfinite map, they are themselves transfinite, in general. After defining the subdomains $\mathcal{E}_j$, interpolation of the dependent variable over each subdomain may be accomplished by any of the formulas derived earlier in this paper. Although the methods used to obtain invertible maps of $\mathcal{S} \to \mathcal{R}$ are easy to describe and to implement in practice, they rely heavily upon geometric and heuristic reasoning and are greatly facilitated by the use of computer generated graphic displays.

References

1. Barnhill, R. E., Birkhoff, G., Gordon, W. J.: Smooth interpolation in triangles. General Motors Research Report GMR-1064, Feb. 1971. To appear in Journal on Approximation Theory
2. Birkhoff, G.: Piecewise bicubic interpolation and approximation in polygons. In: Approximations with Special Emphasis on Spline Functions. Edited by I. J. Schoenberg. Academic Press 1969
3. Carlson, R., Hall, C.: Ritz approximations in two-dimensional boundary value problems. Numer. Math. **18**, 171–181 (1971)
4. Cullen, C. G.: Matrices and linear transformations. Reading: Addison-Wesley 1966
5. Davis, P. J.: Interpolation and approximation. London: Blaisdell 1963
6. Ergatoudis, I., Irons, B., Zienkiewicz, O.: Curved isoparametric, "quadrilateral" elements for finite element analysis. Int. J. Solids Structures, **4**, 31–42 (1968)
7. Gordon, W. J.: Distributive lattices and the approximation of multivariate functions. In: Approximations with special emphasis on spline functions. Edited by I. J. Schoenberg. Academic Press 1969, and references cited therein
8. Gordon, W. J.: "Blending-function" methods of bivariate and multivariate interpolation and approximation. SIAM Numer. Anal. **8**, 158–177 (1971)
9. Gordon, W. J., Hall, C. A.: Geometric aspects of the finite element method: Construction of curvilinear coordinate systems and their applications to mesh generation. GMR-1286, September, 1972, to appear Int. J. Num. Meth. Eng.
10. Griffin, D., Kellogg, R.: A numerical solution for axially symmetrical and plane elasticity problems. Int. J. Solids Structures **3**, 781–794 (1967)
11. Isaacson, E., Keller, H.: Analysis of numerical methods. New York: John Wiley & Sons 1966
12. Mikhlin, S.: The problem of the minimum of a quadratic functional. San Francisco: Holden-Day, Inc. 1950
13. Mikhlin, S., Smolitskiy, K.: Approximate methods for solution of differential and integral equations. Elsevier 1967
14. Tong, P., Pian, T.: The convergence of FEM in solving linear elastic problems. Int. J. Solids Structures **3**, 865–879 (1967)
15. Zienkiewicz, O., Cheung, Y.: The finite element method in structural and continuum mechanics. London: McGraw-Hill Ltd. 1967
16. Zienkiewicz, O., Hollister, G., Eds.: Stress analysis, New York: John Wiley & Sons Ltd. 1965
17. Zienkiewicz, O., Irons, B., Ergatoudis, J., Ahmad, S., Scott, F.: Iso-parametric and associated element families for two- and three-dimensional analysis. In: Finite elements methods in stress analysis. TAPIR, Technical University of Norway, Frondheim, 1970
18. Zienkiewicz, O.: The finite element method. London: McGraw Hill 1971
19. Zienkiewicz, O., Phillips, D.: An automatic mesh generation scheme for plane and curved surfaces by "isoparametric" coordinates. Int. J. Num. Meth. Eng. **3**, 519–528 (1971)
20. Zlamal, M.: On the finite element method. Numer. Math. **12**, 394–409 (1968)
21. Zlamal, M.: On the finite element procedures for solving second order boundary value problems. Numer. Math. **14**, 42–48 (1969)
22. Zlamal, M.: Finite element procedure of the second order of accuracy. Numer. Math. **14**, 394–402 (1970)

23. Ciarlet, P., Raviart, P.: Interpolation theory over curved elements, with applications to finite element methods. Computer Methods in Appl. Mech. and Engineering **1**, 217–249 (1972)
24. Ciarlet, P., Raviart, P.: The combined effect of curved boundaries and numerical integration in isoparametric finite element methods, p. 409–474. In: The mathematical foundations of the finite element method with applications to partial differential equations. Edited by A. K. Aziz. Academic Press 1972
25. Gordon, W. J., Hall, C. A.: Geometric aspects of the finite element method, p. 769–784. In: The mathematical foundations of the finite element method with applications to partial differential equations. Edited by A. K. Aziz. Academic Press 1972

William J. Gordon
Mathematics Department
Research Laboratories
General Motors Corporation
Warren, Michigan 48090/USA

Charles A. Hall
Department of Mathematics
University of Pittsburgh
Pittsburgh, Pennsylvania 15213/USA