Chapter 6

Finite Element Spaces

One of the advantages of the finite element method is that it can be used with relative ease to find approximations to solutions of differential equations on general domains. So far we have only considered approximating in one dimension or in higher dimensions using rectangular elements. The goal of this chapter is to formally define a finite element, present some examples of commonly used elements and to establish a taxonomy for describing elements. Isoparametric elements, which are used for domains with curved boundaries, are discussed in a later chapter.

To precisely describe a particular finite element, it is not enough to give the geometric figure, e.g.,a triangle, rectangle, etc. One must also specify the degree of polynomial that is used. Does describing these two pieces of information uniquely determine the choice? In fact, no. If we recall in $\mathbb{R}^1$ using an interval as the geometric element and specifying a cubic polynomial on each interval does not completely describe the finite element because we can determine the cubic by function values at four points or by function and derivative values (as in Hermite cubic) at two points. Consequently, three pieces of information must be provided to give an adequate description of a finite element; we must specify the geometric element, the degree of polynomial, and the degrees of freedom which are used to uniquely determine the polynomial.

Once we have chosen a particular finite element, we subdivide the domain into a finite number of geometric elements; this meshing must be "admissible", i.e.,satisfy certain properties. We want to construct a finite element space, S^h, over this mesh which possesses specific properties. A basic property which we said is a distinguishing feature of the finite element method is that we use a piecewise polynomial which is a kth degree polynomial when restricted to the specific element. For conforming finite elements we require our finite element space to be a subspace of the underlying Hilbert space. For second order problems this space was $H^1(\Omega)$ or a subspace and for fourth order problems the underlying space was $H^2(\Omega)$. Consequently a second property we require is a global smoothness requirement on the space. Finally, for the finite element method to be computationally efficient we must be able to construct a basis which has small support. Before addressing some

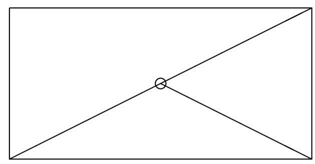


Figure 6.1. *Inadmissible triangulation due to "hanging node"*

of these issues we consider the admissible "triangulations" of a domain.

6.1 Construction of a finite element space

6.1.1 Admissible triangulations

Once a specific geometric element is chosen, we subdivide the domain $\bar{\Omega}$ into a finite number of individual subsets or geometric elements. We will use the terminology *triangulation* to refer to a subdivision of $\bar{\Omega}$ even if the specific geometric element is not a triangle. The subsets form a triangulation of $\bar{\Omega}$, denoted $\mathcal{T}^h$, which must satisfy certain properties. Some of these properties are obvious, such as the fact that their union is $\bar{\Omega}$, while others may not be as obvious. For example, we must add a condition which guarantees there are no "hanging nodes" as indicated in Figure 6.1.

Definition 6.1. *A subdivision $\mathcal{T}^h$ of Ω into subsets $\{\mathcal{K}_1, \mathcal{K}_2, \dots, \mathcal{K}_M\}$ is an admissible triangulation of Ω if it satisfies the following properties:*

(i) $\overline{\Omega} = \cup_{j=1}^{M} \mathcal{K}_j$ *;*

(ii) for each j, $j = 1, 2, \dots, M$, the set $\mathcal{K}_j$ is closed and the interior of $\mathcal{K}_j$ is non-empty;

(iii) for each $\mathcal{K}_j$, $j = 1, 2, \dots, M$, the boundary $\partial\mathcal{K}_j$ is Lipschitz continuous[1] *;*

(iv) if the intersection of two elements $\mathcal{K}_j$ and $\mathcal{K}_\ell$ is nonempty then the intersection must be a common vertex of the elements if the intersection is a single point; otherwise the intersection must be an entire edge or face common to both $\mathcal{K}_\ell$ and $\mathcal{K}_j$.

[1] A domain in Euclidean space with Lipschitz boundary is one whose boundary is "sufficiently regular". Formally, this means that the boundary can be written as, e.g., $z = f(x, y)$ where f is Lipschitz continuous. Recall that a function g is Lipschitz continuous if $\|g(p) - g(q)\| \leq C\,\|p - q\|$. for all p, q.

The penultimate condition in Definition 6.1 allows the application of Green's formula over each element.

The parameter h in the triangulation $\mathcal{T}^h$ is related to the size of the geometric elements and generally gives a measure of the coarseness or fineness of the mesh. It is usually taken to be the diameter of the largest element, i.e.,for $\vec{p}, \vec{q} \in \mathbb{R}^n$

$$h_j = \max_{\vec{p},\vec{q}\in\mathcal{K}_j} \left(\sum_{i=1}^{n} |p_i - q_i|^2 \right)^{1/2} , \quad j = 1, 2, \ldots, M$$

and

$$h = \max\{h_1, h_2, \ldots, h_M\} .$$

If we have a mesh where all the geometric elements are congruent, then the triangulation is *uniform* if all the elements are the same size; otherwise the triangulation is called *nonuniform*.

Clearly, we are interested in obtaining approximations on successively finer triangulations. For this reason, it is important to look at properties of families of triangulations. For example, we know that when we refine a mesh we can't just make the elements smaller in one portion of the domain but rather refine in some uniform way. To define the concept of a shape regular triangulation we introduce the parameter $\rho = \min_j \rho_j$ where ρ_j denotes the diameter of the largest ball contained in an element $\mathcal{K}_j$. Then a triangulation is called *shape regular* provided there exists a constant σ such that

$$\sigma = \frac{h}{\rho} . \tag{6.1}$$

A family of triangulations is called *shape regular* (or just *regular*) provided σ is uniform over the triangulations.

6.1.2 Formal definition of a finite element

From our previous examples in one and two dimensions, we saw that to completely describe a finite element we had to give more information than simply the choice of the geometric element and the degree of the polynomial. In fact, we need three pieces of information – the geometric element, the specific polynomial space defined over the geometric element, and the degrees of freedom needed to uniquely determine the polynomial. We follow Ciarlet's approach for the formal definition of a finite element.

Definition 6.2. *A finite element in $\mathbb{R}^n$ is a triple $(\mathcal{K}, \mathcal{P}_\mathcal{K}, \Theta_\mathcal{K})$ where*

(i) $\mathcal{K}$ is a closed subset of $\mathbb{R}^n$ with nonempty interior and a Lipschitz continuous boundary.

(ii) $\mathcal{P}_\mathcal{K}$ is a space of dimension s of real-valued functions over the set $\mathcal{K}$;

(iii) $\Theta_\mathcal{K}$ is a set of s linearly independent functionals, θ_i, $1 \le i \le s$, defined over the space $\mathcal{P}_\mathcal{K}$.

It is assumed that every $p \in \mathcal{P}_\mathcal{K}$ is uniquely determined by the values of the s functionals in $\Theta_\mathcal{K}$.

The set $\mathcal{K}$ is the specific geometric element in an admissible triangulation. The space $\mathcal{P}_\mathcal{K}$ usually consists of a polynomial defined over $\mathcal{K}$; however, we allow a broader definition so as to include some less common elements. In practice, we take these functions to be our basis for the space $\mathcal{P}_\mathcal{K}$. The set $\Theta_\mathcal{K}$ consists of the degrees of freedom which uniquely determine an element of $\mathcal{P}_\mathcal{K}$.

We can not arbitrarily choose a triple $(\mathcal{K}, \mathcal{P}_\mathcal{K}, \Theta_\mathcal{K})$ to define a finite element because $p \in \mathcal{P}_\mathcal{K}$ may not be uniquely determined by the degrees of freedom specified by $\Theta_\mathcal{K}$. An obvious example is the case where we don't have enough degrees of freedom specified; however, even if we have enough constraints they still may not uniquely determine the polynomials. To demonstrate that the degrees of freedom uniquely determine the polynomial several approaches can be taken. One approach is to show that the system of equations which results from imposing the degrees of freedom on an arbitary $p \in \mathcal{P}_\mathcal{K}$ has a unique solution. An alternate approach which is to actually construct a basis for the space $\mathcal{P}_\mathcal{K}$. We demonstrate both techniques when we consider specific finite elements.

6.1.3 Properties of finite element spaces

We subdivide the domain into a finite number of individual elements $\mathcal{K}_j$. On each $\mathcal{K}_j$ the polynomial space $\mathcal{P}_{\mathcal{K}_j}$ is specified along with the degrees of freedom which uniquely determine a polynomial $p \in \mathcal{P}_{\mathcal{K}_j}$ on the element $\mathcal{K}_j$. Then, an associated *finite element space* is defined through a systematic process. In every instance, this space is a finite-dimensional space of functions defined over $\overline{\Omega}$. An outline of the process is given as follows.

First, one defines the *local* properties with respect to each set $\mathcal{K}_j$ of the finite element space S^h. Restricted to each subset $\mathcal{K}_j \subset \overline{\Omega}$, functions belonging to S^h belong to $\mathcal{P}_{\mathcal{K}_j}$. Furthermore, over each $\mathcal{K}_j$, the functions in S^h are determined by the specified degrees of freedom.

Second, one defines the *global* properties with respect to $\overline{\Omega}$ of the finite element space. In particular, the desired order of global continuity and differentiability for S^h must be specified. For example, one could merely require that $S^h \subset C^0(\overline{\Omega})$ or it may be necessary to require that $S^h \subset C^1(\overline{\Omega})$.

The global properties are dictated by the differential equation which is being approximated. We have seen that for second order differential equations the underlying global smoothness of the finite element space is $S^h \subset H^1(\Omega)$ whereas for fourth order problems we require $S^h \subset H^2(\Omega)$. The question then arises how we can guarantee these global properties. Clearly the choice of local properties of S^h influences the global properties.

The following two propositions give conditions which guarantee the standard global smoothness conditions on S^h. The significance of the first proposition is that imposing the global smoothness $S^h \subset H^1(\Omega)$ does not require the functions in S^h to be continuously differentiable but merely continuous; this should be contrasted with the smoothness requirements for the classical solution of a second order boundary value problem. Similarly, the requirement $S^h \subset H^2(\Omega)$ only requires functions $v^h \in S^h$ to be in $C^1(\Omega)$. In the proposition, the additional assumption that $\mathcal{P}_{\mathcal{K}_j} \subset H^1(\mathcal{K}_j)$ is automatically satisfied when $\mathcal{P}_{\mathcal{K}_j}$ is a polynomial space on $\mathcal{K}_j$.

Proposition 6.3. *Assume that $\mathcal{T}^h$ is an admissible triangulation of $\overline{\Omega} \subset \mathbb{R}^n$ into the subsets $\{\mathcal{K}_j\}$. Let $\mathcal{P}_{\mathcal{K}_j} \subset H^1(\mathcal{K}_j)$ for all j, let $S^h \subset C^0(\overline{\Omega})$, and let $v^h|_{\mathcal{K}_j} \in \mathcal{P}_{\mathcal{K}_j}$ for all $v^h \in S^h$. Then $S^h \subset H^1(\Omega)$. Moreover, if S_0^h consists of those functions in S^h which vanish on the boundary of Ω, then*

$$S_0^h \equiv \{v^h \in S^h \ : \ v^h = 0 \ \textit{on}\ \partial\Omega\} \subset H_0^1(\Omega)\,.$$

Proof. Let $v^h \in S^h$; we must show that $v^h \in H^1(\Omega)$, i.e.,that $v^h \in L^2(\Omega)$ and that its first-order weak derivatives belong to $L^2(\Omega)$. Since $v^h \in C^0(\overline{\Omega})$ we have that $v^h \in L^2(\Omega)$. To demonstrate that its first-order weak derivatives are in $L^2(\Omega)$, we must find functions w_i^h, $i = 1, \ldots, n$, such that

$$\int_\Omega v^h \frac{\partial \phi}{\partial x_i}\, d\Omega = -\int_\Omega w_i^h \phi \, d\Omega \quad \forall\, \phi \in C_0^\infty(\Omega)\,.$$

For each i, we choose the function w_i^h to be the function whose restriction on each finite element $\mathcal{K}_j$ is the function $\partial(v^h|_{\mathcal{K}_j})/\partial x_i$; this is possible since $\mathcal{P}_{\mathcal{K}_j} \subset H^1(\mathcal{K}_j)$. Since each finite element $\mathcal{K}_j$ has a Lipschitz-continuous boundary $\partial\mathcal{K}_j$, we may apply Green's formula to obtain

$$\int_{\mathcal{K}_j} \frac{\partial}{\partial x_i}\left(v^h|_{\mathcal{K}_j}\right)\phi\, dx = -\int_{\mathcal{K}_j}\left(v^h|_{\mathcal{K}_j}\right)\frac{\partial \phi}{\partial x_i}\, dx + \int_{\partial\mathcal{K}_j} v^h|_{\mathcal{K}_j}\, \phi\, n_{i,\mathcal{K}_j}\, dS\,,$$

where $n_{i,\mathcal{K}_j}$ is the ith component of the unit outer normal along the boundary of $\mathcal{K}_j$. Summing over all the elements, we obtain

$$\int_\Omega w_i^h \phi\, d\Omega = -\int_\Omega v^h \frac{\partial\phi}{\partial x_i}\, d\Omega + \sum_j \int_{\partial\mathcal{K}_j} v^h|_{\mathcal{K}_j} \phi\, n_{i,\mathcal{K}_j}\, dS\,.$$

We are done if we can show that the last term vanishes. The boundary of the elements $\partial\mathcal{K}_j$ can be broken up into segments that are part of $\partial\Omega$ and segments that are also part of the boundary of an adjacent subset, say $\mathcal{K}_\ell$. In the first case, $\phi = 0$ so that clearly those terms vanish. In the other case, the boundary integrals from the two adjacent elements cancel since, by hypothesis, $v^h \in C^0(\overline{\Omega})$ and if two elements $\mathcal{K}_j$ and $\mathcal{K}_\ell$ are adjacent then on their common boundary, $n_{i,\mathcal{K}_j} = -n_{i,\mathcal{K}_\ell}$.

The fact that $S_0^h \subset H_0^1(\Omega)$ follows since $\partial\Omega$ is Lipschitz continuous and if $v^h \in S_0^h$, $v^h = 0$ on $\partial\Omega$. ∎

Proposition 6.4. *Assume that $\mathcal{T}^h$ is an admissible triangulation of $\overline{\Omega} \subset \mathbb{R}^n$ into the subsets $\{\mathcal{K}_j\}$. Let $\mathcal{P}_{\mathcal{K}_j} \subset H^2(\mathcal{K}_j)$ for all j, let $S^h \subset C^1(\Omega)$, and let $v^h|_{\mathcal{K}_j} \in \mathcal{P}_{\mathcal{K}_j}$*

for all $v^h \in S^h$. Then, $S^h \subset H^2(\Omega)$. Moreover, if S_b^h consists of all functions that vanish on the boundary, then

$$S_b^h \equiv \{v^h \in S^h \ : \ v^h = 0 \ \text{on} \ \partial\Omega\} \subset H^2(\Omega) \cap H_0^1(\Omega) \tag{6.2}$$

and if S_0^h consists of all functions that vanish on the boundary and whose derivative in the direction of the unit outer normal also vanish on the boundary, then

$$S_0^h \equiv \{v^h \in S^h \ : \ v^h = \frac{\partial v^h}{\partial \vec{n}} = 0 \ \text{on} \ \partial\Omega\} \subset H_0^2(\Omega)\,. \tag{6.3}$$

Proof. The proof is analogous to the proof of Proposition 6.3. The details are left to the exercises. ∎

6.2 Examples of finite elements on n-simplices

In $\mathbb{R}^2$ the common choices for a geometric element are a triangle and a quadrilateral. If the domain is polygonal and not rectangular, then triangular elements are needed to discretize. In $\mathbb{R}^3$ the commonly used elements are tetrahedra and cubes or bricks. In a later chapter we consider isoparametric elements to handle domains with curved boundaries. In this section we look at some of the more commonly used triangular elements and their variants.

We have seen that to completely specify a finite element, it is not enough to just choose a geometric element. We must also specify the degree of polynomial on the element and the degrees of freedom which uniquely determine the polynomial. To use the element we must also specify a basis which has small support. In the last chapter we saw that for rectangular elements we could simply use tensor products of the basis in one-dimension. For triangles or tetrahedra, this approach does not work. In the following section we see that barycentric coordinates are a useful tool in writing basis functions on a triangle or tetrahedron. In addition, in Section **??** we consider the approach of determining the basis functions on a reference element and mapping them to the desired element.

In this section and the next we develop a taxonomy for identifying finite elements whether in one, two or three dimensions. We identify the element by its geometric shape which is called an n-simplex or an $n-$rectangle; by its type which indicates the polynomial space, and by whether it is a Lagrange or Hermite element which indicates the kind of degrees of freedom used.

6.2.1 n-simplices

The first class of finite elements we consider uses subsets $\mathcal{K}$ of $\mathbb{R}^n$ that are *simplices*, e.g.,line segments in $\mathbb{R}^1$, triangles in $\mathbb{R}^2$ or tetrahedra in $\mathbb{R}^3$. Formally, we define an *n-simplex* in the following way.

Definition 6.5. *Let z_k, $k = 1, \ldots, n+1$, denote $n+1$ points in $\mathbb{R}^n$. The convex hull of these $n+1$ points, i.e.,the intersection of all convex sets [2] containing z_k, $k = 1, \ldots, n+1$, is called an n-simplex and the points z_k, $k = 1, \ldots, n+1$, are called the vertices of the n-simplex.*

For example, for $n = 2$ we specify three points $\{z_1, z_2, z_3\}$ and a 2-simplex is simply a triangle with vertices (z_{i_1}, z_{i_2}), $i = 1, 2, 3$, provided the three points are not collinear. To enforce the noncollinearity of the points, we require that the matrix

$$\begin{pmatrix} z_{1_1} & z_{2_1} & z_{3_1} \\ z_{1_2} & z_{2_2} & z_{3_2} \\ 1 & 1 & 1 \end{pmatrix}$$

is nonsingular. Note that the magnitude of the determinant of this matrix is just the area of the parallelogram formed by the vectors $z_2 - z_1$ and $z_3 - z_1$. For $n = 3$, we specify four points $\{z_1, z_2, z_3, z_4\}$ and a 3-simplex is just a tetrahedron with vertices z_i, $i = 1, \ldots, 4$, provided the four points are not coplanar, i.e.,provided the matrix

$$\begin{pmatrix} z_{1_1} & z_{2_1} & z_{3_1} & z_{4_1} \\ z_{1_2} & z_{2_2} & z_{3_2} & z_{4_2} \\ z_{1_3} & z_{2_3} & z_{3_3} & z_{4_3} \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

is nonsingular. Note that the magnitude of the determinant of this matrix is the volume of the parallelepiped formed by the vectors $z_i - z_1$, $i = 2, 3, 4$.

For an integer j such that $1 < j \leq n$, any j-simplex whose vertices are a subset of the $(n+1)$ vertices of a given n-simplex is called a *j-face* of the n-simplex. An (n-1)-face is simply called a *face* and any 1-face is called an *edge*. In $\mathbb{R}^2$, triangles have edges, i.e.,1-faces. In $\mathbb{R}^3$, tetrahedra have faces (2-faces) and edges (1-faces.)

6.2.2 Barycentric coordinates

A geometric concept which is useful in easily writing polynomial basis functions on an n-simplex is the idea of barycentric coordinates which were first defined by Möbius in 1827 (Coexeter 1969, p 27; Fauvel 1993). We know that if we are given a frame in $\mathbb{R}^n$, then we can define a local coordinate system with respect to the frame; e.g.,Cartesian coordinates. If we are given a set of $n+1$ points in $\mathbb{R}^n$ then we can also define a local coordinate system with respect to these points; such coordinate systems are called *barycentric coordinates.*

Suppose we are given a set of $n+1$ points $z_k \in \mathbb{R}^n$, $k = 1, \ldots, n+1$, such that

[2] Recall that a set $\mathcal{S}$ is convex if given any two points x and y in $\mathcal{S}$ then the line segment joining x and y lies entirely in $\mathcal{S}$.

the determinant of the matrix

$$\begin{pmatrix} z_{1_1} & z_{2_1} & \cdots & z_{n+1_1} \\ z_{1_2} & z_{2_2} & \cdots & z_{n+1_2} \\ \vdots & \vdots & \vdots & \vdots \\ z_{1_n} & z_{2_n} & \cdots & z_{n+1_n} \\ 1 & 1 & \cdots & 1 \end{pmatrix} \tag{6.4}$$

is nonzero. As we have seen, this is just the condition which guarantees in $\mathbb{R}^2$ that the points are not collinear and in $\mathbb{R}^3$ that the points are not coplanar. Consider the set of all linear combinations of these points of the form

$$q = \lambda_1 z_1 + \lambda_2 z_2 + \cdots \lambda_{n+1} z_{n+1}$$

where

$$\sum_{j=1}^{n+1} \lambda_j = 1\,.$$

Then the coordinates $(\lambda_1, \lambda_2, \dots, \lambda_{n+1})$ are called the *barycentric coordinates* of points of the space with respect to the given points z_k, $k = 1, \dots, n+1$.

As an example of barycentric coordinates consider three specific points in $\mathbb{R}^2$, $z_1 = (0,0)$, $z_2 = (1,0)$ and $z_3 = (1,1)$; the points form a triangle. Any linear combination of these three points such that $\lambda_1 + \lambda_2 + \lambda_3 = 1$ gives the barycentric coordinates (with respect to z_1, z_2, z_3) of a point in $\mathbb{R}^2$. For example, the barycentric coordinates $(\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$ is the point in space with Cartesian coordinates $(\frac{1}{2}, \frac{1}{4})$ since

$$\frac{1}{2}(0,0) + \frac{1}{4}(1,0) + \frac{1}{4}(1,1) = (\frac{1}{2}, \frac{1}{4})\,.$$

Similarly, the barycentric coordinates $(1,-1,1)$ is the point in space with Cartesian coordinates $(0,1)$ since

$$1 \cdot (0,0) + (-1) \cdot (1,0) + 1 \cdot (1,1) = (0,1)\,.$$

We notice that the point $(\frac{1}{2}, \frac{1}{4})$ with barycentric coordinates $(\frac{1}{2}, \frac{1}{4}, \frac{1}{4})$ lies within the triangle formed by the given z_i, $i = 1, 2, 3$ whereas the point $(0,1)$ with barycentric coordinates $(1,-1,1)$ is not inside the triangle. In general, one can demonstrate that if $0 \le \lambda_1, \lambda_2, \lambda_3 \le 1$ then the point $q = \lambda_1 z_1 + \lambda_2 z_2 + \lambda_3 z_3$ lies inside the triangle. If any λ_k, $k = 1, 2, 3$, is less than zero or greater than one, then the point q lies outside the triangle. If, for example, $\lambda_1 = 0$, then the point q lies on the edge of the triangle through z_2 and z_3. The justifications of these statements are left to the exercises.

Suppose now we are given a point $(x_1, x_2, \dots, x_n)$, in a Cartesian coordinate system or some other frame and want to determine the barycentric coordinates of the point with respect to a given set of $n+1$ points. The barycentric coordinates $(\lambda_1, \lambda_2, \dots, \lambda_{n+1})$ of the point with respect to the prescribed points $n+1$ points z_1,

$z_2, \ldots, z_{n+1}$ are found by solving the system

$$\begin{aligned} \sum_{j=1}^{n+1} z_{j_i}\lambda_j &= x_i \qquad i = 1, \ldots, n \\ \sum_{j=1}^{n+1} \lambda_j &= 1\,. \end{aligned} \tag{6.5}$$

Here z_{j_i} denotes the ith component of the point z_j. The coefficient matrix of (6.5) is just the matrix in (6.4) and hence we are guaranteed a unique solution. If we solve this system for the barycentric coordinates, then we see that the $\lambda_j(x)$, $j = 1, \cdots, n{+}1$, are linear functions of the coordinates of the point $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, i.e.,

$$\lambda_j = \sum_{k=1}^{n} \zeta_{j,k} x_k + \zeta_{j,n+1} \qquad j = 1, \ldots, n+1\,, \tag{6.6}$$

where $\zeta_{i,j}$ denotes the i,j entry of the inverse of the matrix given in (6.4). For example, the barycentric coordinates with respect to the points $z_1 = (0,0)$, $z_2 = (1,0)$ and $z_3 = (1,1)$ for the Cartesian point $(\frac{3}{4}, \frac{1}{2})$ are found by solving the system

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} \frac{3}{4} \\ \frac{1}{2} \\ 1 \end{pmatrix}$$

to get $(\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$. We can write the barycentric coordinates as

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{3}{4} \\ \frac{1}{2} \\ 1 \end{pmatrix}$$

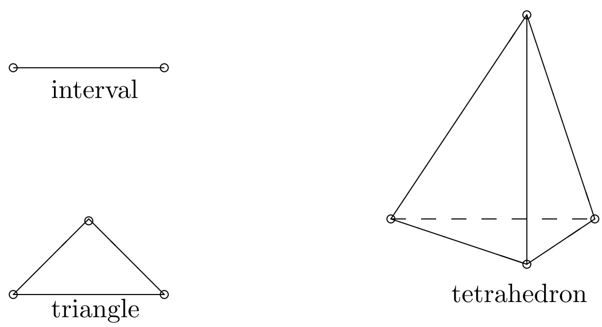
or in the form of (6.6) as $\lambda_1 = (-1)x_1 + (0)x_2 + 1$, $\lambda_2 = (1)x_1 + (-1)x_2 + 0$, etc.

We now want to see how barycentric coordinates can assist us in writing a basis for a polynomial space defined on a triangle or tetrahedron where we require that the basis is a nodal basis, i.e.,it has the property that it is one at one vertex and is zero at the other vertices. Consider the example of a 2-simplex, i.e.,a triangle, with vertices $\{z_1, z_2, z_3\}$. Then, the barycentric coordinates of a point $x = (x_1, x_2) \in \mathbb{R}^2$ are determined by solving the linear system

$$\begin{array}{ccccccc} z_{1_1}\lambda_1 & + & z_{2_1}\lambda_2 & + & z_{3_1}\lambda_3 & = & x_1 \\ z_{1_2}\lambda_1 & + & z_{2_2}\lambda_2 & + & z_{3_2}\lambda_3 & = & x_2 \\ \lambda_1 & + & \lambda_2 & + & \lambda_3 & = & 1\,. \end{array}$$

It is easy to see that if x is one of the vertices of the 2-simplex, say $x = z_k$, , then $\lambda_j(z_k) = \delta_{jk}$ where $\delta_{jk} = 0$ if $j \neq k$ and is equal to one if $j = k$. For example, if $x = z_1$ then the barycentric coordinates of x are $(1,0,0)$. Note also that $\lambda_1(x)$ is zero along the edge formed by z_2 and z_3 since it is a linear function which is zero at z_2 and z_3; thus the side of the triangle formed by the vertices z_2 and z_3 can be described by the equation $\lambda_1 = 0$.

Figure 6.2. *n-simplicies of type(1)*



Summarizing, we have that the barycentric coordinates $(\lambda_1, \lambda_2, \lambda_3)$ are linear functions of x (from (6.6)) which take on the values $(1, 0, 0)$ at $x = z_1$, the values $(0, 1, 0)$ at $x = z_2$ and $(0, 0, 1)$ at $x = z_3$. Consequently, λ_1 is a linear function of x which is one at the vertex z_1 and is zero at the other two vertices z_2 and z_3; similar conditions hold for λ_2 and λ_3. Hence these barycentric coordinates can serve as basis functions for the space of linear polynomials over the triangle formed by the points z_1, z_2, and z_3. When we consider quadratic or higher order basis functions we see that we can simply take appropriate products of the λ_j, $j = 1, \ldots, n + 1$.

6.2.3 Lagrange finite elements on n-simplices

When all the specified degrees of freedom are function values, then the finite element is referred to as a *Lagrange finite element*. Lagrange finite elements on n-simplices lead to finite element spaces that are subspaces of $C^0(\Omega)$ and hence by Proposition 6.3 they are subspaces of $H^1(\Omega)$. Such finite elements are often referred to as "C^0-elements". In the taxonomy of finite elements such elements are called *n-simplices of type (ℓ)* where the qualifier "type (ℓ)" refers to the degree of polynomial specified on the n-simplex.

Lagrange finite element on an n-simplices of type (1)

We first consider an n-simplex of type (1); i.e.,we are using a linear polynomial defined over an interval in $\mathbb{R}^1$, a triangle in $\mathbb{R}^2$ or a tetrahedron in $\mathbb{R}^3$. These are illustrated in Figure 6.2. We choose $\mathcal{P}(\mathcal{K}) = P_1(\mathcal{K})$ to be linear polynomials defined over $\mathcal{K}$. Since the $\dim(P_1(\mathcal{K})) = 3$ in $\mathbb{R}^2$ and $\dim(P_1(\mathcal{K})) = 4$ in $\mathbb{R}^3$, we expect a linear function on $\mathcal{K}$ to be uniquely determined by its values at the $n + 1$ nodes of the n-simplex. This can be proved in several ways; in the following proposition we prove the result using a linear algebra argument and then following the proof we outline an alternate argument.

Proposition 6.6. *Let $\mathcal{K}$ be an n-simplex in $\mathbb{R}^n$, $n = 1, 2, 3$, with vertices $z_1, \ldots, z_{n+1}$. A polynomial $p(x) \in P_1(\mathcal{K})$ is uniquely determined by its values at the vertices.*

Proof. We present a proof for the case $n = 2$ and leave the case $n = 3$ to the exercises; we have already addressed the case of an interval in $\mathbb{R}^1$. Let $p = c_0 + c_1 x_1 + c_2 x_2$ where c_0, c_1, c_2 are constants and let η_i, $i = 1, 2, 3$ be the prescribed values of $p(x)$ at the vertices. Then we must show that there is a unique function $p(x) \in P_1(\mathcal{K})$ such that $p(z_i) = \eta_i$, $i = 1, 2, 3$; i.e.,that the linear system

$$\eta_i = c_0 + c_1 z_{i_1} + c_2 z_{i_2} \quad \text{for } i = 1, 2, 3$$

has a unique solution. Note that the requirement that this coefficient matrix be nonsingular is equivalent to the condition which guaranteed that the vertices were not collinear in $\mathbb{R}^2$. ∎

Alternately, we could have shown that any polynomial $p(x) \in P_1(\mathcal{K})$ can be written in terms of its values η_i at the vertices. Recall that the barycentric coordinates satisfy $\lambda_i(z_k) = \delta_{ik}$ for $1 \leq i, k \leq 3$ so that in $\mathbb{R}^2$ the polynomial

$$\eta_1 \lambda_1(x) + \eta_2 \lambda_2(x) + \eta_3 \lambda_3(x)$$

has the desired property; i.e.,when we evaluate it at the vertices we get the nodal values. Thus any linear polynomial on an n-simplex with vertices $\{z_1, \ldots, z_{n+1}\}$ can be written as

$$p(x) = \sum_{i=1}^{n+1} p(z_i) \lambda_i(x) \,.$$

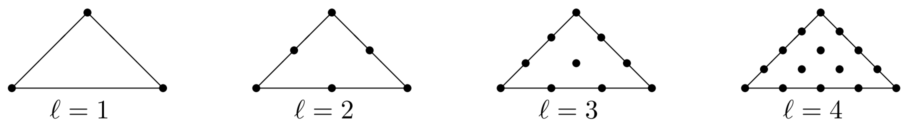
Summarizing, we define the 2-*simplex of type(1)* to be the set $\mathcal{K}$ where $\mathcal{K}$ is a triangle with vertices z_i, $i = 1, 2, 3$ together with the space $P_1(\mathcal{K})$ and the degrees of freedom of the finite element consisting of the values at the three vertices. A *3-simplex of type(1)* is a set $\mathcal{K}$, where $\mathcal{K}$ is a tetrahedron with vertices z_i, $i = 1, 2, \ldots, 4$, together with the space $P_1(\mathcal{K})$ and the degrees of freedom of the finite element being the values at the four vertices.

Lagrange finite element on n-simplices of type (2)

Results for n-simplices of type ℓ for $\ell > 1$ follow in an analogous fashion. In $\mathbb{R}^2$ we know that the dimension of $P_2(\mathcal{K})$ is six so we must specify a second degree polynomial at six points to uniquely determine it; the dimension of $P_3(\mathcal{K})$ is ten so that a third degree polynomial must be specified at ten points on the triangle. The most commonly chosen points are the obvious ones. These points form the *ℓ^{th} order principal lattice of an n-simplex* $\mathcal{K}$ given by

$$\begin{aligned} \mathcal{L}(\ell, n) \quad = \quad & \Big\{ x = \sum_{k=1}^{n+1} \sigma_k \, z_k \; : \; \sum_{k=1}^{n+1} \sigma_k = 1, \\ & \qquad \sigma_k \in \big\{0, \frac{1}{\ell}, \frac{2}{\ell}, \cdots, \frac{\ell - 1}{\ell}, 1\big\}, 1 \leq k \leq n+1 \Big\} \end{aligned} \tag{6.7}$$

Figure 6.3. ℓ^{th} *order lattice for a 2-simplex*



where $z_1, z_2, \ldots, z_{n+1}$ are the vertices of $\mathcal{K}$. It is easy to demonstrate that $\mathcal{L}(\ell, n)$ contains $\binom{\ell+n}{\ell}$ points. For example, in $\mathbb{R}^2$ for $\ell = 1$ $\sigma_k \in \{0, 1\}$ so that the points in $\mathcal{L}(1, 2)$ are $\{z_1, z_2, z_3\}$, i.e.,the vertices of the triangle. For $\ell = 2$, $\sigma_k \in \{0, \frac{1}{2}, 1\}$ so that

$$\mathcal{L}(2,2) = \left\{ z_1, z_2, z_3, \frac{z_1 + z_2}{2}, \frac{z_1 + z_3}{2}, \frac{z_2 + z_3}{2} \right\},$$

i.e.,the verticies of the triangle and the midpoints of the sides. The ℓth order principal lattice for a 2-simplex for $\ell = 1, 2, 3, 4$ is illustrated in Figure 6.3. The following proposition states that an ℓ^{th} order polynomial on an n-simplex is uniquely determined by its values at the points in the corresponding principal lattice.

Proposition 6.7. *Let $\mathcal{K}$ be an n-simplex in $\mathbb{R}^n$ with vertices z_k, $1 \le k \le n+1$. Then for a given integer $\ell \ge 1$, any polynomial $p \in P_\ell$ is uniquely determined by its value at the points in $\mathcal{L}(\ell, n)$ defined by (6.7).*

Proof. The proof is left to the exercies. ∎

We now know that any quadratic polynomial on an n-simplex is uniquely determined by its values at the nodes and the midpoints of the edges of the n-simplex. If we can write any $p \in P_2(\mathcal{K})$ in terms of the specified values at these nodes then we will have a basis for the space. For example, for a 2-simplex we want to write

$$p(x) = \sum_{i=1}^{3} p(z_i) q_i(x) + p(z_{12}) q_{12}(x) + p(z_{13}) q_{13}(x) + p(z_{23}) q_{23}(x),$$

where q_i, $i = 1, 2, 3$, and q_{12}, q_{13}, and q_{23} are quadratic functions on $\mathcal{K}$ and z_{ij} represents the midpoint of the edge joining the nodes z_i and z_j. We use products of the linear barycentric coordinates to write these quadratic functions which serve as our basis functions with small support. First, consider the function q_1 which is a quadratic function which has the properties $q_1(z_1) = 1$ and $q_1(x) = 0$ at the five points z_2, z_3, z_{12}, z_{13}, and z_{23}. Recall that $\lambda_1(x)$ is a linear function such that $\lambda_1(z_1) = 1$, $\lambda_1(z_2) = \lambda_1(z_3) = 0$, so in barycentric coordinates the equation of the line through z_2 and z_3 is just $\lambda_1 = 0$; similarly, the equation through the midpoints z_{12} and z_{13} is $\lambda_1 = 1/2$. Since the point z_{23} lies on the line $\lambda_1 = 0$ we have that the quadratic function

$$\lambda_1(x)\left(\lambda_1(x) - \frac{1}{2}\right)$$

vanishes at the five points z_2, z_3, z_{12}, z_{13}, and z_{23}. Hence we choose $q_1(x) = C\lambda_1(x)\left(\lambda_1(x) - \frac{1}{2}\right)$ and normalize so that $q_1(z_1) = 1$. Since $\lambda_1(z_1) = 1$ we set $C = 2$. In a similar manner $q_i = \lambda_i(x)(2\lambda_i(x) - 1)$, $i = 2, 3$. Now we must construct a quadratic function q_{12} which has the properties that $q_{12}(z_{12}) = 1$ and $q_{12}(x) = 0$ at the vertices and the remaining midpoints. In this case the equation of the line through z_1 and z_3 is $\lambda_2 = 0$ and the line through z_2 and z_3 is $\lambda_1 = 0$. Thus the quadratic $\lambda_1(x)\lambda_2(x)$ has the property that it is zero at the verticies and the midpoints z_{23}, z_{13} and takes on the value one-fourth at z_{12}; consequently we take $q_{12}(x) = 4\lambda_1(x)\lambda_2(x)$. In general, $q_{ij}(x) = 4\lambda_i(x)\lambda_j(x)$. Combining these results we have that for $p \in P_2(\mathcal{K})$ where $\mathcal{K}$ is a 2-simplex

$$p = \sum_{i=1}^{3} p(z_i)\lambda_i(2\lambda_i - 1) + 4p(z_{12})\lambda_1\lambda_2 + 4p(z_{13})\lambda_1\lambda_3 + 4p(z_{23})\lambda_2\lambda_3\,.,$$

For a n-simplex

$$p = \sum_{i=1}^{n+1} p(z_i)\lambda_i(2\lambda_i - 1) + \sum_{\substack{i,k=1\\ i<j}}^{n+1} 4p(z_{ik})\lambda_i\lambda_k \quad \forall\, p \in \mathcal{P}_2(\mathcal{K})\,. \tag{6.8}$$

Recall that to determine the barycentric coordinates with respect to the points z_i, $i = 1, \ldots, n+1$ we had to solve an $(n+1) \times (n+1)$ linear system of equations.

We now define the *n-simplex of type (2)* to be an n-simplex $\mathcal{K}$ together with the space $P_2(\mathcal{K})$ and the degrees of freedom consisting of the values at the vertices and the midpoints of the edges. Properties of the n-simplex of type(2) are summarized in Table 6.1

The cases $\ell \geq 3$ can be handled in a similar manner. Their properties are summarized in Table 6.1. See the exercises for details.

6.2.4 Hermite 2-simplices

In our examples so far in this chapter we have considered Lagrange finite elements whose degrees of freedom were function values at a prescribed set of points and the resulting finite element spaces were subspaces of $H^1(\Omega)$. In the examples in this section, we consider finite elements in which some of the degrees of freedom are partial derivatives, or more generally, directional derivatives. We denote the partial derivative of a function $p(x)$ in the direction of the line segment through two points $a, b \in \mathbb{R}^n$ and evaluated at a point $x = c \in \mathbb{R}^n$ by $D_{[a,b]}p(c)$. Of course, knowledge of the directional derivatives at a point is equivalent to the knowledge of the partial derivatives.

Hermite 2-simplex of type(3)

Recall that in $\mathbb{R}^1$ we used the space of cubic Hermite polynomials to construct a subspace of $H^2(\Omega)$. However, we see that in $\mathbb{R}^2$ (and also in $\mathbb{R}^3$) using Hermite cubics generates a finite element space which is only a subspace of $C^0(\overline{\Omega})$ and thus only a subspace of $H^1(\Omega)$ by Proposition 6.3. In the next section we consider an example of a $C^1(\Omega)$ triangular element in $\mathbb{R}^2$.

To uniquely determine a cubic polynomial on a triangle, we must specify ten conditions since $\dim P_3 = 10$. The following result gives ten degrees of freedom which are combinations of function and derivative values that uniquely determine a polynomial $p \in P_3$.

Proposition 6.8. *Let $\mathcal{K}$ be an 2-simplex with vertices z_i, $1 \leq i \leq 3$, and let $z_{123} = \frac{1}{3}(z_1 + z_2 + z_3)$. Then any polynomial $p(x)$ in the space $P_3(\mathcal{K})$ is uniquely determined by its value at the vertices, $p(z_i)$, $i = 1, 2, 3$ and the value of its two first partial derivatives at the vertices z_j, $1 \leq j \leq 3$, and its value at the point z_{123}.*

Proof. First note we are specifying 10 degrees of freedom and $dim(P_3) = 10$ in $\mathbb{R}^2$. To show uniqueness we demonstrate that if $p \in P_3(\mathcal{K})$ and ξ_i, η_{ij}, ζ are given values then the 10×10 system

$$\begin{aligned} p(z_i) &= \xi_i \quad \text{for } i = 1, 2, 3 \\ \frac{\partial p}{\partial x_j}(z_i) &= \eta_{ij} \quad \text{for } i = 1, 2, 3 \text{ and } j = 1, 2 \\ p(z_{123}) &= \zeta \end{aligned}$$

has a unique solution. An easy way to show this is to set all the given values, ξ_i, η_{ij} and ζ, to zero and prove that $p(x)$ must be identically zero.

If we show that $p \in P_3(\mathcal{K})$ is zero along each edge of the triangle, then we know that $p = \alpha\lambda_1\lambda_2\lambda_3$ for some constant α where λ_j are the barycentric coordinates defined by (6.5). Then, since $p(z_{123}) = 0$ we have that $\alpha = 0$ and thus $p(x)$ must be identically zero in $\mathcal{K}$. To demonstrate that $p \in P_3(\mathcal{K})$ is zero along each edge of the triangle we note that along the line containing the vertices z_i and z_j p is a cubic polynomial of one variable and hence we need four conditions to uniquely determine it. But $p(z_i) = p(z_j) = 0$ and that $D_{[z_i,z_j]}p(z_i) = D_{[z_i,z_j]}p(z_j) = 0$ and thus p is zero on each edge $[z_i, z_j]$. ∎

We can now define the finite element which is called the *Hermite 2-simplex of type(3)* where the partial derivatives at each vertex are degrees of freedom as well as the values at the vertices and the barycenter. Since knowledge of the directional derivatives at each vertex is equivalent to the knowledge of the partial derivatives at each vertex, we can specify either as degrees of freedom. The properties of the Hermite 2-simplex of type (3) are summarized in Table 6.1. Note that in the illustration of the element in the table we indicate the partial derivative degrees of freedom at z_i by a circle centered at z_i.

We now associate a finite element space S^h with a subdivision of $\overline{\Omega} \subset \mathbb{R}^2$ into Hermite 2-simplices of type(3). Then a function $v^h \in S^h$ implies that the restriction $v^h|_{\mathcal{K}_j}$ is in the space $\mathcal{P}_{\mathcal{K}_j} = P_3(\mathcal{K}_j)$ for each $\mathcal{K}_j$ and is defined by its values at all the vertices of the subdivison, its values at the centers of gravity of all the triangles, and the values of its two first partial derivatives at all the vertices of the subdivision. If we assume that we have an admissible triangulation of our domain into 2-simplices then we are able to obtain the following result.

Lemma 6.9. *Let S^h be the finite element space associated with Hermite 2-simplices of type(3). Then the inclusion*

$$S^h \subset C^0(\overline{\Omega}) \cap H^1(\Omega) \tag{6.9}$$

holds.

Proof. Because of Proposition 6.3 it suffices to show that $S^h \subset C^0(\overline{\Omega})$. Along any common side of two adjacent triangles, there is a unique polynomial of degree three in one variable which takes on the prescribed values and prescribed first derviatives at the endpoints of the side yielding a total of four conditions and thus uniquely determines a cubic in one variable. ■

It is tempting to think that the inclusion $S^h \subset C^1(\overline{\Omega})$ holds for Hermite n-simplices of type(3); however, this is not the case. Although the tangential derivative along an edge is continuous from element to element, the normal derivative is not.

Finally, we should produce a basis set consisting of functions of minimial support. As before, we can use the barycentric coordinates to write a polynomial $p \in P_3$ in terms of its values at the vertices and the barycenter, and the six values of its directional derivatives at the vertices; ultimately they are used to construct a basis for our corresponding finite element space. In particular, we want to write any $p \in P_3$ as a linear combination of appropriate cubic polynomials times the value of p and its partial derivatives at the vertices, z_i, $i = 1, 2, 3$, and its value at the barycenter z_{123}. For example, the cubic basis function at the vertex z_1 should have the property that it is one at z_1, zero at z_2, z_3, z_{123} and, in addition, its partial derivatives at nodes z_i, $i = 1, 2, 3$, should be zero. Specifically, for all $p \in P_3(\mathcal{K})$, $\mathcal{K} \subset \mathbb{R}^2$

$$\begin{aligned} p(x) &= \sum_{i=1}^{3} p(z_i)\left(-2\lambda_i^3 + 3\lambda_i^2 - 7\lambda_1\lambda_2\lambda_3\right) \\ &\quad + 27p(z_{123})\lambda_1\lambda_2\lambda_3 + \sum_{i=1}^{3}\sum_{\substack{j=1\\ j\neq i}}^{3} D_{[z_i,z_j]}p(z_i)\lambda_i\lambda_j(2\lambda_i + \lambda_j - 1)\,. \end{aligned} \tag{6.10}$$

It is easy to see that when we evaluate $p(x)$ given by (6.10) at z_i, $1 \le i \le 3$ and at z_{123} we get the corresponding function values $p(z_i)$, $1 \le i \le 3$, and $p(z_{123})$. It is a little more difficult to show that when we evaluate $D_{[z_i,z_j]}p(x)$ at z_i then the terms multiplying $p(z_i)$ and $p(z_{123})$ are zero and the polynomial multiplying $D_{[z_i,z_j]}p(z_i)$ is one. The proof of this is left to the exercises but basically we must show that $D_{[z_i,z_k]}(\lambda_i\lambda_j\lambda_k)(z_i) = 0$, when we differentiate the term $-2\lambda_i^3 + 3\lambda_i^2$ the terms cancel, and a relationship of the form $D_{[z_i,z_k]}\lambda_j = \delta_{jk} - \lambda_j(z_i)$, $1 \le k \le 3$, $k \neq i$ then if $k = j \neq i$ we get the desired result.

6.2.5 C^1 elements on n-simplices

For fourth order differential equations, the inclusion $S^h \subset H^2(\Omega)$ is needed; however, none of the examples presented so far satisfy this condition. Recall that the difficulty in the Hermite 2-simplex was the fact that the normal derivatives did not agree along an edge common to two adjacent elements.

The Argyris triangle

The first C^1 element which we consider is the *Argyris triangle* which uses a complete polynomial of degree five. The degrees of freedom consist of function values and first and second derivatives at the vertices in addition to normal derivatives at the midpoints of the sides. It can be shown that in $\mathbb{R}^2$ any $p(x) \in P_5$ is uniquely determined by the 21 degrees of freedom given by

$$\Theta_K = \{D^\alpha p(z_i), |\alpha| \leq 2, 1 \leq i \leq 3, \frac{\partial}{\partial n_i} p(z_{jk}), 1 \leq i \leq 3\},$$

where n_i denotes the normal along the edge of the triangle formed by z_j, z_k, $j \neq k \neq i$ and z_{jk} denotes the midpoint of that edge. Note that we have used multi-index notation to denote the derivatives to simplify the statement of the degrees of freedom. The *Argyris 21-degree of freedom triangle* is illustrated in Table 6.1 where we use | to indicate normal derivatives and a circle to indicate derivatives at vertices.

A finite element space is constructed in the usual manner. Since we require the normal derivative at the midpoint of each edge to be a degree of freedom, we expect the normal derivative as well as the tangential derivative along an edge to be continuous. The following result demonstrates that the finite element space generated by using the Argyris triangle is a subspace of $H^2(\Omega)$ and thus can be used to approximate fourth order problems.

Proposition 6.10. *Let S^h be the finite element space associated with the Argyris triangle. Then the inclusion*

$$S^h \subset C^1(\overline{\Omega}) \cap H^2(\Omega)$$

holds.

Proof. By Proposition 6.4, it suffices to show that $S^h \subset C^1(\overline{\Omega})$. Let $\mathcal{K}_i$ and $\mathcal{K}_j$ be two adjacent triangles with a common side $[b_k, b_\ell]$ where b_k, b_ℓ denote vertices of the triangulation and let $v^h \in S^h$. Considered as functions of an abscissa t along $[b_k, b_\ell]$ the functions $v^h|_{K_i}$ and $v^h|_{\mathcal{K}_j}$ are polynomials of degreee five in the variable t. Call these polynomials q_1 and q_2. Since, by the definition of the space S^h, we have

$$q(b_k) = q'(b_k) = q''(b_k) = q(b_\ell) = q'(b_\ell) = q''(b_\ell) = 0$$

where $q = q_1 - q_2$; it then follows that $q = 0$ and hence the inclusion $S^h \subset C^0(\overline{\Omega})$ holds. Likewise, call r_1 and r_2, the restrictions to the side $[b_k, b_\ell]$ of the functions

$\dfrac{\partial}{\partial \vec{n}} v^h|_{K_i}$ and $\dfrac{\partial}{\partial \vec{n}} v^h|_{K_j}$. Then r_1 and r_2 are polynomials of degree four in the variable t and again, from the definition of S^h, we have the five conditions

$$r(b_k) = r'(b_k) = r(b_{k\ell}) = r(b_2) = r'(b_\ell) = 0$$

where $r = r_1 - r_2$ and $b_{k\ell}$ is the midpoint of the side $[b_k, b_\ell]$. Therefore, $r = 0$. We have thus shown the continuity of the normal derivative. Since $q = 0$ along $[b_k, b_\ell]$, $q' = 0$ along $[b_k, b_\ell]$ also. Therefore, the first derivatives are also continuous on $\overline{\Omega}$. ■

One difficulty with the Argyris triangle is that there are 21 degrees of freedom. A modification to the Argyris triangle is the *Bell element* which suppresses the values of the normal slopes at the nodes at the three midpoint sides, reducing the degrees of freedom to 18. Functions in the finite element space associated with the Bell element are in a space P_B where $P_4 \subset P_B \subset P_5$. Here P_B denotes the space of all fifth degree polynomials whose normal derivatives along each side of the triangle are third degree polynomials. Note that, in general, in the Argyris triangle the normal derivative of $p \in P_5$ along each edge is a fourth degree polynomial. In this element the degrees of freedom are

$$\Theta_K = \{D^\alpha p(z_i), |\alpha| \leq 2, 1 \leq i \leq 3\}.$$

The determination of the basis functions for both the Argyris and Bell triangles is somewhat involved. The reader is referred to [**?**] for details.

Hsieh-Clough-Toucher triangles

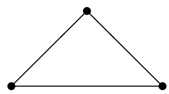
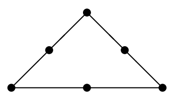
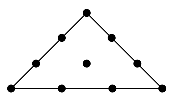
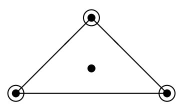
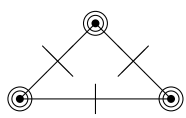
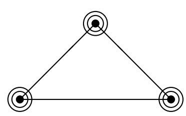
In an effort to create an element which generates a finite element space that is a subspace of $H^2(\Omega)$ but which has fewer degrees of freedom, researchers have developed composite type elements commonly called *macro elements*. In the Hsieh-Clough-Tocher triangle, the triangle is first decomposed into three triangles by connecting the barycenter of the given triangle with each of its vertices. On each of the subtriangles a cubic polynomial is constructed so that the resulting function is C^1 on the original triangle. There are a total of 12 degrees of freedom which consist of the function values and first partial derivatives at the three vertices of the original triangle in addition to the normal derivative at the midpoints of the sides of the original triangle.

There is also a reduced Hsieh-Clough-Toucher triangle where the degrees of freedom have been reduced to nine. Once again, the construction of the basis functions are involved; the reader is referred to [**?**, **?**] for details.

6.3 Examples of finite elements on n-rectangles

In this section we assume that $\overline{\Omega} \subset \mathbb{R}^n$ is a region that can be subdivided into rectangular elements. Many of the results are analogous to those when we subdivide a polyhedral region into n-simplices.

Table 6.1. *Triangular elements*

degrees of freedom	element	$\mathcal{P}_\ell(\mathcal{K})$	$\dim \mathcal{P}_\mathcal{K}$
	2-simplex of type (1)	$\mathcal{P}_1(\mathcal{K})$	3
	2-simplex of type (2)	$\mathcal{P}_2(\mathcal{K})$	6
	2-simplex of type (3)	$\mathcal{P}_3(\mathcal{K})$	10
	Hermite cubic 2-simplex	$\mathcal{P}_3(\mathcal{K})$	10
	Argyris triangle	$\mathcal{P}_5(\mathcal{K})$	21
	Bell triangle	$\mathcal{P}_B \subset \mathcal{P}_5(\mathcal{K})$	18

We let Q_ℓ, for positive integers ℓ, be the space of all polynomials of degree less than or equal to ℓ with respect to *each* of the n variables $x_1, x_2, \ldots, x_n$. For example, if $n = 2$ and $\ell = 1$ Q_1 =span$\{1, x_1, x_2, x_1 x_2\}$. We note that we always have the inclusion $P_\ell \subset Q_\ell$ and in general,

$$\dim(Q_\ell) = (\ell + 1)^n . \tag{6.11}$$

We formally define an *n-rectangle* in $\mathbb{R}^n$ as a product of compact intervals with non-empty interiors.

Definition 6.11. *An n-rectangle, $\mathcal{K}$ in $\mathbb{R}^n$ is defined by*

$$\mathcal{K} = \prod_{i=1}^{n} [a_i, b_i] = \{\vec{x} = (x_1, x_2, \ldots, x_n) \ : \ a_i \le x_i \le b_i, 1 \le i \le n\} \tag{6.12}$$

for finite a_i, b_i for each $i = 1, \ldots, n$.

6.3.1 n-rectangles of type(ℓ)

As in the case of n-simplices, once we have chosen the degree of Q_ℓ then we must specify points for the degrees of freedom, i.e.,points where if we prescribe a polynomial of degree ℓ in the n-rectangle then the polynomial is uniquely determined. An easy way to specify the degrees of freedom is to consider a particular n-rectangle, namely the *unit hypercube* $[0,1]^n$ and specify the points on it. Then a linear mapping gives the points on an arbitrary n-rectangle. The following proposition gives a set of points which guarantees that a polynomial in Q_ℓ is uniquely determined by its values on the set.

Proposition 6.12. *A polynomial $p \in Q_\ell$ is uniquely determined by its values on the set*

$$\mathcal{M}(\ell, n) = \left\{ x = \left(\frac{i_1}{\ell}, \frac{i_2}{\ell}, \cdots, \frac{i_n}{\ell} \right) \in \mathbb{R}^n \ : \ i_j \in \{0, 1, \cdots, \ell\}, 1 \le j \le n \right\} . \tag{6.13}$$

Proof. See exercises. □

For example, in $\mathbb{R}^2$

$$\mathcal{M}(1,2) = \{(0,0), (0,1), (1,0), (1,1)\}$$

and in $\mathbb{R}^3$

$$\mathcal{M}(1,3) = \{(0,0,0), (0,1,0), (0,0,1), (1,0,0), (0,1,1), (1,0,1), (1,1,0), (1,1,1)\} .$$

Thus a 2-rectangle of type(1) consists of a rectangular element $\mathcal{K}$, the space of linear polynomials on $\mathcal{K}$, $Q_1(\mathcal{K})$, whose dimension is 4 and whose degrees of freedom consist of the values at the four vertices. Similarly a 3-rectangle of type(1) consists of a rectangular element $\mathcal{K}$, the linear polynomials on $\mathcal{K}$, $Q_1(\mathcal{K})$, whose dimension is 8 and whose degrees of freedom consist of the values at the eight vertices.

For n-rectangles of type(2)

$$\mathcal{M}(2,2) = \{(0,0), (0,1), (1,0), (1,1), (0, \frac{1}{2}), (\frac{1}{2}, 0), (\frac{1}{2}, \frac{1}{2}), (1, \frac{1}{2}), (\frac{1}{2}, 1)\}$$

Thus a 2-rectangle of type(2) consists of a rectangular element $\mathcal{K}$, the space of quadratic polynomials on $\mathcal{K}$, $Q_2(\mathcal{K})$, whose dimension is 9 and whose degrees of freedom consist of the values at the four vertices, the midpoints of the edges and the barycenter of the rectangle. Similar properties hold for a 3-rectangle of type(2), 2- and 3-rectangles of type(3).

6.3.2 Example of a rectangular C^1 element

For fourth order problems, the inclusion $S^h \subset H^2(\Omega)$ is needed. We can easily define a rectangular element in $\mathbb{R}^2$ for which $S^h \subset H^2(\Omega)$ holds. The element is

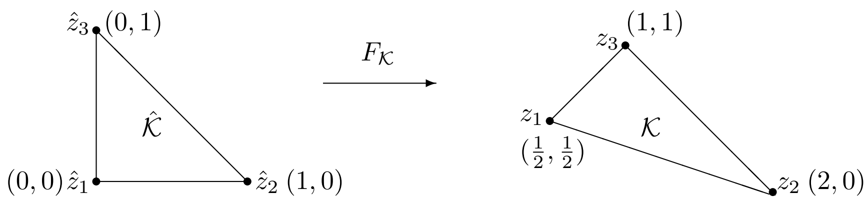


Figure 6.4. *Example of an affine transformation in a 2-simplex*

defined by prescribing $p(z_i)$, $\dfrac{\partial p}{\partial x_1}(z_i)$, $\dfrac{\partial p}{\partial x_2}(z_i)$, $\dfrac{\partial^2}{\partial x_1 \partial x_2}p(z_i)$ at the four vertices of the rectangular element. The resulting polynomial p is in the space $\mathcal{Q}_3$ which has dimension 16. The element is referred to as the *Bogner-Fox-Schmit rectangle.* The proof that the finite element space constructed in the usual manner using this element is a subspace of $C^1(\overline{\Omega})$ is left to the exercises.

6.4 Affine families of finite elements

In this section we want to demonstrate that for many choices of finite elements, instead of specifying a finite element discretization by the data $\mathcal{K}$, $P_{\mathcal{K}}$, and $\Theta_{\mathcal{K}}$, we can prescribe one reference finite element and the affine or linear function which maps the vertices of the reference element into the vertices of the geometric element in the admissible triangulation of the domain. We begin discussion of affine families of finite elements with an example.

We first consider the specific situation depicted in Figure 6.4 where we wish to find an affine mapping which maps the vertices of triangle $\widehat{\mathcal{K}}$ into the vertices of triangle K; i.e.,we seek $F_{\mathcal{K}}$ such that $F_{\mathcal{K}}(\hat{z}_i) = z_i$, $i = 1, 2, 3$ where $\hat{z}_i$ are the vertices of triangle $\widehat{\mathcal{K}}$ and z_i the vertices of triangle $\mathcal{K}$. In this case, $F_{\mathcal{K}}(\hat{x})$ can be explicitly written as

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = F_{\mathcal{K}}(\hat{x}_1, \hat{x}_2) = \frac{1}{2}\begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Clearly $F_{\mathcal{K}}$ maps the vertices in the reference triangle $\hat{\mathcal{K}}$ into the corresponding vertices in triangle $\mathcal{K}$. Moreover, since the mapping is linear, $F_{\mathcal{K}}(\frac{1}{2}, 0) = (\frac{5}{4}, \frac{1}{4})$, $F_{\mathcal{K}}(0, \frac{1}{2}) = (\frac{3}{4}, \frac{3}{4})$, and $F_{\mathcal{K}}(\frac{1}{2}, \frac{1}{2}) = (\frac{3}{2}, \frac{1}{2})$; i.e.,the midpoints are preserved under the transformation. In addition, the center of mass is preserved as well other points which we may use as degrees of freedom.

Suppose now that we choose $P_{\mathcal{K}} = P_1(\mathcal{K})$ and $P_{\hat{\mathcal{K}}} = P_1(\hat{\mathcal{K}})$ and we want to compare a basis function $\hat{\phi}_i \in P_1(\hat{\mathcal{K}})$ evaluated at a point $\hat{x}$ with the corresponding basis function in $P_{\mathcal{K}}$ evaluated at $x = F_{\mathcal{K}}(\hat{x})$. For example, the basis function $\hat{\phi}_3$ defined on $\hat{\mathcal{K}}$ which is associated with node $z_3 = (0, 1)$ is $\hat{\phi}_3 = \hat{x}_2$ and the basis function ϕ_3 defined on $\mathcal{K}$ which is associated with node $z_3 = (1, 1)$ is $\phi_3 = \frac{1}{2}x_1 + \frac{3}{2}x_2 - 1$. If we evaluate each basis function at, e.g.,the barycenter we get the same value, i.e.,$\hat{\phi}_3(\frac{1}{3},\frac{1}{3}) = \frac{1}{3}$ and $\phi_3(\frac{7}{6},\frac{1}{2}) = \frac{1}{3}$. This is because $(\frac{7}{6},\frac{1}{2}) = F_{\mathcal{K}}(\frac{1}{3},\frac{1}{3})$. Consequently, to evaluate basis functions on $\mathcal{K}$ at quadrature points on $\mathcal{K}$, we simply evaluate the corresponding basis function on the reference triangle at the corresponding quadrature point. However, this is not true when we deal with derivatives of basis functions as when we construct a stiffness matrix. For example, $\frac{\partial\hat{\phi}_3}{\partial\hat{x}_1} = 0$ and $\frac{\partial\phi_3}{\partial x_1} = \frac{1}{2}$. We shouldn't expect this to hold because we are differentiating with respect to different variables so clearly we must consider the transformation. The Jacobian of our transformation is given by

$$J = \begin{pmatrix} \frac{\partial x_1}{\partial\hat{x}_1} & \frac{\partial x_1}{\partial\hat{x}_2} \\ \frac{\partial x_2}{\partial\hat{x}_1} & \frac{\partial x_2}{\partial\hat{x}_2} \end{pmatrix} = \begin{pmatrix} \frac{3}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

By the chain rule we have

$$\frac{\partial\phi}{\partial\hat{x}_i} = \frac{\partial\phi}{\partial x_1}\frac{\partial x_1}{\partial\hat{x}_i} + \frac{\partial\phi}{\partial x_2}\frac{\partial x_2}{\partial\hat{x}_i}$$

so that

$$\begin{pmatrix} \frac{\partial\phi}{\partial\hat{x}_1} \\ \frac{\partial\phi}{\partial\hat{x}_2} \end{pmatrix} = \begin{pmatrix} \frac{\partial x_1}{\partial\hat{x}_1} & \frac{\partial x_2}{\partial\hat{x}_1} \\ \frac{\partial x_1}{\partial\hat{x}_2} & \frac{\partial x_2}{\partial\hat{x}_2} \end{pmatrix} \begin{pmatrix} \frac{\partial\phi}{\partial x_1} \\ \frac{\partial\phi}{\partial x_2} \end{pmatrix} = J^T \begin{pmatrix} \frac{\partial\phi}{\partial x_1} \\ \frac{\partial\phi}{\partial x_2} \end{pmatrix}.$$

Thus

$$\begin{pmatrix} \frac{\partial\phi}{\partial x_1} \\ \frac{\partial\phi}{\partial x_2} \end{pmatrix} = J^{-T} \begin{pmatrix} \frac{\partial\phi}{\partial\hat{x}_1} \\ \frac{\partial\phi}{\partial\hat{x}_2} \end{pmatrix}.$$

For our problem this just becomes

$$\begin{pmatrix} \frac{\partial\phi}{\partial x_1} \\ \frac{\partial\phi}{\partial x_2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{3}{2} \end{pmatrix} \begin{pmatrix} \frac{\partial\phi}{\partial\hat{x}_1} \\ \frac{\partial\phi}{\partial\hat{x}_2} \end{pmatrix}.$$

so that with $\hat{\phi}_3 = \hat{x}_2$ we have

$$\frac{\partial\phi_3}{\partial x_1} = \frac{1}{2}\frac{\partial\hat{\phi}_3}{\partial\hat{x}_1} - \frac{1}{2}\frac{\partial\hat{\phi}_3}{\partial\hat{x}_2} = 0 - \frac{1}{2} = \frac{1}{2}$$

$$\frac{\partial\phi_3}{\partial x_2} = \frac{1}{2}\frac{\partial\hat{\phi}_3}{\partial\hat{x}_1} + \frac{3}{2}\frac{\partial\hat{\phi}_3}{\partial\hat{x}_2} = 0 + \frac{3}{2} = \frac{3}{2}$$

which agrees with what we would get if we differentiated $\phi_3(x_1, x_2) = \frac{1}{2}x_1 + \frac{3}{2}x_2 - 1$.

In summary, we have seen that if we have a reference element and an affine function which maps the reference element into a particular $\mathcal{K}$ of our admissible triangulation, then all of the calculations can be performed on the reference element. Moreover, using a reference element and the linear map is a simple way to describe a family of finite elements.

Consider the case where we are given a family $(\mathcal{K}, \mathcal{P}_{\mathcal{K}}, \Theta_{\mathcal{K}})$ of triangles of type(2) and our goal is to describe this family as simply as possible. Let $\widehat{\mathcal{K}}$ be a reference triangle with vertices $\hat{z}_i$ and edge midpoints $\hat{z}_{ij} = (\hat{z}_i + \hat{z}_j)/2$, $1 \le i < j \le 3$, and let

$$\Theta_{\widehat{\mathcal{K}}} = \{\hat{p}(\hat{z}_i), 1 \le i \le 3;\ \hat{p}(\hat{z}_{ij}), 1 \le i < j \le 3\}$$

so that the element $(\widehat{\mathcal{K}}, P_{\widehat{\mathcal{K}}}, \Theta_{\widehat{\mathcal{K}}})$ with $P_{\widehat{\mathcal{K}}} = P_2(\widehat{\mathcal{K}})$ is also a triangle of type(2). Given any finite element $\mathcal{K}$ in the family, there exists a unique invertible affine mapping

$$F_{\mathcal{K}} : \hat{x} \in \mathbb{R}^2 \to F_{\mathcal{K}}(\hat{x}) = B_{\mathcal{K}}\hat{x} + b_{\mathcal{K}}$$

such that

$$F_{\mathcal{K}}(\hat{z}_i) = z_i,\ 1 \leq i \leq 3\,;$$

that is, $B_{\mathcal{K}}$ is an invertible 2×2 matrix and $b_{\mathcal{K}}$ a vector in $\mathbb{R}^2$. In the previous example we constructed a specific $F_{\mathcal{K}}$ of this form. Then it automatically follows that

$$F_{\mathcal{K}}(\hat{z}_{ij}) = z_{ij} \quad 1 \leq i < j \leq 3$$

since the property of a point being the midpoint of a line segment is preserved under an affine mapping. Likewise the points such as $z_{ijk} = \frac{1}{3}(z_i+z_j+z_k)$, $z_{iij} = \frac{2}{3}z_i+\frac{1}{3}z_j$, etc. keep their geometrical definitions through affine transformations. Once we have established the relation $\hat{x} \in \widehat{\mathcal{K}} \to x = F_{\mathcal{K}}(\hat{x}) \in \mathcal{K}$, between the points of the sets $\mathcal{K}$ and $\widehat{\mathcal{K}}$, it is natural to associate the spaces

$$\mathcal{P}^*_{\mathcal{K}} = \{p\ :\ \mathcal{K} \to \mathbb{R}^1;\ p = \hat{p}[F^{-1}_{\mathcal{K}}(x)],\ \hat{p} \in \mathcal{P}_{\widehat{\mathcal{K}}}\}$$

with the space $\mathcal{P}_{\mathcal{K}}$. Then it follows that

$$\mathcal{P}^*_{\mathcal{K}} = \mathcal{P}_{\mathcal{K}} = \mathcal{P}_2(\mathcal{K})$$

since the mapping $F_{\mathcal{K}}$ is affine.

In other words, rather than prescribing the family by the data $\mathcal{K}, \mathcal{P}_{\mathcal{K}}, \Theta_{\mathcal{K}}$, one cas prescribe one reference finite element $(\widehat{\mathcal{K}}, \mathcal{P}_{\widehat{\mathcal{K}}}, \Theta_{\widehat{\mathcal{K}}})$ and the affine mappings $F_{\mathcal{K}}$. Then for our example of a 2-simplex of type(2), a typical element in the family $(\mathcal{K}, \mathcal{P}_{\mathcal{K}}, \Theta_{\mathcal{K}})$ is such that

$$\left\{\begin{array}{rcl} \mathcal{K} & = & F_{\mathcal{K}}(\widehat{\mathcal{K}}) \\ \mathcal{P}_{\mathcal{K}} & = & \{p : \mathcal{K} \to \mathbb{R}^1\ :\ p = \hat{p}[F^{-1}_{\mathcal{K}}(x)],\ \hat{p} \in \mathcal{P}_{\widehat{\mathcal{K}}}\} \\ \Theta_{\mathcal{K}} & = & \{p[F_{\mathcal{K}}(\hat{z}_i)], 1 \leq i \leq 3;\ p[F_{\mathcal{K}}(\hat{z}_{ij})], 1 \leq i < j \leq 3\}\,. \end{array}\right.$$

With this example in mind, we can now give the general definition that two finite elements $(\widehat{\mathcal{K}}, \mathcal{P}_{\widehat{\mathcal{K}}}, \Theta_{\widehat{\mathcal{K}}})$ and $(\mathcal{K}, \mathcal{P}_{\mathcal{K}}, \Theta_{\mathcal{K}})$, with degrees of freedom of the form (**??**), are said to be *affine-equivalent* if there exists an *invertile affine mapping*

$$F : \hat{x} \in \mathbb{R}^n \to F(\hat{x}) = B\hat{x} + b \in \mathbb{R}^n$$

such that the following relations hold:

$$\mathcal{K} = F(\widehat{\mathcal{K}}) \tag{6.14}$$

$$\mathcal{P}_{\mathcal{K}} = \{p : \mathcal{K} \to \mathbb{R}^1; p = \hat{p}[F^{-1}(x)], \hat{p} \in \mathcal{P}_{\widehat{\mathcal{K}}}\} \tag{6.15}$$

$$\tag{6.16}$$

whenever the nodes z_i ($\hat{z}_i$ occur in the definitions of the set $\Theta_{\mathcal{K}}$ ($\Theta_{\widehat{\mathcal{K}}}$). It is clear that two n-simplices of type(ℓ) for a given $\ell \geq 1$ are affine-equivalent. Likewise, two n-rectangles of type(ℓ) are affine-equivalent through diagonal affine transformations.

Indeed, any two identical Lagrange finite elements that we have considered are affine-equivalent. The situation for Hermite elements is less simple. For example, consider two Hermite n-simplices of type(3) with sets of degrees of freedom involving $D_{[z,z_j]}p(z_i)$. Then it is clear that they are affine-equivalent because the relations

$$z_j - z_i = F(\hat{z}_j) - F(\hat{z}_i) = B(\hat{z}_j - \hat{z}_i), 1 \leq i, j \leq n, i \neq j \, .$$

On the other hand,the Argyris 21-degee of freedom triangle, is *not, in general, affine-equivalent* unless they are equilateral triangles since the normal derivative degrees of freedom are not preserved through an affine transformation, i.e.,the property of a vector that it be perpendicular to a hyperplane is not, in general, preserved through an affine mapping.

A family of finite elements is called an *affine family* if all its finite elements are affine-equivalent to a single finite element, which is called the *reference finite element* of the family. Note that the reference element, which we denote by $(\widehat{\mathcal{K}}, \mathcal{P}_{\widehat{\mathcal{K}}}\Theta_{\widehat{\mathcal{K}}})$ need not belong to the family. In the case of an affine family consisting of n-simplices, it is customary to choose the set $\widehat{\mathcal{K}}$ to be the *unit n-simplex* with vertices

$$\hat{z}_1 = (1, 0, \ldots, 0), \hat{z}_2 = (0, 1, 0, \ldots) \cdots \hat{z}_n = (0, 0, \ldots, 0, 1), \hat{z}_{n+1} = (0, 0, \ldots, 0)$$

for which the barycentric coordinates take the simple form

$$\lambda_i = x_i 1 \leq i \leq n, \text{and} \quad \lambda_{n+1} = 1 - \sum_{i=1}^{n} x_i \, .$$

In the case of an affine family of rectangular elements, the usual choice for the refence set $\widehat{\mathcal{K}}$ is either the unit hypercube $[0, 1]^n$ or the hypercube $[-1, 1]^n$.

The concept of affine family of finite elements is important because (i) in practical computations the calculations for the matrix entries are performed on the reference element; and (ii) for such families an elegant interpolation theory can be developed, which in turn is the basis for most of the convergence theorems concerning finite element approximations.