Let us now switch to triangular finite elements, which are better adapted for problems that are posed in open sets that are not just rectangles. First we need a quick review of affine geometry.

## 5.7  Barycentric coordinates

Triangular finite elements are much easier to work with using a system of coordinates in the plane that is quite different from the usual Cartesian system, namely barycentric coordinates. Actually, barycentric coordinates are a natural system of coordinates for affine geometry.

We will be given three points $A^1$, $A^2$ and $A^3$ in the plane. We first define the weighted barycenter of these points.

**Definition 5.7.1** *Let $\lambda_1$, $\lambda_2$ and $\lambda_3$ be three scalars such that $\lambda_1 + \lambda_2 + \lambda_3 = 1$. The barycenter of the points $A^j$ with weights $\lambda_j$ is the unique point M in the plane such that $\overrightarrow{OM} = \sum_{j=1}^{3} \lambda_j \overrightarrow{OA^j}$, where O is a given point. This point does not depend on the choice of O and we thus write*

$$M = \sum_{j=1}^{3} \lambda_j A^j.$$

One statement in this definition needs to be checked, namely that $M$ does not depend on $O$. Indeed, let $O'$ be another choice of point, and $M'$ be such that $\overrightarrow{O'M'} = \sum_{j=1}^{3} \lambda_j \overrightarrow{O'A^j}$. We have

$$\overrightarrow{O'M'} = \sum_{j=1}^{3} \lambda_j (\overrightarrow{O'O} + \overrightarrow{OA^j}) = \Big( \sum_{j=1}^{3} \lambda_j \Big) \overrightarrow{O'O} + \sum_{j=1}^{3} \overrightarrow{OA^j} = \overrightarrow{O'O} + \overrightarrow{OM} = \overrightarrow{O'M}$$

hence $M' = M$.

Now of course, barycenters are likewise defined for any finite family of points and weights of sum equal to 1, and in any affine space, but we will only use three points in the plane.

From now on, we assume that the three points $A^j$ are *not aligned*. In this case, we have the following basic result.

**Proposition 5.7.1** *For all points M in the plane, there exists a unique triple $(\lambda_1, \lambda_2, \lambda_3)$ of real numbers with $\lambda_1 + \lambda_2 + \lambda_3 = 1$ such that*

$$M = \sum_{j=1}^{3} \lambda_j A^j.$$

*The scalars $\lambda_i = \lambda_i(M)$ are called the* barycentric coordinates *of M, with respect to points $A^1, A^2, A^3$.*

*Proof.* We use Cartesian coordinates. Let $(x_1^j, x_2^j)$ be the Cartesian coordinates of $A^j$ in some Cartesian coordinate system, and $(x_1, x_2)$ be the Cartesian coordinates of point $M$. We have $M = \sum_{j=1}^3 \lambda_j A^j$ if and only if $x_k = \sum_{j=1}^3 \lambda_j x_k^j$ for $k = 1, 2$. Moreover, we have the condition $1 = \sum_{j=1}^3 \lambda_j$. We thus find a system of three linear equations in the three unknowns $\lambda_j$:

$$\begin{cases} \lambda_1 + \lambda_2 + \lambda_3 = 1, \\ x_1^1 \lambda_1 + x_1^2 \lambda_2 + x_1^3 \lambda_3 = x_1, \\ x_2^1 \lambda_1 + x_2^2 \lambda_2 + x_2^3 \lambda_3 = x_2. \end{cases}$$

The determinant of this system is

$$\Delta = \begin{vmatrix} 1 & 1 & 1 \\ x_1^1 & x_1^2 & x_1^3 \\ x_2^1 & x_2^2 & x_2^3 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ x_1^1 & x_1^2 - x_1^1 & x_1^3 - x_1^1 \\ x_2^1 & x_2^2 - x_2^1 & x_2^3 - x_2^1 \end{vmatrix} = (x_1^2 - x_1^1)(x_2^3 - x_2^1) - (x_1^3 - x_1^1)(x_2^2 - x_2^1) \neq 0$$

since it is equal to $\det(\overrightarrow{A^2 A^1}, \overrightarrow{A^3 A^1})$ and the points are not aligned.

Therefore, for any right-hand side, *i.e.*, for any point $M$, the system has one and only one solution. $\square$

**Remark 5.7.1** Going from barycentric coordinates to Cartesian coordinates is just done by applying the definition. Conversely, to compute barycentric coordinates from Cartesian coordinates, we just need to solve the above linear system.

If the three points are aligned, then we get a system which has a solution only if $M$ is on the line spanned by the points, and there are infinitely many solutions, and if the three points are equal, the system only has a solution if $M$ is equal to the other points, again with an infinity of solutions. $\square$

Let us give a few miscellaneous properties of barycentric coordinates.

**Proposition 5.7.2** *We have*

i) $\lambda_i(A^j) = \delta_{ij}$.

ii) *The functions $\lambda_i$ are affine in $(x_1, x_2)$ and conversely.*

iii) *Let $(A^i, A^j)$ denote the straight line passing through $A^i$ and $A^j$ for $i \neq j$. Then $(A^i, A^j) = \{M; \lambda_k(M) = 0, k \neq i, k \neq j\}$.*

iv) *Let $T$ be the closed triangle determined by the three points $A^j$. Then $T = \{M, 0 \leq \lambda_i(M) \leq 1, i = 1, 2, 3\}$.*

*Proof.* i) We have $A^1 = 1 \times A^1 + 0 \times A^2 + 0 \times A^3$ with $1 + 0 + 0 = 1$, hence by uniqueness of the barycentric coordinates, $\lambda_i(A^1) = \delta_{i1}$.

ii) Use Cramer's rule for solving the above linear system.

iii) The function $\lambda_k$ is a nonzero affine function by i) and ii), thus it vanishes on a straight line. By i), this straight line contains $A^i$ and $A^j$, so it is equal to $(A^i, A^j)$.

iv) We have just seen by iii) that $\lambda_k(M) = 0$ is the equation of the straight line opposite to vertex $A^k$. Moreover, by i) the half-plane containing $A_k$ is the half-plane $\{M; \lambda_k(M) \geq 0\}$. The triangle $T$ is the intersection of these three half-planes, so it is the set of points whose barycentric coordinates are all positive. Since their sum is equal to 1, they are also less than 1. □
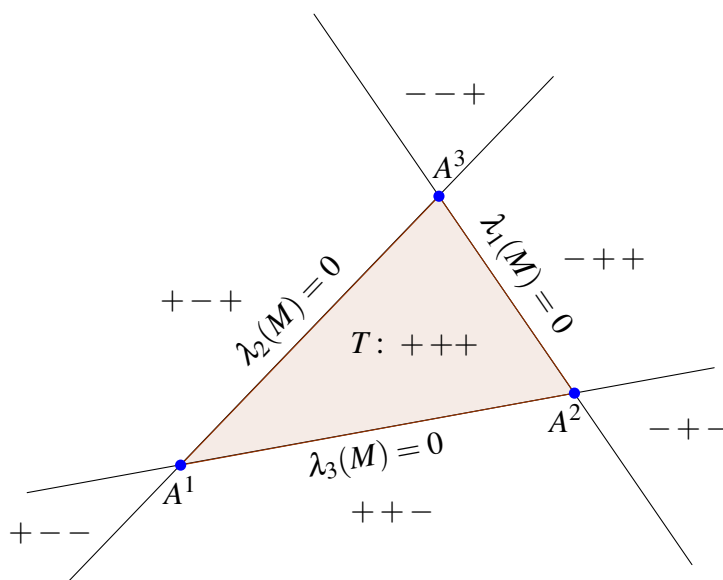


Figure 26. Signs of the barycentric coordinates in order $\lambda_1, \lambda_2, \lambda_3$.

Note that there is no $- - -$ region, it would be hard to have $\sum \lambda_i = 1$ in such a region... The last important feature of barycentric coordinates is their invariance under affine transformations. For this we modify the notation a bit by indicating the dependence on the points $A^j$ by writing $\lambda_i^{A^1, A^2, A^3}(M)$, which is admittedly cumbersome, and will thus not be used after this.

**Proposition 5.7.3** *Let F be an bijective affine transformation of the plane. Then we have*

$$\lambda_i^{F(A^1), F(A^2), F(A^3)}(F(M)) = \lambda_i^{A^1, A^2, A^3}(M)$$

*for $i = 1, 2, 3$ and all M.*

*Proof.* This is clear as affine transformations conserve barycenters. □

Let us give the barycentric coordinates of a few points of interest in a triangle:

- Middle of $A^1A^2$: $\left(\frac{1}{2}, \frac{1}{2}, 0\right)$,

- Middle of $A^2A^3$: $\left(0, \frac{1}{2}, \frac{1}{2}\right)$,

- Middle of $A^1A^3$: $\left(\frac{1}{2}, 0, \frac{1}{2}\right)$,

- Center of gravity of the triangle: $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$.

## 5.8   Triangular $P_1$ Lagrange elements

Let us be given a triangular mesh $\mathscr{T}$ on $\Omega$. We recall that $P_1$ denotes the space of polynomials of total degree less than 1, *i.e.*, affine functions. We define the corresponding approximation spaces

$$W_h = \{v_h \in C^0(\bar{\Omega}), v_{h|T_k} \in P_1 \text{ for all } T_k \in \mathscr{T}\},$$

without boundary conditions and

$$V_h = \{v_h \in h, v_h = 0 \text{ on } \partial\Omega\},$$

with boundary conditions. The general approximation theory applies and we thus just need to describe the approximation spaces in terms of finite elements and basis functions.

Let $T$ be a triangle with non aligned vertices $A^1$, $A^2$ and $A^3$.

**Proposition 5.8.1** *The finite element* $(T, P_1, \{p(A^1), p(A^2), p(A^3)\})$ *is unisolvent.*

*Proof.* We have $\dim P_1 = 3$ so the numbers match. The basis polynomials are obvious: $\lambda_1, \lambda_2, \lambda_3$, by Proposition 5.7.2, i) and ii).   □

**Proposition 5.8.2** *A function of $V_h$ is uniquely determined by its values at the internal nodes of the mesh and all sets of values are interpolated by an element of $V_h$.*

*Proof.* By unisolvence, three values for the three nodes of an element determine one and only one $P_1$ polynomial that interpolates these nodal values (we take the value 0 for the nodes located on the boundary). Therefore, if we are given a set of values for each node in the mesh, this set determines one $P_1$ polynomial per element. Let us check that they combine into a globally $C^0$ function.

Since the mesh is admissible, an edge common to two triangles $T_k$ and $T_{k'}$ is delimited by two vertices $A^1$ and $A^2$ which are also common to both triangles.

We thus have two $P_1$ polynomials $p$ and $p'$ such that $p(A^1) = p'(A^1)$ and $p(A^2) = p'(A^2)$. We parametrize the segment $[A^1, A^2]$ as $M = \mu A^1 + (1 - \mu)A^2$ with $\mu \in [0, 1]$. Then the restriction of $p - p'$ to this segment is a first degree polynomial in the variable $\mu$ that has two roots, $\mu = 0$ and $\mu = 1$. Therefore, $p - p' = 0$ on this segment, and the function defined by $p(x)$ if $x \in T_k$, $p'(x)$ if $x \in T_{k'}$ is continuous on $T_k \cup T_{k'}$. $\qquad\square$

**Corollary 5.8.1** *Let us be given a numbering of the nodes $S^k$, $k = 1, \ldots, N_i$. There is a basis of $V_h$ composed of the functions $w_h^i$ defined by $w_h^i(S^j) = \delta_{ij}$ and and for all $v_h \in V_h$, we have*

$$v_h = \sum_{i=1}^{N_i} v_h(S^i) w_h^i. \tag{5.13}$$

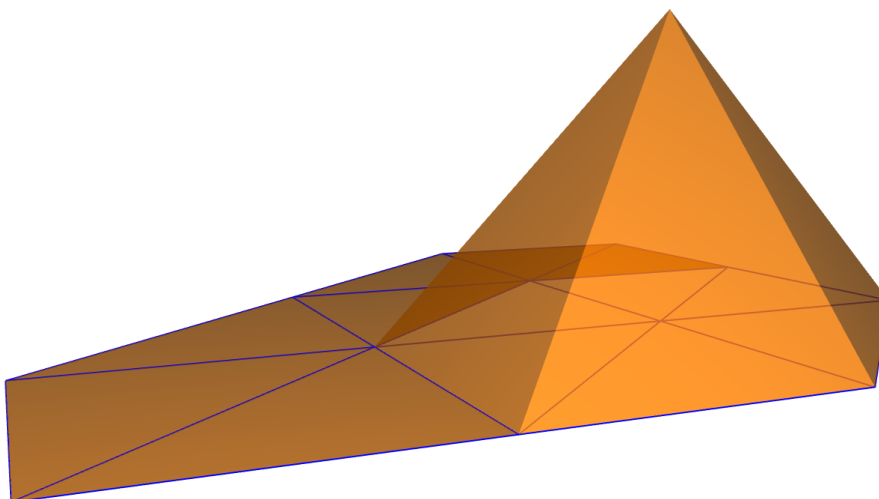*Proof.* Same as before. $\qquad\square$



Figure 27. A $P_1$ basis function on a triangular mesh.

Other pictures can be found on the class Web page.

Let us now talk a little bit about matrix assembly. We will not touch on the node numbering issue, which is clearly more complicated in a triangular mesh than in a rectangular mesh, especially in an *unstructured* triangular mesh, such as that shown in Figure 2, in which there is no apparent natural numbering.

We will however see how the use of a reference triangle and of barycentric coordinates simplifies the computation of matrix coefficients. We have the same element-wise decomposition as in the rectangular case

$$A_{ij} = \sum_{k=1}^{N_i} A_{ij}(T_k),$$

with

$$A_{ij}(T_k) = \int_{T_k} (\nabla w_h^j \cdot \nabla w_h^i + c w_h^j w_h^i) \, dx.$$

On each triangle $T_k$, the basis functions either vanish or are equal to one barycentric coordinate. So we need to compute the integral of the product of two barycentric coordinates (for $c$ constant) and the integral of the scalar product of their gradient.

We thus introduce a reference triangle

$$\widehat{T} = \{(\hat{x}_1, \hat{x}_2) \in \mathbb{R}^2, \hat{x}_1 \geq 0, \hat{x}_2 \geq 0, \hat{x}_1 + \hat{x}_2 \leq 1\}.$$

Let $\hat{A}^1 = (0,0)$, $\hat{A}^2 = (1,0)$ and $\hat{A}^3 = (0,1)$ be its vertices. Let $T_k$ be a generic triangle in the mesh, with vertices $A_k^1$, $A_k^2$, $A_k^3$. Now, there exists one and only one affine mapping $F_k$ such that $F_k(\hat{A}^j) = A_k^j$, $j = 1, 2, 3$. Indeed, since affine mappings conserve barycenters, we simply have

$$F_k(\widehat{M}) = \hat{\lambda}_1(\widehat{M}) A_k^1 + \hat{\lambda}_2(\widehat{M}) A_k^2 + \hat{\lambda}_3(\widehat{M}) A_k^3,$$

or in other words, $\lambda_i(F_k(\widehat{M})) = \hat{\lambda}_i(\widehat{M})$, where the first barycentric coordinates are taken relative to the vertices of $T_k$ in increasing superscript order.

Now the expression of barycentric coordinates in the reference triangle in terms of Cartesian coordinates is particularly simple:

$$\hat{\lambda}_1 = 1 - \hat{x}_1 - \hat{x}_2, \quad \hat{\lambda}_2 = \hat{x}_1, \quad \hat{\lambda}_3 = \hat{x}_2,$$

whereas they are fairly disagreeable in the generic triangle.

Let us give an example of computation with the integral $\int_{T_k} \lambda_2^2 \, dx$. We are going to use the change of variables $x = F_k(\hat{x})$. Since this change of variable is affine, its Jacobian $J$ is constant, and we have

$$\text{area } T_k = \int_{T_k} dx = \int_{\widehat{T}} J \, d\hat{x} = \frac{J}{2}$$

therefore $J = 2 \, \text{area } T_k$. Now we can compute

$$\begin{aligned}
\int_{T_k} \lambda_2^2(x) \, dx &= \int_{\widehat{T}} \hat{\lambda}_2^2(\hat{x}) J \, d\hat{x} \\
&= 2 \, \text{area } T_k \int_{\widehat{T}} \hat{x}_1^2 \, d\hat{x} \\
&= 2 \, \text{area } T_k \int_0^1 \hat{x}_1^2 \left( \int_0^{1-\hat{x}_1} d\hat{x}_2 \right) d\hat{x}_1 \\
&= 2 \, \text{area } T_k \int_0^1 \hat{x}_1^2 (1 - \hat{x}_1) \, d\hat{x}_1 \\
&= 2 \, \text{area } T_k \left( \frac{1}{3} - \frac{1}{4} \right) \\
&= \frac{\text{area } T_k}{6}.
\end{aligned}$$

Exchanging the vertices, we find $\int_{T_k} \lambda_1^2(x)\,dx = \int_{T_k} \lambda_3^2(x)\,dx = \frac{\text{area}\,T_k}{6}$. A similar computation shows that $\int_{T_k} \lambda_i \lambda_j(x)\,dx = \frac{\text{area}\,T_k}{12}$ for all $i \neq j$. Such terms are thus of the order of $h^2$.

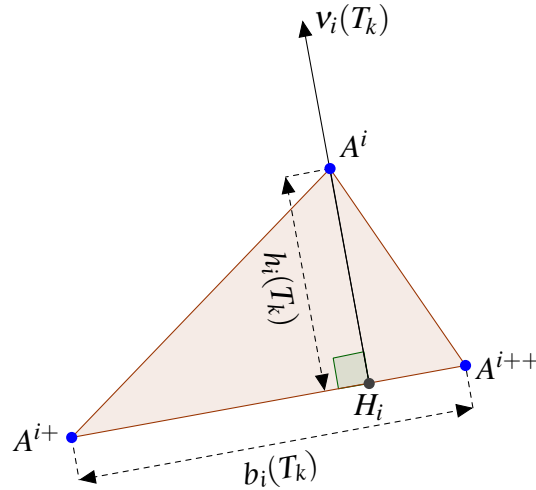Let us now turn to the gradient terms. We first need to compute $\nabla \lambda_i$, which is a constant vector.



Figure 28. Geometric elements of a generic triangle.

We introduce $h_i(T_k)$ and $b_i(T_k)$ respectively the height and base of $T_k$ relative to $A^i$, $H_i$ the foot of the altitude of $A^i$ and $v_i(T_k)$ the unit vector perpendicular to the base and pointing from the base toward $A^i$. Since $\lambda_i$ is affine, we have for all points $M$

$$\lambda_i(M) = \lambda_i(H_i) + \nabla \lambda_i \cdot \overrightarrow{H_i M}.$$

Now $H_i$ lies on the straight line $(A^{i+}, A^{i++})$, thus $\lambda_i(H_i) = 0$. Since $\lambda_i$ vanishes on this straight line, it follows that $\nabla \lambda_i = \mu v_i(T_k)$ for some scalar $\mu$. Taking $M = A^i$, we obtain

$$1 = \mu v_i(T_k) \cdot \overrightarrow{H_i M} = \mu h_i(T_k).$$

Therefore, we have

$$\nabla \lambda_i = \frac{1}{h_i(T_k)} v_i(T_k) = \frac{b_i(T_k)}{2\,\text{area}\,T_k} v_i(T_k).$$

It follows from instance that

$$\|\nabla \lambda_i\|^2 = \frac{b_i(T_k)^2}{4(\text{area}\,T_k)^2},$$

so that

$$\int_{T_k} \|\nabla\lambda_i\|^2 \, dx = \frac{b_i(T_k)^2}{4\,\text{area}\,T_k}.$$

These terms are of the order of 1. We could likewise compute $\int_{T_k} \nabla\lambda_i \cdot \nabla\lambda_j \, dx$ without difficulty, with expressions that involve the angles of $T_k$.

## 5.9   Triangular $P_2$ Lagrange elements

Let us go one step up in degree and consider $P_2$ elements. We have $\dim P_2 = 6$ as is shown by its canonical basis $(1, X, Y, X^2, XY, Y^2)$. This canonical basis is useless for our purposes and it is again much better to work in barycentric coordinates.

**Proposition 5.9.1** *The family $(\lambda_1^2, \lambda_2^2, \lambda_3^2, \lambda_1\lambda_2, \lambda_2\lambda_3, \lambda_1\lambda_3)$ is a basis of $P_2$.*

*Proof.* The functions $\lambda_i$ are affine, thus products $\lambda_i\lambda_j$ belong to $P_2$. We have a family of 6 vectors in a 6-dimensional space, it thus suffices to show that it is linearly independent. Let is be given a family of 6 scalars $\alpha_{ij}$ such that

$$\sum_{i\le j=1}^{3} \alpha_{ij}\lambda_i\lambda_j = 0.$$

Evaluating first this relation at point $A^k$, we obtain

$$0 = \sum_{i\le j=1}^{3} \alpha_{ij}\delta_{ik}\delta_{jk} = \alpha_{kk}$$

for all $k$. we are thus left with

$$\alpha_{12}\lambda_1\lambda_2 + \alpha_{13}\lambda_1\lambda_3 + \alpha_{23}\lambda_2\lambda_3 = 0.$$

We evaluate this relation at point $\frac{A^1+A^2}{2}$, the middle of $A^1$ and $A^2$, for which $\lambda_1 = \lambda_2 = \frac{1}{2}$ and $\lambda_3 = 0$. Hence
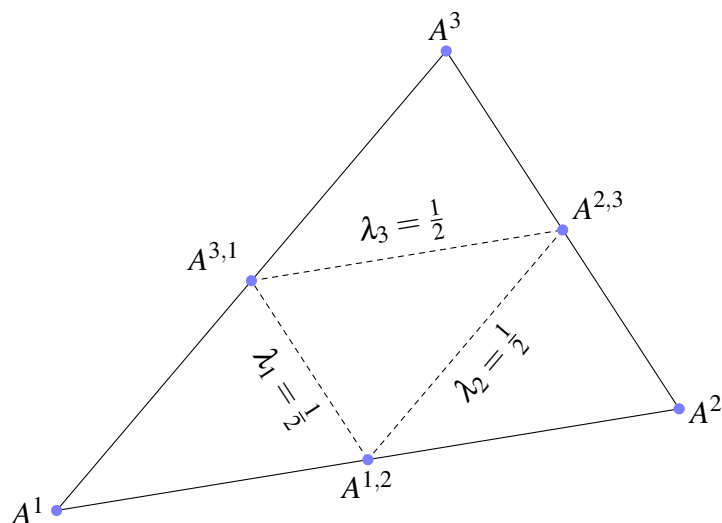
$$\frac{\alpha_{12}}{4} = 0,$$

and similarly $\alpha_{13} = \alpha_{23} = 0$.                                    □

We need 6 degrees of freedom of Lagrange interpolation. We take the three vertices $A^i$ and the three edge middles $A^{i,i+}$. Then we have

**Proposition 5.9.2** *The finite element $\left(T, P_2, \{p(A^i), p(A^{i,i+})\}_{i=1,2,3}\right)$ is unisolvent.*

Figure 29. The $P_2$ Lagrange triangle.

*Proof.* We have the right number of degrees of freedom with respect to the dimension of the polynomial space. It is thus sufficient to construct the basis polynomials. Everything being invariant by permutation of the vertices, it is clearly sufficient to construct for example, the basis polynomial corresponding to $A^1$ and that corresponding to $A^{1,2}$.

Let us start with $A^1$. We thus need a polynomial $p_1 \in P_2$ such that $p_1(A^1) = 1$ and $p_1$ vanishes at all the other nodes. We will freely use the obvious fact that the restriction of a polynomial of total degree at most $n$ in two variables to a straight line is a polynomial of degree at most $n$ in any affine parametrization of the straight line. Here, $p_1$ is of degree at most 2 on $(A^2, A^3)$, with three roots corresponding to points $A^2$, $A^{2,3}$ and $A^3$, thus it vanishes on $(A^2, A^3)$. The equation of the straight line is $\lambda_1 = 0$, hence $p_1$ is divisible by $\lambda_1$, *i.e.*, there exists a polynomial $q$ such that $p_1 = q\lambda_1$. Now $\lambda_1$ is of degree 1, therefore $q$ is of degree at most one. Moreover, since $\lambda_1(A^{1,2}) = \lambda_1(A^{3,1}) = \frac{1}{2} \neq 0$, we have $q(A^{1,2}) = q(A^{3,1}) = 0$. Therefore, by the same token, $q$ vanishes on the straight line $(A^{1,2}, A^{3,1})$, of equation $\lambda_1 - \frac{1}{2} = 0$. Thus $q$ is divisible by $\lambda_1 - \frac{1}{2}$, so that $q = c(\lambda_1 - \frac{1}{2})$ with $c$ of degree at most 0, *i.e.*, a constant. Finally, the relation $p_1(A^1) = 1$ yields $1 = \frac{c}{2}$, hence $p_1 = \lambda_1(2\lambda_1 - 1)$. Conversely, it is easy—but necessary—to check that this polynomial is in $P_2$ and satisfies the required interpolation relations.

To sum up, we have

$$p_1 = \lambda_1(2\lambda_1 - 1), \quad p_2 = \lambda_2(2\lambda_2 - 1), \quad p_3 = \lambda_3(2\lambda_3 - 1),$$

for the basis polynomials associated with the vertices.

Next we deal with $A^{1,2}$. Let us go faster. The polynomial must vanish on both lines $(A^1, A^3)$ and $(A^2, A^3)$, hence $p_{12} = c\lambda_1\lambda_2$ where $c$ is a constant. Using $p_{12}(A^{1,2}) = 1$ gives $c = 4$.

To sum up, we have

$$p_{12} = 4\lambda_1\lambda_2, \quad p_{23} = 4\lambda_2\lambda_3, \quad p_{31} = 4\lambda_1\lambda_3,$$

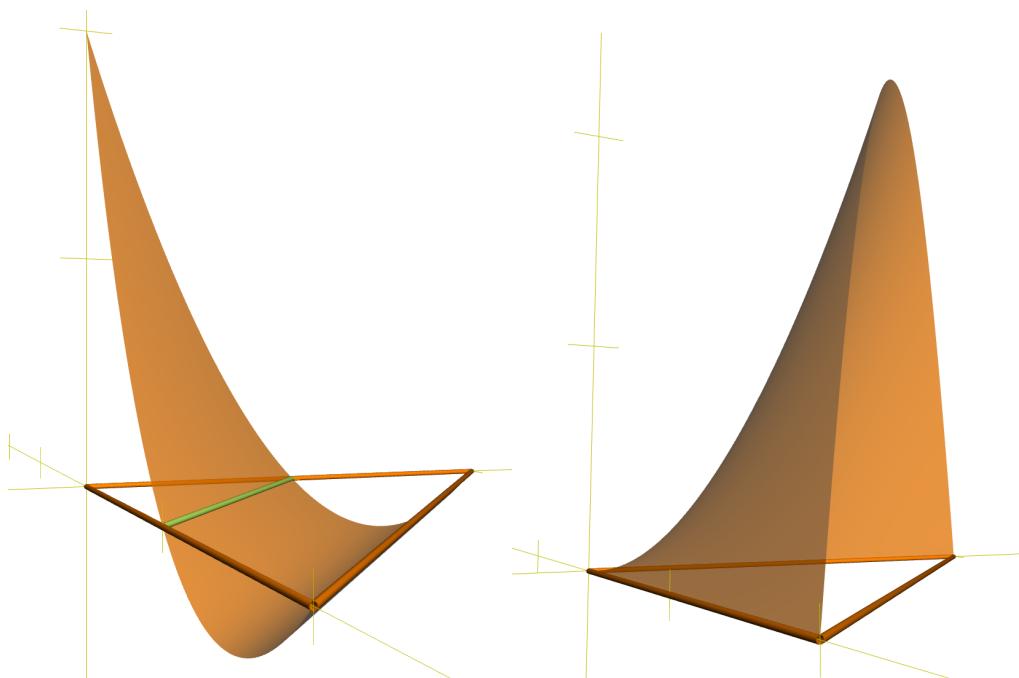for the basis polynomials associated with the middles and the $P_2$ Lagrange triangular element is unisolvent. □



Figure 32. The two different kinds of $P_2$ basis polynomials.

The approximation space

$$V_h = \{v_h \in C^0(\bar{\Omega}); v_{h|T_k} \in P_2, \forall T_k \in \mathscr{T}, v_h = 0 \text{ on } \partial\Omega\}$$

is of course endowed with set of basis functions that interpolate values at all nodes (vertices and middles). Let us quickly check the continuity across an edge. We thus have two polynomials of degree at most two, one on each side of the edge, that coincide at the vertices and the middle. Their restriction to the edge is of degree two in one variable, their difference has three roots, hence they are equal on the edge. The rest follows as before.
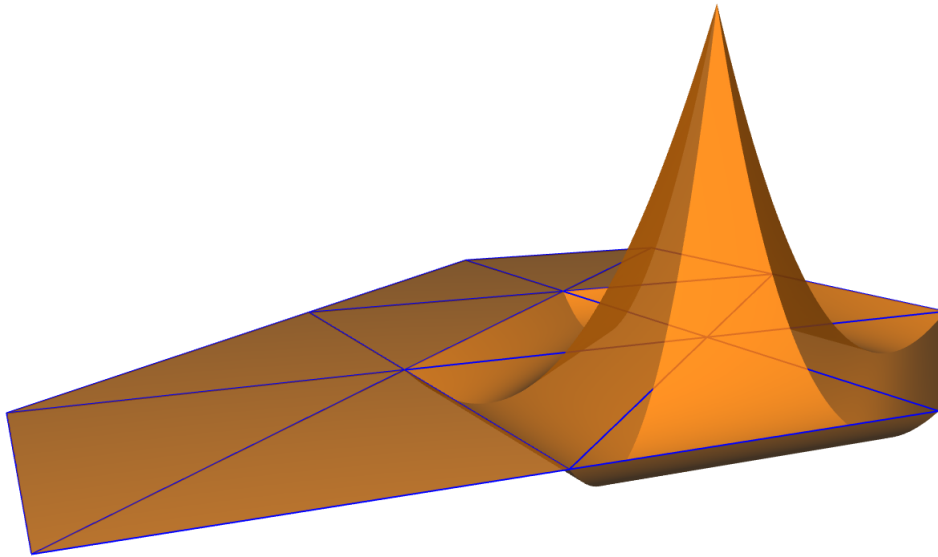
Figure 33. A basis function associated with a vertex.
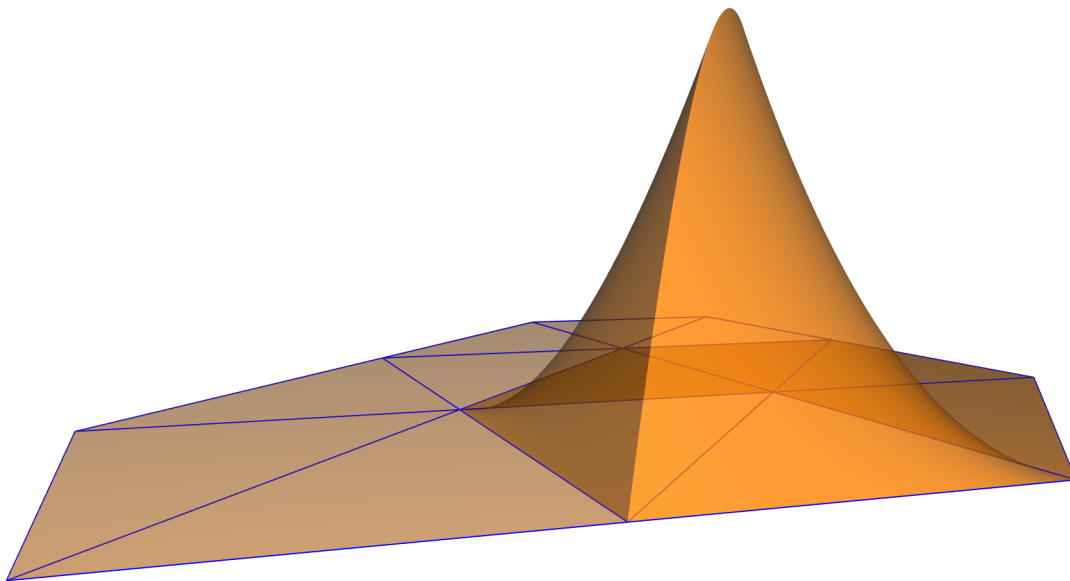


Figure 34. A basis function associated with a middle.

More pictures on the Web page.

Let us say a few words about $P_3$ Lagrange triangles. We have $\dim P_3 = 10$, thus 10 interpolation points are needed. We take the 3 vertices plus two points per edge, located at the thirds (this will obviously imply global continuity). That makes 9 points. A simple choice for the tenth point is then the center of gravity.
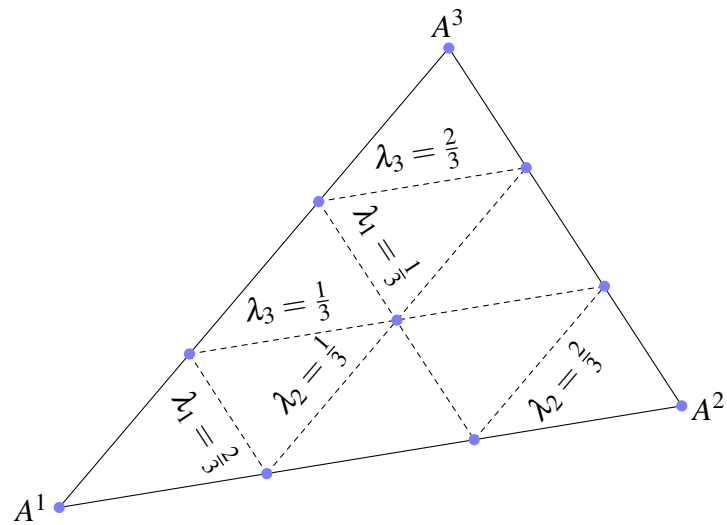
Figure 35. The 10 nodes of a $P_3$ Lagrange triangle.

Of course, this finite element is unisolvent. We list the basis polynomials:

$$p_0 = 27\lambda_1\lambda_2\lambda_3,$$

corresponding to the center of gravity, also called a *bubble* due to the shape of its graph.

$$p_1 = \frac{1}{2}\lambda_1(3\lambda_1 - 1)(3\lambda_1 - 2), \text{ etc.}$$

associated with the 3 vertices, and

$$p_{112} = \frac{9}{2}\lambda_1\lambda_2(3\lambda_1 - 1), \text{ etc.}$$
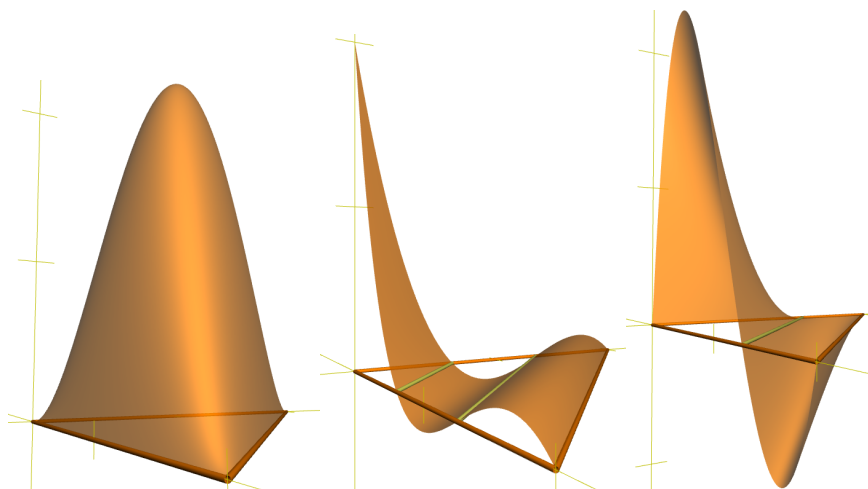
associated with the 6 edge nodes.



Figure 36. The three different kinds of $P_3$ basis polynomials.

Also of course, the approximation space

$$V_h = \{v_h \in C^0(\bar{\Omega}); v_{h|T_k} \in P_3, \forall T_k \in \mathcal{T}, v_h = 0 \text{ on } \partial\Omega\}$$

has the usual basis made of basis functions which we picture below.
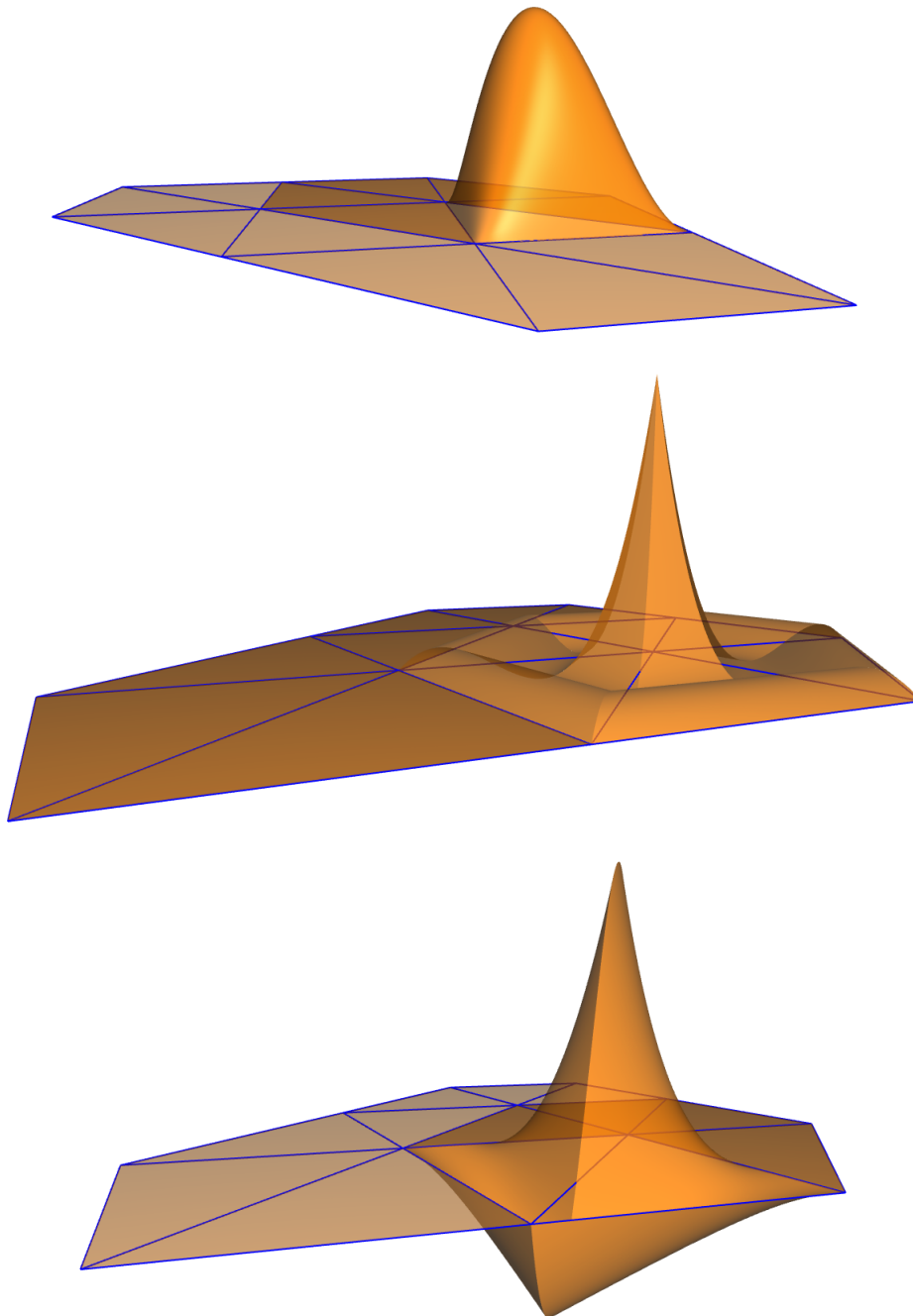
Figure 37. A few $P_3$ basis functions.

As in the rectangular case, the reason for facing the added complexity of taking higher degree polynomials is to achieve faster convergence. Indeed, we have the general result.

**Theorem 5.9.1** *Let us be given a regular family of triangulations indexed by h. We consider $P_k$ Lagrange elements on the triangulations. If $u \in H^{k+1}(\Omega)$, then we have*

$$\|u - u_h\|_{H^1(\Omega)} \leq Ch^k |u|_{H^{k+1}(\Omega)},$$

*where $|u|_{H^{k+1}(\Omega)} = |\nabla^{k+1} u|_{L^2(\Omega)}.$*

The proof is along the same lines as the proof in the $Q_1$ case, but with a lot more technicality due to the affine changes of variables between the reference triangle and the generic triangle.

All the above Lagrange triangular elements are adequate for $H^1$ approximation and are adapted to $C^0$ approximation spaces. It is also possible to define $C^1$ Hermite triangular elements for fourth order problems. One possible construction uses $P_5$ polynomials, hence 21 degrees of freedom. More generally, there is a very large diversity of triangular elements in the literature, sometimes especially crafted for one (class of) boundary value problem(s).