Chapter 3

Finite Element spaces

In the previous lectures we have studied the properties of coercive problems in an abstract setting and described Ritz and Galerkin methods for the approximation of the solution to a PDE, respectively in the case of symmetric and non-symmetric bilinear forms.

The abstract setting reads:

\[
\begin{align*}
\text{Find } u_h \in V_h \subset H \text{ such that:} \\
a(u_h, v_h) &= L(v_h), \quad \forall v_h \in V_h
\end{align*}
\]

such that:

- \( V_h \) is a finite dimensional approximation space characterized by a discretization parameter \( h \),
- \( a(\cdot, \cdot) \) is a continuous bilinear form on \( V_h \times V_h \), coercive w.r.t \( \| \cdot \|_V \),
- \( L(\cdot) \) is a continuous linear form.

Under these assumptions existence and uniqueness of a solution to the approximate problem holds owing to the Lax–Milgram Theorem and \( u_h \) is called \textit{discrete solution}. Provided this abstract framework which allows us to seek approximate solutions to PDEs, we need now to define the discrete space \( V_h \) and construct a basis \( (\varphi_1, \cdots, \varphi_{N_{V_h}}) \) of \( V_h \), \( N_{V_h} = \dim(V_h) \), on which the discrete solution is decomposed as

\[
u_h = \sum_{j=1}^{N_{V_h}} u_j \varphi_j
\]

with \( \{u_j\} \) a family of \( N_{V_h} \) real numbers called \textit{global degrees of freedom} and \( \{\varphi_j\} \) a family of \( N_{V_h} \) elements of \( V_h \) called \textit{global shape functions}.

Previously no assumption was made on the finite dimensional space \( V_n \) aside from that \( V_n \subset V \). The change of notation to \( V_h \) is to reflect that the discrete space \( V_h \) will be characterized more carefully as an \textit{approximation space} by constructing the shape functions and by defining the degrees of freedom.

To construct the Finite Element space \( V_h \), three ingredients are introduced:
1. An admissible mesh $\mathcal{T}_h$ generated by a tesselation of domain $\Omega$.

2. A reference Finite Element $(\hat{K}, \hat{P}, \hat{\Sigma})$ to construct a basis of $V_h$ and define the meaning of $u_j$.

3. A mapping that generates a Finite Element $(K, P, \Sigma)$ for any cell in the mesh from the reference element $(\hat{K}, \hat{P}, \hat{\Sigma})$.

As a preliminary step the approximation of the Poisson problem in one dimension by linear Lagrange Finite Elements is described to give an overview of the methodology without hitting the technical difficulties. Concepts and notations for the discretization of the physical domain are then introduced. Provided that all the requirements are identified, a framework for all the Finite Element methods is introduced by stating the definition of a Finite Element. Finally, the generation of the Finite Element space from a reference Finite Element will be described. Some examples of Finite Element spaces are listed at the end of the chapter.

### 3.1 A preliminary example in one dimension of space

For the sake of completeness, steps performed to derive a Galerkin method for the Poisson Problem 1.3 on $\Omega = (0, 1)$ are sketched below to recapitulate the methodology.

#### 3.1.1 Weak formulation

A solution is sought in the distributional sense by testing the equation against smooth functions,

$$ -\int_{\Omega} u''(x)v(x) \, dx = \int_{\Omega} f(x)v(x) \, dx , \quad \forall \, v \in C_0^\infty(\Omega) $$

then reporting derivatives on the test functions using the integration by part

$$ -\int_{\Omega} u''(x)v(x) \, dx = -\left[ u'(x)v(x) \right]_0^1 + \int_{\Omega} u'(x)v'(x) \, dx $$

the weak formulation consists of finding $u \in V$ such that

$$ \int_{\Omega} u'(x)v'(x) \, dx = \int_{\Omega} f(x)v(x) \, dx , \quad \forall \, v \in V $$

given $f \in L^2(\Omega)$. The choice of solution space and test space is guided by the equation and the data: in this case $V = H^1_0(\Omega)$ since $v$ and $v'$ should be controlled in $L^2(\Omega)$, and homogeneous Dirichlet boundary conditions are imposed.
3.1. A PRELIMINARY EXAMPLE IN ONE DIMENSION OF SPACE

3.1.2 Galerkin method

The approximate problem by a Galerkin method consists of seeking a discrete solution $u_h$ in a finite dimensional space $V_h \in V$, such that

$$\int_{\Omega} u_h'(x)v'(x) \, dx = \int_{\Omega} f(x)v(x) \, dx, \quad \forall \, v \in V_h$$

and given a basis $\{\varphi_j\}$ of $V_h$ any function $w \in V_h$ can be written as

$$w_h = \sum_{j=1}^{N_{V_h}} w_j \varphi_j$$

with $N_{V_h} = \dim V_h$, and moreover

$$w'_h = \sum_{j=1}^{N_{V_h}} w'_j \varphi'_j$$

by simple application of the derivation on the linear combination. Bounds of the sum will be omitted to simplify the notation,

$$\sum_{j=1}^{N_{V_h}} \sim \sum_j$$

when there is no possible confusion.

Inserting the Galerkin decomposition in the weak formulation and using the commutativity of the derivation with the linear combinations,

$$\int_{\Omega} \left( \sum_j u_j \varphi'_j(x) \right) \left( \sum_i v_i \varphi'_i(x) \right) \, dx = \int_{\Omega} f(x) \left( \sum_i v_i \varphi_i(x) \right) \, dx$$

and the integration can also commute with the linear combinations,

$$\sum_j \sum_i u_j v_i \int_{\Omega} \varphi'_j(x) \varphi'_i(x) \, dx = \sum_i v_i \int_{\Omega} f(x) \varphi_i(x) \, dx$$

so that up to some cosmetic reordering, for any $v \in V$

$$\sum_i \sum_j v_i \int_{\Omega} \varphi'_i(x) \varphi'_j(x) \, dxu_j = \sum_i v_i \int_{\Omega} f(x) \varphi_i(x) \, dx$$

the relation to the linear system of algebraic equations becomes evident,

$$v^T A u = v^T b$$

for any $v = [v_i] \in \mathbb{R}^{N_{V_h}}$, with

$$A = \left[ \int_{\Omega} \varphi'_j(x) \varphi'_i(x) \, dx \right]_{ij}, \quad b = \left[ \int_{\Omega} f(x) \varphi_i(x) \, dx \right]_i$$

and the discrete solution is represented by the solution vector $u = [u_j] \in \mathbb{R}^{N_{V_h}}$. Computing contributions $A_{ij}$ and $b_i$ is possible as soon as the basis $\{\varphi\}_j$ is constructed explicitly.
Remark 3.1.1. The choice of indices $i$ and $j$ in the previous expressions follows the usual convention for row and column indices. The matrix $A$ represents a linear application from the solution space $H$ to the trial space $V$: therefore the solution space is the column space, while the trial space is the row space. In the case of Galerkin approximations where $H = V$ – and even more when the bilinear form is symmetric – it is tempting to choose the indices arbitrarily, but for the sake of consistency following the convention is recommended.

### 3.1.3 Construction of the discrete space

A discretization of the computational domain $ar{\Omega} = [0, 1]$ is constructed by partitioning the interval into disjoints subintervals $K_i = [x_i, x_{i+1}]$, $1 \leq i \leq N_K$.

$$\bar{\Omega} = \bigcup_{i=1}^{N_K} K_i, \quad K_i \cap K_j = \emptyset$$

The family of cells $\{K_i\}$ defines a mesh noted $\mathcal{T}_h$. A partition of $[0, 1]$ into $N_K$ subintervals $[x_i, x_{i+1}]$ corresponds to a family of $N_K + 1$ points $\{x_i\}$ which are the vertices of the mesh $\mathcal{T}_h$.

In this example the approximation space is constructed with piecewise linear Lagrange polynomials. The chosen discrete space is

$$V_h = \{ v \in C^0(\bar{\Omega}) \cap H^1_0(\Omega) : \forall K \in \mathcal{T}_h, v|_K \in P_1(K) \} \quad (3.1)$$

which consists of functions continuous over $\bar{\Omega}$ that are linear on each cell $K$, and $V_h \subset V$ so that the approximation is $H^1_0$–conformal. A continuous piecewise linear function $v_h \in V_h$ is the linear interpolate of $v \in V$ on $\mathcal{T}_h$ if $v_h = \mathcal{I}V_h$ with the interpolation operator

$$\mathcal{I}V_h : V \rightarrow V_h$$

$$v \mapsto \sum_{i=1}^{N_{V_h}} v(\xi_i) \varphi_i \quad (3.2)$$

such that $N_{V_h} = \text{dim} V_h$ and $\{\xi_i\}$ is a family of $N_{V_h}$ distinct nodes. In this particular case of a linear interpolation $N_{V_h} = N_K + 1$ and the family $\{\xi_i\}$ is identified with the vertices $\{x_i\}$.

The construction of the basis $\{\varphi_j\}$ of $V_h$ should produce a linear system that can be solved easily so the matrix should be as sparse as possible. Given the expression of the contributions $A_{ij}$ the requirement is that functions $\varphi_j$ overlap as little as possible with each other: the support of $\varphi_j$ should be reduced so that contributions are non-zero only for neighbouring $\varphi_i$ functions; this choice is consistent with the locality of differential operators.

For any $x_i$ the shape function is defined as

$$\varphi_i(x) = \begin{cases} 
\frac{x_{i+1} - x}{x_{i+1} - x_i} = 1 - \frac{x - x_i}{x_{i+1} - x_i}, & x_{i-1} \leq x \leq x_i \\
\frac{x - x_i}{x_{i+1} - x_i}, & x_i \leq x \leq x_{i+1} \\
0 & \text{otherwise}
\end{cases}$$
3.1. A PRELIMINARY EXAMPLE IN ONE DIMENSION OF SPACE

Figure 3.1: The function $v(x) = 0.8 \sin(2\pi x) - 0.32 \sin(4\pi x)$ and its linear interpolate $\mathcal{I}_{V_h} v$ with 6 equidistributed nodes.

Figure 3.2: Shape functions for Lagrange $P_1$ on a unit interval discretized into subintervals $K_i = [x_i, x_{i+1}]$.

so that the support of constructed functions $\varphi_i$ overlaps only with $\varphi_{i-1}$ and $\varphi_{i+1}$, and the expression of all the functions can be obtained from one another by an affine transformation; in the case of a uniform grid, $|x_{i+1} - x_i| = h$, shape functions are obtained simply by translation. Therefore it would be convenient to define shape functions on a reference interval, then apply an affine transformation to generate any $\varphi_i$. Shape functions are considered on a reference cell $\hat{K}$ which is the unit interval, then transported to any subinterval $K_i$ using the geometric transformation from $\hat{K}$ to $K$.

In the reference cell $\hat{K} = [0, 1]$ shape functions depicted in Figure 3.1.3 have expressions

\[
\begin{aligned}
\hat{\varphi}_0(\hat{x}) &= 1 - \hat{x} \\
\hat{\varphi}_1(\hat{x}) &= \hat{x}
\end{aligned}
\]
CHAPTER 3. FINITE ELEMENT SPACES

and the affine mapping $T_K$ mapping $\hat{K}$ to $K$ in one dimension is given by $T_K: \hat{x} \mapsto b + a\hat{x}$, for some $a, b \in \mathbb{R}$. The affine mapping to any subinterval $K_i$ is given by the change of coordinates

$$T_K: \hat{x} \in \hat{K} \mapsto x = x_i + (x_{i+1} - x_i)\hat{x} \in K_i$$

which is consistent with the definition of the global shape functions and reference shape functions, since

$$\begin{align*}
\hat{\varphi}_0(\hat{x}) &= 1 - \hat{x} \\
\varphi_i(x) &= \frac{x_{i+1} - x}{x_{i+1} - x_i} = 1 - \frac{x - x_i}{x_{i+1} - x_i}
\end{align*}$$

$$\begin{align*}
\hat{\varphi}_1(\hat{x}) &= \hat{x} \\
\varphi_{i+1}(x) &= \frac{x - x_i}{x_{i+1} - x_i}
\end{align*}$$

so that $\hat{x} = (x - x_i)(x_{i+1} - x_i)^{-1}$ as expected.

**Remark 3.1.2** (Link to higher dimensions of space). The affine mapping in $\mathbb{R}^d$ is given a relation of the type $T_K: \hat{x} \mapsto b_K + B_K\hat{x}$, with $b_K \in \mathbb{R}^d$ and $B_K \in \mathbb{R}^{d \times d}$. The matrix $B_K$ is the Jacobian of $T_K$, noted $J_{T_K}$. In one dimension of space $J_{T_K}$ contains only one entry $\partial T_K / \partial \hat{x} = (x_{i+1} - x_i) = |K|$, and its inverse $J_{T_K}^{-1}$ is obviously $(x_{i+1} - x_i)^{-1} = |K|^{-1}$.

3.1.4 Transport of Finite Element contributions

In the case of linear Lagrange elements the transport of Finite Element contributions corresponds to the change of coordinates $T_K$. The change of variable $x = T_K(\hat{x})$ is sketched for the mass matrix and the stiffness matrix, given a cell $K = [x_{i+1} - x_i]$.

Firstly, let us consider contributions to the mass matrix $M$,

$$M_{ij} = \int_K \varphi_j(x)\varphi_i(x)\,dx$$
which can be rewritten, given that \( x = T_K(\hat{x}) \) and \( \mathrm{d}x = (x_{i+1} - x_i) \, \mathrm{d}\hat{x} \),

\[
M_{ij} = \int_K \varphi_j(T_K(\hat{x})) \varphi_i(T_K(\hat{x}))(x_{i+1} - x_i) \, \mathrm{d}\hat{x}
\]

Since \( \varphi_i \circ T_K = \hat{\varphi}_i \) and \( (x_{i+1} - x_i) = |K| \), then

\[
M_{ij} = |K| \int_K \hat{\varphi}_j(\hat{x}) \hat{\varphi}_i(\hat{x}) \, \mathrm{d}\hat{x}
\]

so that contributions for \( K \in T_h \) to the mass matrix can obtained by a scaling of the contribution on \( \hat{K} \).

Secondly, let us consider contributions to the stiffness matrix \( K \),

\[
K_{ij} = \int_K \partial_x \varphi_j(x) \partial_x \varphi_i(x) \, \mathrm{d}x
\]

which can be rewritten using the chain rule for the derivation of \( \varphi_i = \hat{\varphi}_i \circ T_K^{-1} \), formally \( \partial_x \varphi_i = \partial_{\hat{x}} \hat{\varphi}_i \partial_x T_K^{-1} \) with \( \partial_x T_K^{-1} = |K|^{-1} \)

\[
K_{ij} = \int_K \frac{1}{|K|} \partial_{\hat{x}} \hat{\varphi}_j(\hat{x}) \frac{1}{|K|} \partial_{\hat{x}} \hat{\varphi}_i(\hat{x}) |K| \, \mathrm{d}\hat{x}
\]

therefore,

\[
K_{ij} = \frac{1}{|K|} \int_K \hat{\varphi}_j(\hat{x}) \hat{\varphi}_i(\hat{x}) \, \mathrm{d}\hat{x}
\]

so that contributions for \( K \in T_h \) to the stiffness matrix can also obtained by a scaling of the contribution on \( \hat{K} \).

### 3.1.5 Generalization of the methodology

The example of linear Lagrange in one dimension provides a good overview of the methodology for constructing discrete spaces using Finite Elements. The steps can be summarized as:

1. Choose a function space suggested by the weak formulation.
2. Discretize the computational domain into a mesh.
3. Choose a polynomial approximation space.
4. Define an interpolation operator.
5. Provide a definition of a reference element.
6. Construct a mapping to generate the discrete space.

The generalization of this approach can take different directions:

- Extending to multiple dimensions in space.
- Increasing the polynomial order.
- Using a different approximation space.
- Controlling the solution in other means than pointwise values.
3.2 Admissible mesh

**Definition 3.2.1 (Mesh).** Let $\Omega$ be polygonal ($d = 2$) or polyhedral ($d = 3$) subset of $\mathbb{R}^d$, we define $\mathcal{M}$ (a triangulation $T_h$ in the simplicial case) as a finite family $\{K_i\}$ of disjoint convex non-empty subsets of $\Omega$ named cells. Moreover $\mathcal{V}(T_h) = \{v_i\}$ denotes the set a vertices of $\mathcal{M}$, $\mathcal{E}(\mathcal{M}) = \{\sigma_{KL} = K \cap L\}$ denotes the set of facets, which are edges ($d = 2$) or faces ($d = 3$).

**Definition 3.2.2 (Mesh size).**

\[ h_T = \max_{K \in T_h} (\text{diam}(K)) \]

with diam$(K)$ with diameter of the cell, *i.e.* the maximum distance between two points of $K$.

**Definition 3.2.3 (Geometrically conforming mesh).** A mesh is said geometrically conforming if two neighbouring cells share either exactly one vertex, exactly one edge in the case $d = 2$, or in the case $d = 3$ exactly one face.

The meaning of the previous condition is that there should not be any “hanging node” on a facet. Moreover some theoretical results require that the mesh satisfies some regularity condition: for example, bounded ratio of equivalent ball diameter, Delaunay condition on the angles of a triangle, . . .

3.3 Definition of a Finite Element

In the frame of the Galerkin method, given a function space $V$, a discrete space $V_h = \text{span}\{\phi_j\} \subset V$ was introduced, this section describes how to build such space by constructing an abstraction called *Finite Element*.

**Definition 3.3.1 (Finite Element – [4] page 19, [2] page 69).** A Finite Element consists of a triple $(K, \mathcal{P}, \Sigma)$, such that

- $K$ is a compact, connected subset of $\mathbb{R}^d$ with non-empty interior and with regular boundary (typically Lipshitz continuous),
- $\mathcal{P}$ is a finite dimensional vector space of functions $p : K \to \mathbb{R}$, which is the space of shape functions; $\text{dim}(\mathcal{P}) = N_\mathcal{P}$,
- $\Sigma$ is a set $\{\sigma\}_j$ of linear forms,

\[ \sigma_j : \mathcal{P} \to \mathbb{R}, \quad \forall j \in [1, N_\mathcal{P}] \]

\[ p \mapsto p_j = \sigma_j(p) \]

which is a basis of $\mathcal{L}(\mathcal{P}, \mathbb{R})$, the dual of $\mathcal{P}$. 


3.3. DEFINITION OF A FINITE ELEMENT

Practically, the definition requires first to consider the Finite Element on a cell $K$ which can be an interval ($d = 1$), a polygon ($d = 2$) or a polyhedron ($d = 3$) (Example: triangle, quadrangle, tetrahedron, hexahedron), then an approximation space $\mathcal{P}$ (Example: polynomial space) and the local degrees of freedom $\Sigma$ are chosen (Example: value at $N_p$ geometrical nodes $\{\xi_i\}$, $\sigma_i(\varphi_j) = \varphi_j(\xi_i)$). The local shape functions $\{\varphi_i\}$ are then constructed so as to ensure unisolvence. The dimension of $\mathcal{P}$ is usually called dimension of the Finite Element.

**Proposition 3.3.2** (Determination of the local shape functions). Let $\{\sigma_i\}_{1\leq i \leq N_p}$ be the set of local degrees of freedoms, the local shape functions are defined as $\{\varphi_i\}_{1\leq i \leq N_p}$ a basis of $\mathcal{P}$ such that,

$$\sigma_i(\varphi_j) = \delta_{ij}, \; \forall \, i, j \in [1, N_p]$$

**Definition 3.3.3** (Unisolvency). A Finite Element is said unisolvent if for any vector $(\alpha_1, \cdots, \alpha_{N_p}) \in \mathbb{R}^N$ there exists a unique representant $p \in \mathcal{P}$ such that $\sigma_i(p) = \alpha_i, \; \forall \, i \in [1, N_p]$.

The unisolvency property of a Finite Element is equivalent to construct $\Sigma$ as dual basis of $\mathcal{P}$, thus any function $p \in \mathcal{P}$ can be expressed as

$$p = \sum_{j=1}^{N} \sigma_j(p) \varphi_j$$

the unique decomposition on $\{\varphi_j\}$, with $p_j = \sigma_j(p)$ the $j$-th degree of freedom. In other words, the choice of $\Sigma = \{\sigma_j\}$ ensures that the vector of degree of freedoms $(p_1, \cdots, p_{N_p})$ uniquely represents a function of $\mathcal{P}$. Defining $\Sigma$ as dual basis of $\mathcal{P}$ is equivalent to:

$$\dim(\mathcal{P}) = \text{card}(\Sigma) = N_p$$

$$(3.3a)$$

$$\forall \, p \in \mathcal{P}, \; (\sigma_i(p) = 0, 1 \leq i \leq N) \Rightarrow (p = 0)$$

$$(3.3b)$$

in which Property $(3.3a)$ ensures that $\Sigma$ generates $\mathcal{L}(\mathcal{P}, \mathbb{R})$ and Property $(3.3b)$ that $\{\sigma_i\}$ are linearly independent. Usually the unisolvency is part of the definition of a Finite Element since chosing the shape functions such that $\sigma_i(\varphi_j) = \delta_{ij}$ is equivalent.

As a first step, the Galerkin decomposition of functions $u_h \in \mathcal{V}_h$ was introduced given a basis of $\mathcal{V}_h$ and degrees of freedom $u_j$, but without giving a proper definition of the latter; we just assumed that they are real coefficients. In a second step, the Lagrange interpolation operator of Definition 3.2 was introduced to interpolate functions in $V$ as continuous piecewise linear functions in $\mathcal{V}_h$. Degrees of freedom $u_j$ were then defined as $u(\xi_j)$ the value of the function $u$ at the Lagrange node $\xi_j$. Now that a definition a Finite Element was stated, in a similar fashion a global interpolation operator can be defined for general Finite Elements, i.e. with degrees of freedom defined by linear forms $\sigma_j$. 

Definition 3.3.4 (Global interpolation operator).

\[ \mathcal{I}_{V_h} : V \rightarrow V_h \]

\[ v \mapsto \sum_{j=1}^{N_{V_h}} \sigma_j(v) \varphi_j \]

Given that the Finite Element is defined on a cell \( K \), the corresponding local interpolation operator can be introduced.


\[ \mathcal{I}_{K,V} : V(K) \rightarrow \mathcal{P} \]

\[ v \mapsto \sum_{j=1}^{N_P} \sigma_j(v) \varphi_j \]

Remark 3.3.6. The notation using the dual basis can be confusing but with the relation \( \sigma_i(p) = p(\xi_i) \) in the nodal Finite Element case it is easier to understand that the set \( \Sigma \) of linear forms defines how the interpolated function \( \mathcal{I}_{K,V} u \) represents its infinite dimensional counterpart \( u \) through the definition of the degrees of freedom. In the introduction, we defined simply \( u_i = \sigma_i(u) \) without expliciting it. A natural choice is the pointwise representation \( u_i = u(\xi_i) \) at geometrical nodes \( \{\xi_i\} \), which is the case of Lagrange elements, but it is not the only possible choice! For example, \( \sigma_i \) can be:

- a mean flux through each facet of the element (Raviart–Thomas)
  \[ \sigma_i(v) = \int_{\xi_i} v \cdot n_{\xi} \, ds \]

- a mean value over each facet of the element (Crouzeix–Raviart)
  \[ \sigma_i(v) = \int_{\xi_i} v \, ds \]

- a mean value of the tangential component over each facet of the element (Nédelec)
  \[ \sigma_i(v) = \int_{\xi_i} v \cdot \tau_{\xi} \, ds \]

A specific choice of linear form allows a control on a certain quantity: divergence for the first two examples, and curl for the third. The approximations will then not only be \( H^1 \)-conformal but also include the divergence or the curl in the space.

Remark 3.3.7. The Finite Element approximation is said \( H \)-conformal if \( V_h \subset H \) and is said non-conformal is \( V_h \not\subset H \). In this latter case the approximate problem can be constructed by building an approximate bilinear form

\[ a_h(\cdot, \cdot) = a(\cdot, \cdot) + s(\cdot, \cdot) \]

as described, for instance, in the case of stabilized methods for advection-dominated problems in Section ??.
3.4 Transport of the Finite Element

In practice to avoid the construction of shape functions for any Finite Element \((K, \mathcal{P}, \Sigma)\), \(K \in \mathcal{T}_h\), the local shape functions are evaluated for a reference Finite Element \((\hat{K}, \hat{\mathcal{P}}, \hat{\Sigma})\) defined on a reference cell \(\hat{K}\) and then transported onto any cell \(K\) of the mesh. For example, in the case of simplicial meshes the reference cell in one dimension is the unit interval \([0,1]\), in two dimensions the unit triangle with vertices \(\{(0,0), (1,0), (0,1)\}\), and in three dimensions the unit tetrahedron with vertices \(\{(0,0,0), (1,0,0), (0,1,0), (0,0,1)\}\).

Using this approach any Finite Element \((K, \mathcal{P}, \Sigma)\) on the mesh can be generated from \((\hat{K}, \hat{\mathcal{P}}, \hat{\Sigma})\) provided that a mapping can be constructed such that \((K, \mathcal{P}, \Sigma)\) and \((\hat{K}, \hat{\mathcal{P}}, \hat{\Sigma})\) possess equivalent properties. In particular, an important property is the unisolvence: if a Finite Element is equivalent to another Finite Element which is unisolvent, then it is also unisolvent.

**Definition 3.4.1 (Affine-equivalent Finite Elements).** Two Finite Elements \((K, \mathcal{P}, \Sigma)\) and \((\hat{K}, \hat{\mathcal{P}}, \hat{\Sigma})\) are said affine-equivalent if there exists a bijection \(T_K\) from \(\hat{K}\) onto \(K\) such that:

\[
\forall p \in \mathcal{P}, \ p \circ T_K \in \hat{\mathcal{P}}
\]

and

\[
\Sigma = T_K(\hat{\Sigma})
\]

By collecting the local shape functions and local degrees of freedom from all the generated \((K, \mathcal{P}, \Sigma)\) on the mesh, we then construct global shape functions and global degrees of freedom and thus the approximation space \(V_h\).

For Lagrange elements the transformation used to transport the Finite Element on the mesh is the affine mapping \(T_K\) but this is not suitable in general. An auxiliary mapping is needed to transfer the approximation space on \(\hat{K}\) to the approximation space on \(K\). The following definition extends the affine-equivalence to a general equivalence property between Finite Elements.

**Definition 3.4.2 (Equivalent Finite Elements).** Two Finite Elements \((K, \mathcal{P}, \Sigma)\) and \((\hat{K}, \hat{\mathcal{P}}, \hat{\Sigma})\) are said equivalent if there exists a bicontinuous bijection \(\Phi_K\) from \(V(K)\) onto \(V(\hat{K})\) such that \((K, \mathcal{P}, \Sigma)\) is generated from \((\hat{K}, \hat{\mathcal{P}}, \hat{\Sigma})\):

\[
\begin{align*}
K &= T_K(\hat{K}) \\
\mathcal{P} &= \{\Phi_K^{-1}(\hat{p}), \forall \hat{p} \in \hat{\mathcal{P}}\} \\
\Sigma &= \{\sigma_{K,j} : \sigma_{K,j}(\hat{p}) = \hat{\sigma}_{K,j}(\Phi_K^{-1}(\hat{p})), \forall \hat{p} \in \hat{\mathcal{P}}\}
\end{align*}
\]

Given that any Finite Element \((K, \mathcal{P}, \Sigma)\) can be generated from a reference Finite Element \((\hat{K}, \hat{\mathcal{P}}, \hat{\Sigma})\), then global interpolation properties for the entire space \(V_h\) (spanned by collecting all the shape functions) can be inferred from local interpolation properties on \(\hat{K}\). More precisely, without going into the details, the following result is specified in [4]:
which means that the interpolation operators and the transport of the elements commute: \( \mathcal{I}_{K,V} \circ \Phi_K = \Phi_K \circ \mathcal{I}_{K,V} \).

**Remark 3.4.3.** Note that in the literature the mapping between spaces is defined from \( V(K) \) to \( V(\hat{K}) \), while the geometric mapping between cells is defined from \( \hat{K} \) to \( K \). The affine-equivalence consists in the case \( \Phi_K \) coinciding with \( T_K^{-1} \). For Lagrange we defined \( \Phi_K : C^0(K) \rightarrow C^0(\hat{K}) \) given by \( v \mapsto v \circ T_K \). In that case \( \mathcal{P} = \text{span} \{ \hat{\phi} \circ T_K^{-1} \} \) and the degrees of freedom are defined by \( \Sigma = \{ \sigma_i : \sigma_i(v) = \hat{\sigma}_i(\Phi_K(v)) = \Phi_K(v)(\hat{\xi}_i) = v \circ T_K(\hat{\xi}_i) \} \). If \( \xi_i = T_K(\hat{\xi}_i) \) then the definition is consistent.

### 3.5 Method

**Algorithm 3.5.1** (Finite Element Method). Solving a problem by a Finite Element Method is defined by the following procedure:

1. Choose a reference Finite Element \((\hat{K}, \hat{P}, \hat{\Sigma})\).
2. Construct an admissible mesh \( T_h \) such that any cell \( K \in T_h \) is in bijection with the reference cell \( \hat{K} \).
3. Define a mapping to transport the reference Finite Element defined on \( \hat{K} \) onto any \( K \in T_h \) to generate \((K, P, \Sigma)\).
4. Construct a basis for \( V_h \) by collecting all the shape functions of Finite Elements \( \{(K, P, \Sigma)\}_{K \in T_h} \) sharing the same degree of freedom.