Galerkin Approximations and Finite Element Methods

Ricardo G. Durán[1]

[1]Departamento de Matemática, Facultad de Ciencias Exactas, Universidad de Buenos Aires, 1428 Buenos Aires, Argentina.

Chapter 1

Galerkin Approximations

1.1 A simple example

In this section we introduce the idea of Galerkin approximations by considering a simple 1-d boundary value problem. Let u be the solution of

$$\begin{cases} -u'' + u = f & \text{in } (0,1) \\ u(0) = u(1) = 0 \end{cases} \tag{1.1}$$

and suppose that we want to find a computable approximation to u (of course, it is not very interesting to solve this problem approximately but the ideas we are going to introduce are quite general and can be applied in many situations as we are going to see later on).

Multiplying equation (1.1) by a test function and integrating by parts we obtain the weak formulation of (1.1)

$$\int_0^1 (u'v' + uv)\,dx = \int_0^1 fv\,dx \qquad \forall v \in H_0^1(0,1) \tag{1.2}$$

where $H_0^1(0,1)$ is the Sobolev space

$$H_0^1(0,1) = \{v \in L^2(0,1) : v' \in L^2(0,1) \ \text{ and}, v(0) = v(1) = 0\}$$

If u is regular (for example with two continuous derivatives) then problems (1.1) and (1.2) are equivalent. We can use (1.2) in order to define an approximation to u. We are going to construct polygonal approximations to u. With this purpose let us introduce a uniform partition of the domain

$(0,1)$ into $N+1$ subintervals (x_j, x_{j+1}) with

$$x_j = \frac{j}{N+1} \text{for} \quad j = 0, \ldots, N+1$$

and consider the space V_N of polygonal functions vanishing at the boundary of $(0,1)$, i.e.,

$$V_N = \{v \in C^0 : v|_{(x_j, x_{j+1})} \text{ is linear and } \ v(0) = v(1) = 0\}$$

where C^0 denotes the space of continuous functions.

Observe that, $\forall N$, V_N is a subspace of $H_0^1(0,1)$ and that V_N has finite dimension. Indeed, a polygonal function $v \in V_N$ is uniquely determined by its values at the finite number of points $x_1, \ldots, x_N$.

We define the Galerkin approximation $u_N \in V_N$ to u by imposing (1.2) but only for functions $v \in V_N$, i.e., $u_N \in V_N$ is such that:

$$\int_0^1 (u_N' v' + u_N v)\, dx = \int_0^1 f v\, dx \qquad \forall v \in V_N \tag{1.3}$$

We are going to see that there is a unique u_N satisfying (1.3) and moreover, since V_N is finite dimensional, that it can be computed by solving a linear system of equations. Indeed, given a basis ϕ_j of V_N, for example, the usual Lagrange basis defined by $\phi_j(x_i) = \delta_{ij}$ for $i, j = 1, \ldots, N$, u_N can be written as,

$$u_N = \sum_{j=1}^{N} U_j \phi_j \quad , U_j \in \mathbb{R} \tag{1.4}$$

Note that with this choice of basis we have $U_j = u_N(x_j)$. Now, since any $v \in V_N$ is a linear combination of the ϕ_j it is easy to see that (1.3) is equivalent to

$$\int_0^1 (u_N' \phi_k' + u_N \phi_k)\, dx = \int_0^1 f \phi_k\, dx \quad \text{for} \quad k = 1, \ldots, N \tag{1.5}$$

and using (1.4) we have

$$\sum_{j=1}^{N} U_j \int_0^1 (\phi_j' \phi_k' + \phi_j \phi_k)\, dx = \int_0^1 f \phi_k\, dx \quad \text{for} \quad k = 1, \ldots, N$$

Therefore, we can find $U = (U_j) \in \mathbb{R}^N$ (and then u_N) by solving the linear system of equations

$$AU = F$$

where $A = (a_{kj}) \in \mathbb{R}^{N\times N}$ with $a_{kj} = \int_0^1 (\phi_j'\phi_k' + \phi_j\phi_k)\, dx$ and $F \in \mathbb{R}^N$ with $F_k = \int_0^1 f\phi_k\, dx$.

An easy computation shows that A is the tridiagonal symmetric matrix such that

$$a_{jj} = \frac{2}{h} + \frac{2}{3}h \quad \text{and} \quad a_{jj-1} = a_{jj+1} = -\frac{1}{h} + \frac{h}{6}$$

Therefore, the system of equations to be solved is

$$\frac{-U_{j-1} + 2U_j - U_{j+1}}{h} + \frac{h}{6}U_{j-1} + \frac{2h}{3}U_j + \frac{h}{6}U_{j+1} = F_j \quad \text{for} \quad j = 1, \ldots, N$$

where we define $U_0 = U_{N+1} = 0$.

In particular the matrix A is invertible and moreover, it is positive definite (a property that is inherited from the coercivity of the bilinear form associated with the differential equation). Consequently, there is a unique solution U and therefore the Galerkin approximation u_N is well defined.

Note that dividing by h we obtain a finite difference scheme for problem (1.1), i.e.,

$$\frac{-U_{j-1} + 2U_j - U_{j+1}}{h^2} + \frac{1}{6}U_{j-1} + \frac{2}{3}U_j + \frac{1}{6}U_{j+1} = \frac{1}{h}F_j \quad \text{for} \quad j = 1, \ldots, N$$

where $u''(x_j)$ is approximated by a standard centered difference scheme and, $u(x_j)$ and $f(x_j)$ are replaced by averages. Therefore, in this particular case, the Galerkin approximation is related with a known finite difference approximation.

For any N we have defined the Galerkin approximation $u_N \in V_N$ to u and one would expect that u_N will converge to u when $N \to \infty$ because any continuous function can be approximated by polygonals with an increasing number of nodes. In other words, one would expect that the Galerkin approximations converge to u whenever the family of spaces V_N approximates u in the following sense:

$$d(u, V_N) = \inf_{v\in V_N} d(u, v) \to 0 \quad \text{when} \quad N \to \infty$$

where $d(u, v) = \|u - v\|$ is the distance measured in some appropriate norm. In the next section we are going to see that this is true in a general context.

1.2 The general case

In this section we define and analyze the convergence of Galerkin approximations of a general problem given by a bilinear form in a Hilbert space. Let V be a Hilbert space and let $a(\,.\,,\,.\,)$ and L be continuous bilinear and linear forms respectively defined on V. We want to find a computable approximation to the solution $u \in V$ of the problem

$$a(u,v) = \langle L, v\rangle \quad \forall v \in V \tag{1.6}$$

where $\langle\,.\,,\,.\,\rangle$ denotes the duality product between V' and V. Below we will recall general conditions on the form a which ensure the existence of a unique solution u, which in particular, applies to the very important class of the coercive forms.

Definition 1.2.1 *We say that a is coercive on V if there exists a constant $\alpha > 0$ such that*

$$a(u,u) \geq \alpha \|u\|_V^2 \quad \forall u \in V \tag{1.7}$$

Examples of problems like (1.6) are given by the variational formulation of differential equations.

Example 1.2.1 *Scalar linear elliptic equations of second order.*

$$\begin{cases} -\sum_{i,j=1}^n \frac{\partial}{\partial x_i}(a_{ij}\frac{\partial u}{\partial x_j}) = f & in \;\; \Omega \subset \mathbb{R}^n \\ u = 0 & on \;\; \partial\Omega \end{cases}$$

where the coefficients $a_{ij} = a_{ij}(x)$ are bounded functions and there exist $\gamma > 0$ such that

$$\gamma|\xi|^2 \leq \sum_{i,j=1}^n a_{ij}\xi_i\xi_j \quad \forall x \in \Omega \quad \forall \xi \in \mathbb{R}^n \tag{1.8}$$

This problem can be written as (1.6) with

$$V = H_0^1(\Omega) = \{v \in L^2(\Omega) : \frac{\partial v}{\partial x_j} \in L^2(\Omega) \;\; for \;\; j = 1, \dots, n \;\; and \;\; , v = 0 \;\; on \;\; \partial\Omega\}$$

which is a Hilbert space with the norm

$$\|v\|_{H^1} = \|v\|_{L^2} + \|\nabla v\|_{L^2} \quad ,$$

and a and L defined by

$$a(u,v) = \sum_{i,j=1}^{n} \int_{\Omega} a_{i,j} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} \, dx$$

and

$$\langle L, v \rangle = \int_{\Omega} f v \, dx$$

By using the ellipticity condition (1.8), the boundedness of the coefficients and the Poincaré inequality (see for example [8]) it can be seen that the form a is coercive and continuous. The linear form L is continuous if we assume, for example, that $f \in L^2$.

Example 1.2.2 *The linear elasticity equations.*

If we consider, for simplicity, homogeneous Dirichlet conditions, the equations are

$$\begin{cases} -\mu \Delta \mathbf{u} - (\lambda + \mu) \nabla div \mathbf{u} = \mathbf{f} & in \;\; \Omega \subset \mathbb{R}^3 \\ \mathbf{u} = 0 & on \;\; \partial\Omega \end{cases}$$

where μ and λ are positive constants (the Lamé elasticity parameters). Now the unknown $\mathbf{u}$ and the right hand side $\mathbf{f}$ are vector functions. The weak formulation of this problem can be written as (1.6) with $V = H_0^1(\Omega)^3$ and,

$$a(\mathbf{u}, \mathbf{v}) = \int_{\Omega} \{2\mu \varepsilon_{i,j}(\mathbf{u}) \varepsilon_{i,j}(\mathbf{v}) + \lambda div \mathbf{u} div \mathbf{v}\} \, dx$$

where

$$\varepsilon_{i,j}(\mathbf{v}) = \frac{1}{2}(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i})$$

In this case, it can be seen that the bilinear form a is coercive by using the Korn's inequality (see for example [15])

The continuity and coercivity of the form imply the existence of a unique solution of (1.6) (this result is known as Lax-Milgram theorem, see [8, 34]). As we are going to see, these conditions also imply the convergence of Galerkin approximations (of course, provided that they are defined on "good" approximation spaces). However, there are important examples (such as the Stokes equations) in which the associated bilinear form is not coercive but it satisfies a weaker condition known as “the inf-sup condition”. This condition also ensures the existence of a unique solution of (1.6), and in fact it is also necessary (actually, if the form is not symmetric it has to satisfy two inf-sup conditions). We will recall this fundamental theorem below and in the next section we will analyze the convergence of Galerkin approximations for this kind of bilinear forms.

Definition 1.2.2 *We say that the bilinear form a satisfies the inf-sup conditions on V if there exists $\alpha > 0$ such that*

$$\sup_{v\in V} \frac{a(u,v)}{\|v\|_V} \geq \alpha \|u\|_V \quad \forall u \in V \tag{1.9}$$

and

$$\sup_{u\in V} \frac{a(u,v)}{\|u\|_V} \geq \alpha \|v\|_V \quad \forall v \in V \tag{1.10}$$

Remark 1.2.1 *Clearly, if a is symmetric both conditions are the same.*

Remark 1.2.2 *Note that condition (1.9) (and analogously (1.10)) can be written as*

$$\inf_{u\in V} \sup_{v\in V} \frac{a(u,v)}{\|u\|_V \|v\|_V} > 0$$

which justifies the usual terminology.

Remark 1.2.3 *If a is coercive it satisfies the inf-sup conditions. In fact,*

$$\sup_{v\in V} \frac{a(u,v)}{\|v\|_V} \geq \frac{a(u,u)}{\|u\|_V} \geq \alpha \|u\|_V$$

Remark 1.2.4 *The inf-sup condition can be written in terms of the linear operators A and its adjoint A^* associated with a,*

$$A : V \to V' \quad \text{and} \quad A^* : V \to V'$$

defined by

$$\langle Au, v\rangle_{V'\times V} = a(u, v) \quad \text{and} \quad \langle u, A^* v\rangle_{V\times V'} = a(u, v)$$

In fact, (1.9) and (1.10) are equivalent to

$$\|Au\|_{V'} \geq \alpha \|u\|_V \qquad \forall u \in V \tag{1.11}$$

and

$$\|A^* v\|_{V'} \geq \alpha \|v\|_V \qquad \forall v \in V \tag{1.12}$$

Remark 1.2.5 *For example, when $V = \mathbb{R}^n$ the coercivity of a means that the associated matrix A is positive definite while the inf-sup condition means that A is invertible.*

In the next theorem we will use the following well known result of functional analysis (see [8, 34]). For $W \subset V$ we define $W^0 \subset V'$ by

$$W^0 = \{L \in V' : \langle L\,,\, v\rangle = 0, \quad \forall v \in W\}$$

then,

$$(KerA)^0 = \overline{ImA^*} \tag{1.13}$$

and

$$(KerA^*)^0 = \overline{ImA} \tag{1.14}$$

Theorem 1.2.1 *The continuous bilinear form a satisfies the inf-sup conditions (1.9) and (1.10) if and only if the operator A is bijective (i.e., problem (1.6) has a unique solution for any L and therefore, A has a continuous inverse,i.e., $\|u\|_V \leq C\|L\|_{V'}$).*

Proof. Assume first that a satisfies the inf-sup conditions. It follows from (1.11) that A is injective and from (1.12) that A^* is injective. So, in view of (1.14) the proof concludes if we show that ImA is closed. Suppose that $Au_n \to w$ then, it follows from (1.11) that

$$\|A(u_n - u_m)\|_{V'} \geq \alpha \|u_n - u_m\|_V$$

and therefore $\{u_n\}$ is a Cauchy sequence and so convergent to some $u \in V$ and, by continuity of A, $w = Au \in ImA$.

Conversely, if A is bijective, then A^* is bijective too and therefore both have a continuous inverse (see [8, 34]) and so (1.9) and (1.10) hold. □

Now we introduce the Galerkin approximations to the solution of problem (1.6). Assume that we have a family V_N of finite dimensional subspaces of V. Then, the Galerkin approximation $u_N \in V_N$ is defined by

$$a(u_N, v) = \langle L, v \rangle \qquad \forall\, v \in V_N \tag{1.15}$$

In order to have u_N well defined we need to ask some condition on the form a. From the Theorem above we know that u_N satisfying (1.15) exists and is unique if and only if a satisfies the inf-sup conditions on V_N. In particular, the Galerkin approximations are well defined for coercive forms. At this point, it is important to remark a fundamental difference between coercive forms on V and forms which satisfy the inf-sup on V but are not coercive:

If a is coercive on V, then, it is also coercive on any subspace, and in particular on V_N and the Galerkin approximation u_N is well defined. Instead, the inf-sup condition on V is not inherited by subspaces, and so, when the form is not coercive, the inf-sup (or something equivalent!) has to be verified on V_N in order to have u_N well defined. We will come back to this point when we analyze the convergence of Galerkin approximations.

1.3 Convergence for the case of coercive forms

Assume now that the form a is continuous and coercive. We will call M the continuity constant, i.e.,

$$a(u, v) \leq M \|u\|_V \|v\|_V \qquad \forall u, v \in V \tag{1.16}$$

A natural question is whether $\lim_{N\to\infty} u_N = u$ provided the spaces V_N are chosen in an appropriate way. Clearly, if the Galerkin approximations converge to u we have that

$$d(u, V_N) = \inf_{v \in V_N} \|u - v\|_V \to 0 \quad \text{when} \quad N \to \infty \tag{1.17}$$

therefore, (1.17) is a natural property to ask on the subspaces (it means that they approximate u), and we would like to know if it is also a sufficient condition for convergence. The answer is yes and it follows from the following Lemma (known as Cea's lemma).

Lemma 1.3.1 *If a is continuous and coercive then,*

$$\|u - u_N\|_V \leq \frac{M}{\alpha} \inf_{v \in V_N} \|u - v\|_V$$

Proof. Subtracting (1.15) from (1.6) we have the error equation

$$a(u - u_N, v) = 0 \qquad \forall v \in V_N \tag{1.18}$$

Now, using (1.18), (1.16) and (1.7) we have that for any $v \in V_N$

$$\alpha \|u - u_N\|^2 \leq a(u - u_N, u - u_N) = a(u - u_N, u - v) \leq M \|u - u_N\|_V \|u - v\|_V$$

and therefore

$$\|u - u_N\| \leq \frac{M}{\alpha} \|u - v\|_V \qquad \forall v \in V_N$$

and the lemma is proved. □

The lemma says that the Galerkin approximation u_N is like the best approximation in V_N to u up to a constant depending only on the form a (i.e., independent of the subspaces). In particular we have the following convergence result.

Theorem 1.3.2 *If a is continuous and coercive and the spaces V_N are such that (1.17) holds then* $\lim_{N \to \infty} u_N = u$

Remark 1.3.1 *In the particular case in which the form a is symmetric, it defines a scalar product on the space V which is equivalent to the original one and, (1.18) shows that the Galerkin approximation u_N is exactly the orthogonal projection of u onto V_N with the scalar product given by a. Therefore, it is the best approximation in the norm corresponding to that scalar product. In particular, it is easy to see that in this case, the constant $\frac{M}{\alpha}$ in the estimate of Lemma 1.3.1 can be replaced by $\sqrt{\frac{M}{\alpha}}$*

1.4 Convergence for forms satisfying the inf-sup condition

Suppose now that the form a is not coercive but it satisfies the inf-sup conditions (1.9) and (1.10) on V. Then, we know that problem (1.6) has a unique solution $u \in V$ and, as before, we are interested in the convergence of its Galerkin approximations. As we mentioned in Section 1.2, the inf-sup condition is not inherited by subspaces (note that the sup will be taken in a smaller set). Therefore, in order to have the Galerkin approximations well defined we have to assume (and in concrete cases it has to be proved!) that a satisfies the inf-sup condition also on V_N, i.e., that there exists $\beta > 0$ such that

$$\sup_{v \in V_N} \frac{a(u,v)}{\|v\|_V} \geq \beta \|u\|_V \quad \forall u \in V_N \tag{1.19}$$

Note that, since V_N is finite dimensional the second inf-sup condition follows from this one.

In order to prove convergence, we will also ask that the constant β be independent of N. Under this assumption we have the following generalization of Cea's lemma due to Babuska [2] and, as a consequence, a convergence result which generalizes Theorem 1.3.2 for this case.

Lemma 1.4.1 *If the form a is continuous and satisfies the inf-sup condition (1.19) then,*

$$\|u - u_N\|_V \leq \left(1 + \frac{M}{\beta}\right) \inf_{v \in V_N} \|u - v\|_V$$

in particular, if β is independent of N, the constant in this error estimate is independent of N.

Proof. Take $v \in V_N$. From (1.19) and the error equation (1.18) we have,

$$\beta \|v - u_N\|_V \leq \sup_{w \in V_N} \frac{a(v - u_N, w)}{\|w\|_V} = \sup_{w \in V_N} \frac{a(v - u, w)}{\|w\|_V} \leq M \|v - u\|_V$$

and the proof concludes by using the triangle inequality.□

As an immediate consequence we have the following convergence result,

Theorem 1.4.2 *If a is continuous and satisfies the inf-sup condition (1.19) with β independent of N, and the spaces V_N are such that (1.17) holds then,*

$$\lim_{N \to \infty} u_N = u$$

Remark 1.4.1 *Condition (1.19) is a "stability condition", indeed, it says that the solution is bounded by the right hand side, i.e.,* $\|u_N\|_V \leq \frac{1}{\beta}\|L\|_{V'}$ *and this estimate is valid uniformly in* N *if* β *is independent of* N*. Therefore, the Theorem above can be thought of as the finite element version of the classical Lax Theorem for Finite Differences which states that stability plus consistency implies convergence. In the case we are considering here the consistency follows from the fact that* V_N *is a subspace of* V*. It is possible to construct approximations on spaces* V_N *which are not contained in* V *and in that case, the consistency has to be verified. In the Finite Element context this kind of methods are called "non conforming" (we will not treat them here, we refer for example to [14]).*

Chapter 2

Finite element spaces, interpolation and error estimates

In this chapter we apply the results obtained above to the numerical solution of elliptic boundary value problems. Among the most important and widely used Galerkin approximations are those based on spaces of piecewise polynomial functions. Let $\Omega \in \mathbb{R}^n$ (with $n = 2$ or 3) be a polygonal (or polyhedral) domain and u be the solution of the elliptic equation of Example 1.2.1 of Section 1.2. As we have seen in that section, u is the solution of a problem like (1.6) with a coercive form a (Indeed, all what we are going to say applies to Example (1.2.2) (the elasticity equations)). Therefore, the convergence result of Theorem 1.3.2 applies to this problem and the question is how to construct good approximation subspaces (i.e., such that they satisfy (1.17)) V_N of $V = H_0^1(\Omega)$ (the space where the exact solution belongs). The Finite Element Method provides a systematic way of constructing this kind of subspaces. The domain Ω is divided into a finite number of subsets (or elements) in an appropriate way to be specified below and the approximation to u is such that restricted to each element it is a polynomial of a certain class. A simple example is the one given for 1-d problems in the first section. We are going to see some classical examples of finite element spaces in 2 dimensions (for extensions to 3-d we refer to [14]).

2.1 Triangular elements of order k

Assume that we have a triangulation $\mathcal{T} = \{T\}$ of $\Omega \in \mathbb{R}^2$, i.e., $\Omega = \cup_{T\in\mathcal{T}} T$. The triangulation is admissible if the intersection of two triangles is either empty, or a vertex, or a common side, and from now on, all the triangulations considered are assumed to be admissible. Given a natural number k we associate with $\mathcal{T}$ the space $V^k(\mathcal{T})$ of continuous piecewise polynomials of degree k, i.e.,

$$V^k(\mathcal{T}) = \{v \in C^0(\Omega) : v|_T \in \mathcal{P}_k \,, \forall T \in \mathcal{T}\}$$

where $\mathcal{P}_k$ denotes the space of polynomials of degree k (i.e., $p \in \mathcal{P}_k \Leftrightarrow p(x_1, x_2) = \sum_{0\le i+j\le k} a_{ij} x_1^i x_2^j$).

It is not difficult to see that $V^k(\mathcal{T})$ is a subspace of $H^1(\Omega)$. Therefore, the subset $V_0^k(\mathcal{T}) \subset V^k(\mathcal{T})$ of functions vanishing at the boundary $\partial\Omega$ is a subspace of $H_0^1(\Omega)$. Therefore, we can define the finite element approximation $u_{\mathcal{T}} \in V_0^k(\mathcal{T})$ to the exact solution u as its Galerkin approximation, i.e.,

$$a(u_{\mathcal{T}}, v) = \langle L, v\rangle \quad \forall v \in V_0^k(\mathcal{T})$$

where a and L are the forms associated with the differential equation (see Example 1.2.1). Since a is continuous and coercive we can apply Lemma 1.3.1 to obtain that there exists a constant $C > 0$, depending only on the differential equation and the domain Ω (indeed, it will depend on the bounds for the coefficients, on the ellipticity constant and on the domain via the constant in the Poincaré inequality), such that

$$\|u - u_{\mathcal{T}}\|_{H^1} \le C \inf_{v\in V_0^k(\mathcal{T})} \|u - v\|_{H^1} \tag{2.1}$$

In order to have convergence we need a family of spaces satisfying (1.17). There are two natural ways of defining finite element spaces with this property: changing the triangulation making the size of the elements go to zero or increasing the degree k of the polynomials. Here, we restrict our analysis to the first strategy, known as the "h version"of the Finite Element Method. For the other method, known as "p version" (where p is what here we call k) we refer to [4].

As is standard in the finite element literature we introduce the parameter h, which measures the size of the triangulation. Assume that for $h \to 0$ we have a family of triangulations $\mathcal{T}_h$ of Ω such that, if we denote by h_T the diameter of T then, $h = max_{T \in \mathcal{T}_h} h_T$. Let ρ_T be the inner diameter of T (i.e., the diameter of the largest ball contained in T). We say that the family of triangulations $\{\mathcal{T}_h\}$ is regular if there exists a constant $\sigma > 0$ such that

$$\frac{h_T}{\rho_T} \leq \sigma \quad \forall T \in \mathcal{T}_h, \quad \forall h \tag{2.2}$$

Associated with $\mathcal{T}_h$ we have the FE space $V_0^k(\mathcal{T}_h)$ that, to simplify notation, we will denote by V_h (we drop the k since it is fixed). Analogously we set $u_h = u_{\mathcal{T}_h}$ for the FE approximation to u.

With these notations, estimate (2.1) reads as follows,

$$\|u - u_h\|_{H^1} \leq C \inf_{v \in V_h} \|u - v\|_{H^1} \tag{2.3}$$

So, in order to prove convergence of u_h to u we need to verify property (1.17) (of course with $N \to \infty$ replaced by $h \to 0$). It is enough to show that there are good approximations to u from V_h. A usual and natural way of doing this is by means of Lagrange interpolation. On each triangle, a set of nodes $P_1, \ldots, P_m$ (with $m = dim \mathcal{P}_k$) for which the Lagrange interpolation is well defined can be given. In other words, these interpolation nodes are such that for any continuous function u there is a unique $u^I \in \mathcal{P}_k$ such that $u(P_i) = u^I(P_i)$ for $i = 1, \ldots, m$. Moreover, these interpolation nodes can be chosen such that the global interpolation $\Pi_h u$, defined to agree with u^I in each triangle, is continuous (note that it is enough to have $k+1$ nodes on each side of the triangle). Figure 2.1 below shows the usual interpolation nodes for $k = 1, 2$ and 3 on a reference triangle (for a general one the nodes are obtained by an affine transformation of this triangle). It is not difficult to see what may be the nodes for any k (see [14]).

The following error estimates for Lagrange interpolation are known (see [14, 7]).

Theorem 2.1.1 *There exists a constant $C > 0$ depending on the degree k and the constant σ in (2.2) but independent of u and h_T such that*

$$\|u - \Pi_h u\|_{L^2(T)} \leq C h_T^{k+1} \|D^{k+1} u\|_{L^2(T)}$$

$$\|u - \Pi_h u\|_{H^1(T)} \leq C h_T^{k} \|D^{k+1} u\|_{H^1(T)}$$

for any triangle T and any $u \in H^{k+1}(T)$, where $D^{k+1}u$ denotes the tensor of all derivatives of order $k+1$ of u.

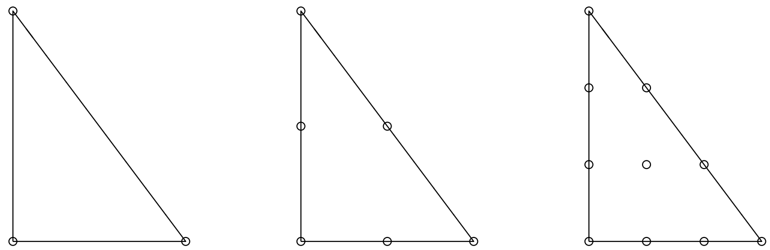

Figure 2.1: Interpolation points for degrees $k = 1, 2$ and 3

Adding the estimates of the theorem over all the triangles of a partition $\{\mathcal{T}_h\}$ we obtain the following global error estimates for the interpolation error.

Corollary 2.1.2 *If the family of triangulations $\{\mathcal{T}_h\}$ is regular then, there exists a constant $C > 0$ independent of h and u such that*

$$\|u - \Pi_h u\|_{L^2(\Omega)} \leq Ch^{k+1}\|D^{k+1}u\|_{L^2(\Omega)}$$
$$\|u - \Pi_h u\|_{H^1(\Omega)} \leq Ch^{k}\|D^{k+1}u\|_{L^2(\Omega)}$$

for any $u \in H^{k+1}(\Omega)$.

Remark 2.1.1 *The regularity assumption (2.2) can be relaxed. For example, in 2-d it can be replaced by a maximum angle condition (see for example [5, 25] and also [18, 26, 31] where results for the 3-d case are obtained).*

2.2 Error estimates for the finite element approximation

Corollary 2.1.2 together with (2.3) yields the following error estimate for the finite element approximation of degree k to u.

Theorem 2.2.1 *If the solution $u \in H^{k+1}(\Omega)$ and the family of triangulations $\{\mathcal{T}_h\}$ is regular, then there exists a constant $C > 0$ independent of h and u such that*

$$\|u - u_h\|_{H^1(\Omega)} \leq Ch^k \|D^{k+1}u\|_{L^2(\Omega)}$$

Theorem 2.2.1 gives an error estimate provided the exact solution is in the Sobolev space $H^{k+1}(\Omega)$ (i.e., the solution is regular enough). Unfortunately, this is not true in general. Let us consider $k = 1$ (linear elements), in this case the theorem says that the error in H^1-norm is of order h whenever the solution is in $H^2(\Omega)$. For example, for the Laplace equation

$$\begin{cases} -\Delta u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases} \tag{2.4}$$

this can be proved if the polygonal domain is convex and, moreover, in this case, the following a priori estimate holds (see [24]),

$$\|u\|_{H^2(\Omega)} \leq C\|f\|_{L^2(\Omega)} \tag{2.5}$$

and consequently we have an error estimate depending only on the right hand side f, i.e., there exists a constant $C > 0$ such that

$$\|u - u_h\|_{H^1(\Omega)} \leq Ch\|f\|_{L^2(\Omega)}$$

(note that we use the letter C as a generic constant, not necessarily the same at each ocurrence, but always independent of h and the functions involved).

When the polygonal domain is not convex the solution is not in general in $H^2(\Omega)$ due to the presence of corner singularities (see [24]) and the error is not of order h. By using more general estimates for the interpolation error and a priori estimates for u in fractional order Sobolev spaces it can be shown that the error is bounded by a constant times h^η where $0 < \eta < 1$ depends on the maximum interior angle of the domain. On the other hand, when the solution has singularities one has to work in practice with locally refined meshes and so, the local mesh size h_T will be very different from one region to another. Therefore, it is reasonable to look at the error in terms of a parameter different than h, for example the number N of nodes in the mesh (see [24] for some results in this direction).

On the other hand, for $k > 1$ and polygonal domain Ω the solution is not in general in $H^{k+1}(\Omega)$ (even if Ω is convex!) and therefore the order of convergence is less than k. The estimate given by Theorem 2.2.1 for $k > 1$ is of interest for the case of a domain with a smooth boundary (where, of course, the triangulation would not cover exactly the domain and so we would have to analyze the error introduced by this fact (see for example [32]). In this case, the a priori estimate (2.5) can be generalized (see [1, 22]) for any k (provided $\partial\Omega$ is C^∞) and an estimate in terms of f can be obtained for the error, showing in particular that the optimal order k is obtained in the H^1-norm, whenever f is in H^{k-1}.

Theorem 2.2.1 gives in particular an error estimate for the L^2-norm. However, in view of Corollary 2.1.2 a natural question is whether the error for the finite element approximation is also of order $k+1$ for regular solutions. The following theorem shows that the answer is positive provided Ω is a convex polygon (or has a smooth boundary). The proof is based on a duality argument due to Aubin and Nitsche (see [14]) and the a priori estimate (2.5), and is in fact very general and has been applied to many situations although, for the sake of simplicity, we consider here the model problem (2.4).

Theorem 2.2.2 *If Ω is a convex polygon, the solution $u \in H^{k+1}(\Omega)$ and the family of triangulations $\{\mathcal{T}_h\}$ is regular, then there exists a constant $C > 0$ independent of h and u such that*

$$\|u - u_h\|_{L^2(\Omega)} \le Ch^{k+1}\|D^{k+1}u\|_{L^2(\Omega)}$$

Proof. Set $e = u - u_h$ and let ϕ be the solution of the problem

$$\begin{cases} -\Delta\phi = e & \text{in } \Omega \\ \phi = 0 & \text{on } \partial\Omega \end{cases} \tag{2.6}$$

Then, using the error equation (1.18) and the estimate for the interpolation error in H^1 given by Theorem 2.1.2 we have

$$\begin{aligned} \|e\|^2_{L^2(\Omega)} &= \int_\Omega e(-\Delta\phi) = \int_\Omega \nabla e \nabla\phi = \int_\Omega \nabla e \nabla(\phi - \Pi_h\phi) \\ &\le \|\nabla e\|_{L^2(\Omega)}\|\nabla(\phi - \Pi_h\phi)\|_{L^2(\Omega)} \le Ch\|\phi\|_{H^2(\Omega)}\|\nabla e\|_{L^2(\Omega)} \end{aligned}$$

and using the a priori estimate (2.5) we obtain

$$\|e\|_{L^2(\Omega)} \leq Ch\|\nabla e\|_{L^2(\Omega)}$$

which, together with Theorem 2.2.1 concludes the proof.□

2.3 Quadrilateral elements

The results obtained in the previous section apply to other finite element spaces. We consider here the case of piecewise polynomials on partitions made of quadrilaterals. First, assume that the elements are rectangles. For a given k the natural space of polynomials on a rectangular element is that of polynomials of degree k in each variable.

For example, consider the case $k = 1$. In order to have continuity between neighboring rectangles the value at a vertex has to be the same for any element sharing that vertex. Therefore, we need a space of, at least, dimension 4 (note that $dim\,\mathcal{P}_1 = 3$ and so it is not an adequate space for rectangles). The appropriate space is that of bilinear functions, i.e., polynomials of the form

$$p(x_1, x_2) = a + bx_1 + cx_2 + dx_1x_2$$

For a general value of k we define

$$\mathcal{Q}_k = \{p \in C^0 : p(x_1, x_2) = \sum_{0\leq i,j\leq k} a_{ij}x_1^i x_2^j\}$$

Note that $\mathcal{Q}_k$ is the tensor product of the spaces of polynomials of degree k in each variable (a property that is useful for computational purposes).

Observe that $\dim \mathcal{Q}_k = (k+1)^2$ and so, in order to define the Lagrange interpolation nodes for $\mathcal{Q}_k$ we can take $(k+1)^2$ equidistributed points in the rectangle. Figure 2.2 shows the interpolation nodes for $k = 1$ and 2. Since, on each side there are $k+1$ nodes, the Lagrange interpolation will be continuous from one element to another.

The error estimates for the Lagrange interpolation given for triangular elements are valid in this case. Indeed, a general proof of Theorem 2.1.2 can be given which is based on the fact that the interpolation is exact for polynomials in $\mathcal{P}_k$, plus approximation properties of $\mathcal{P}_k$, namely, the so called Bramble-Hilbert lemma (see [14]). So, the important point here is that $\mathcal{P}_k \subset \mathcal{Q}_k$.

Consequently, all the convergence results obtained for triangular elements (Theorems 2.2.1 and 2.2.2) hold for rectangular partitions also.

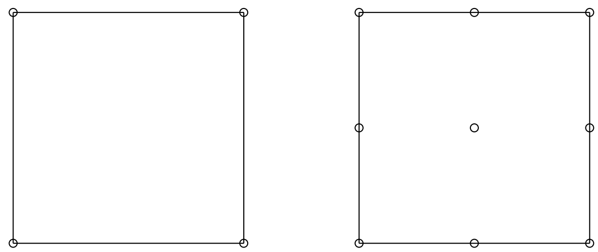

Figure 2.2: Interpolation points for $k = 1$ and 2

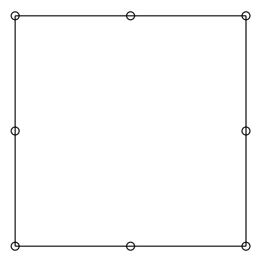

Figure 2.3: Interpolation points for Serendipity elements of order $k = 2$

The space $\mathcal{Q}_k$ can be reduced to a subspace $\mathcal{Q}_k^{red}$ preserving the same convergence properties provided $\mathcal{P}_k \subset \mathcal{Q}_k^{red}$ and that there are enough nodes left on the boundary in order to ensure continuity. To give an example, we consider $k = 2$. In this case, one can eliminate the term corresponding to $x_1^2 x_2^2$ and the interior node. So, $dim\, \mathcal{Q}_k^{red} = 8$ and the interpolation nodes can be taken as those in Figure 2.3. This kind of spaces are called Serendipity elements (see [14] for the general case). Observe that in this way, we reduce the size of the algebraic problem and so the computational cost, still providing the same order of convergence (in fact Theorems 2.2.1 and 2.2.2 hold also in this case).

More generally, we can consider partitions including non rectangular quadrilaterals. Let us analyze the case $k = 1$. A general quadrilateral can be obtained by a bilinear transformation of a reference rectangle $\hat{K}$ with vertices $\hat{P}_j\,, \quad j = 1, \ldots, 4$, i.e., given a quadrilateral K with vertices $P_j\,, \quad j = 1, \ldots, 4$ we can find a transformation $\mathbf{F} = (F_1, F_2)$ such that $F_j \in \mathcal{Q}_1\,, \quad j = 1, 2$, $\mathbf{F}(\hat{P}_j) = P_j\,, \quad j = 1, \ldots, 4$ and $\mathbf{F}(\hat{K}) = K$.

Using $\mathbf{F}$ we can define the space on K by transforming $\mathcal{Q}_1$ in the following way:

$$\tilde{\mathcal{Q}}_1 = \{p \in C^0 : p \circ F \in \mathcal{Q}_1\}$$

Note that $\tilde{\mathcal{Q}}_1$ is not a space of polynomials. However, for computational purposes one can work on the reference element via the transformation F. The convergence results are also valid in this case.

The space $\tilde{\mathcal{Q}}_1$ is an example of the so called isoparametric finite elements (note that the transformation $\mathbf{F}$ has the same form as the interpolation functions on the reference element). Higher order isoparametric finite elements would produce curved boundaries. For example, if we transform a triangle using a quadratic $\mathbf{F}$ we will obtain a "curved side" triangle. So, this kind of elements are useful to approximate curved boundaries (see [14] for more examples and a general analysis).

Chapter 3

Mixed finite elements

Finite element methods in which two spaces are used to approximate two different variables receive the general denomination of mixed methods. In some cases, the second variable is introduced in the formulation of the problem because of its physical interest and it is usually related with some derivatives of the original variable. This is the case, for example, in the elasticity equations, where the stress can be introduced to be approximated at the same time as the displacement. In other cases there are two natural independent variables and so, the mixed formulation is the natural one. This is the case of the Stokes equations, where the two variables are the velocity and the pressure.

The mathematical analysis and applications of mixed finite element methods have been widely developed since the seventies. A general analysis for this kind of methods was first developed by Brezzi [9]. We also have to mention the papers by Babuska [3] and by Crouzeix and Raviart [16] which, although for particular problems, introduced some of the fundamental ideas for the analysis of mixed methods. We also refer the reader to [21, 20], where general results were obtained, and to the books [13, 30, 23].

In this chapter we analyze first the mixed approximation of second order elliptic problems and afterwards we introduce the general abstract setting for mixed formulations and prove general existence and approximation results.

3.1 Mixed approximation of second order elliptic problems

In this section we analyze the mixed approximation of the scalar second order elliptic problem

$$\begin{cases} -div(a\nabla p) = f & \text{in } \Omega \\ p = 0 & \text{on } \partial\Omega \end{cases} \tag{3.1}$$

where $\Omega \subset \mathrm{I\!R}^n \;\; n = 2, 3$ is a polygonal (or polyhedral) domain and $a = a(x)$ is a function bounded by above and below by positive constants (we take this problem to simplify notation but all what we are going to see applies to the case in which a is a matrix like in Example 1.2.1).

In many applications the variable of interest is

$$\mathbf{u} = -a\nabla p$$

and then, it could be desirable to use a mixed finite element method which approximates $\mathbf{u}$ and p simultaneously. With this purpose the problem (3.1) is decomposed into a first order system as follows:

$$\begin{cases} \mathbf{u} + a\nabla p = 0 & \text{in } \Omega \\ div\mathbf{u} = f & \text{in } \Omega \\ p = 0 & \text{on } \partial\Omega \end{cases} \tag{3.2}$$

Writing $\mu = \frac{1}{a(x)}$ the first equation in (3.2) reads

$$\mu\mathbf{u} + \nabla p = 0 \;\; \text{in } \Omega$$

therefore, multiplying by test functions and integrating by parts we obtain the following weak formulation of problem (3.2) appropriate for mixed finite element methods,

$$\begin{cases} \int_\Omega \mu\mathbf{u}\mathbf{v}\, dx - \int_\Omega p\, div\, \mathbf{v}\, dx = 0 & \forall \mathbf{v} \in H(div, \Omega) \\ \int_\Omega q\, div\, \mathbf{u}\, dx = \int_\Omega f q\, dx & \forall q \in L^2(\Omega) \end{cases} \tag{3.3}$$

where

$$H(div, \Omega) = \{\mathbf{v} \in L^2(\Omega)^n : div\, \mathbf{v} \in L^2(\Omega)\}$$

is the Hilbert space with the norm

$$\|\mathbf{v}\|_{H(div,\Omega)} = \|\mathbf{v}\|_{L^2(\Omega)} + \|div\, \mathbf{v}\|_{L^2(\Omega)}$$

Observe that the weak formulation (3.3) involves the divergence of the solution and test functions but not arbitrary first derivatives. This fact allows us to work on the space $H(div, \Omega)$ instead of the smaller $H^1(\Omega)^n$ and this will be important for the finite element approximation because piecewise polynomials vector functions do not need to have both components continuous to be in $H(div, \Omega)$, but only their normal component.

Problem (3.3) can be written as problem (1.6) on the space $H(div, \Omega) \times L^2(\Omega)$ with the symmetric bilinear form

$$c((\mathbf{u}, p), (\mathbf{v}, q)) = \int_\Omega \mu \mathbf{u}\mathbf{v}\, dx - \int_\Omega p\, div\, \mathbf{v}\, dx - \int_\Omega q\, div\, \mathbf{u}\, dx$$

and the linear form

$$L((\mathbf{v}, q)) = -\int_\Omega f q\, dx$$

Indeed, $(\mathbf{u}, p) \in H(div, \Omega) \times L^2(\Omega)$ is the solution of (3.3) if and only if

$$c((\mathbf{u}, p), (\mathbf{v}, q)) = L((\mathbf{v}, q)) \quad \forall (\mathbf{v}, q) \in H(div, \Omega) \times L^2(\Omega)$$

(taking $(\mathbf{v}, 0)$ and $(0, q)$ we recover the two equations (3.3)).

Therefore, we can define Galerkin approximations to $(\mathbf{u}, p)$ using the general method described in Chapter 1. The bilinear form c is not coercive but it can be shown that it satisfies the inf-sup condition (1.9) (and so (1.10) since it is symmetric) and therefore we can apply the results of Chapter 1. Problem (3.3) corresponds to the optimality conditions of a saddle point problem. In the next section we will analyze this kind of problems in an abstract setting to find sufficient conditions for the form c to satisfy the inf-sup condition (both continuous and discrete).

However, the problem considered in this section has some particular properties which allow to simplify the analysis and to obtain better results than those provided by the general theory. We will follow the analysis of [17] (see also [19] where a similar analysis is applied to obtain error estimates in other norms).

In order to define finite element approximations to the solution $(\mathbf{u}, p)$ of (3.3) we need to have finite element subspaces of $H(div, \Omega)$ and $L^2(\Omega)$. Using the notation of Chapter 2 we assume that we have a family $\mathcal{T}_h$ of Ω and so we have to construct piecewise polynomials spaces V_h and Q_h associated with $\mathcal{T}_h$ such that

$$V_h \subset H(div, \Omega) \quad \text{and} \quad Q_h \subset L^2(\Omega)$$

The general theory will show us, in particular, that in order to have stability (and so convergence) V_h and Q_h can not be chosen arbitrarily but they have to be related. For the problem considered here several choices of spaces have been introduced for 2 and 3 dimensional problems and we will recall some of them in the next sections.

Now, we give an error analysis assuming some properties on the spaces that, as we will see, are verified in many cases.

The mixed finite element approximation $(\mathbf{u}_h, p_h) \in V_h \times Q_h$ is defined by

$$\left\{ \begin{array}{rl} \int_\Omega \mu \mathbf{u}_h \mathbf{v}\, dx - \int_\Omega p_h\, div\, \mathbf{v}\, dx = 0 & \forall \mathbf{v} \in V_h \\ \int_\Omega q\, div\, \mathbf{u}_h\, dx = \int_\Omega f q\, dx & \forall q \in Q_h \end{array} \right. \tag{3.4}$$

We assume that the finite element spaces satisfy the following properties:

$$div\, V_h = Q_h \tag{3.5}$$

and that there exists an operator $\Pi_h : H^1(\Omega)^n \to V_h$ such that

$$\int_\Omega div(\mathbf{u} - \Pi_h \mathbf{u}) q = 0 \quad \forall \mathbf{u} \in H^1(\Omega)^n \;\; , \;\; \forall q \in Q_h \tag{3.6}$$

Introducing the L^2-projection $P_h : L^2(\Omega) \to Q_h$, properties (3.5) and (3.6) can be sumarized in the following commutative diagram,

$$\begin{array}{ccccc} H^1(\Omega)^n & \xrightarrow{div} & L^2(\Omega) & & \\ \Pi_h \big\downarrow & & \big\downarrow P_h & & \\ V_h & \xrightarrow{div} & Q_h & \longrightarrow 0 & \end{array}$$

Before starting with the error analysis let us see that under these conditions on the spaces, the discrete solution exists and is unique. Since this is a finite dimensional problem it is enough to show uniqueness. So, assume that

$$\left\{ \begin{array}{rl} \int_\Omega \mu \mathbf{u}_h \mathbf{v}\, dx - \int_\Omega p_h\, div\, \mathbf{v}\, dx = 0 & \forall \mathbf{v} \in V_h \\ \int_\Omega q\, div\, \mathbf{u}_h\, dx = 0 & \forall q \in Q_h \end{array} \right.$$

then, since $div\, V_h \subset Q_h$, we can take $q = div\, \mathbf{u}_h$ in the second equation to conclude that $div\, \mathbf{u}_h = 0$ and taking $\mathbf{v} = \mathbf{u}_h$ in the first equation we obtain $\mathbf{u}_h = 0$. Therefore, $\int_\Omega p_h\, div\, \mathbf{v}\, dx = 0 \ \forall \mathbf{v} \in V_h$. But $div\, V_h \supset Q_h$ and so, taking $\mathbf{v} \in V_h$ such that $div\, \mathbf{v} = p_h$ we obtain that $p_h = 0$.

The following theorem gives an estimate that will provide convergence with optimal order error estimates in the concrete examples.

Theorem 3.1.1 *If the spaces V_h and Q_h are such that properties (3.5) and (3.6) hold, then there exists a constant $C > 0$ depending only on the bounds of the coefficient a of the differential equation such that*

$$\|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)} \le C\|\mathbf{u} - \Pi_h \mathbf{u}\|_{L^2(\Omega)}$$

Proof. Subtracting (3.4) from (3.3) we obtain the error equations

$$\int_\Omega \mu(\mathbf{u} - \mathbf{u}_h)\mathbf{v}\, dx - \int_\Omega (p - p_h)\, div\, \mathbf{v}\, dx = 0 \quad \forall \mathbf{v} \in V_h \tag{3.7}$$

and,

$$\int_\Omega q\, div\,(\mathbf{u} - \mathbf{u}_h)\, dx = 0 \quad \forall q \in Q_h \tag{3.8}$$

Using (3.6) and (3.8) we obtain

$$\int_\Omega q\, div\,(\Pi_h \mathbf{u} - \mathbf{u}_h)\, dx = 0 \quad \forall q \in Q_h$$

and, since (3.5) holds we can take $q = div\,(\Pi_h \mathbf{u} - \mathbf{u}_h)$ to conclude that

$$div\,(\Pi_h \mathbf{u} - \mathbf{u}_h) = 0$$

therefore, taking $\mathbf{v} = \Pi_h \mathbf{u} - \mathbf{u}_h$ in (3.7) we obtain

$$\int_\Omega \mu(\mathbf{u} - \mathbf{u}_h)(\Pi_h \mathbf{u} - \mathbf{u}_h)\, dx = 0$$

and so,

$$\|(\Pi_h \mathbf{u} - \mathbf{u}_h)\|^2_{L^2(\Omega)} \le \|a\|_\infty \int_\Omega \mu(\Pi_h \mathbf{u} - \mathbf{u})(\Pi_h \mathbf{u} - \mathbf{u}_h)\, dx$$

$$\le \|a\|_\infty \|\mu\|_\infty \|(\Pi_h \mathbf{u} - \mathbf{u})\|_{L^2(\Omega)} \|(\Pi_h \mathbf{u} - \mathbf{u}_h)\|_{L^2(\Omega)}$$

and the proof concludes by using the triangle inequality.□

In the next theorem we obtain error estimates for the scalar variable p. For the case in which Ω is convex and the coefficient a is smooth enough to have the a priori estimate

$$\|p\|_{H^2(\Omega)} \leq C\|f\|_{L^2(\Omega)} \tag{3.9}$$

we also obtain a higher order error estimate for $\|P_h p - p_h\|_{L^2(\Omega)}$ by using a duality argument. For the proof of this result we will also assume that the following estimates hold,

$$\|q - P_h q\|_{L^2(\Omega)} \leq Ch^2\|q\|_{H^2(\Omega)} \quad \forall q \in H^2(\Omega) \tag{3.10}$$

and,

$$\|\mathbf{v} - \Pi_h \mathbf{v}\|_{L^2(\Omega)} \leq Ch\|\mathbf{v}\|_{H^1(\Omega)} \quad \forall \mathbf{v} \in H^1(\Omega) \tag{3.11}$$

In particular,

$$\|\Pi_h \mathbf{v}\|_{L^2(\Omega)} \leq C\|\mathbf{v}\|_{H^1(\Omega)} \tag{3.12}$$

The first estimate will be true if the space of polynomials defining Q_h on each element contains $\mathcal{P}_1$. Therefore, this hypothesis excludes only the lowest order cases. The estimate for Π_h holds in all the examples as we are going to see.

The estimate for $\|P_h p - p_h\|_{L^2(\Omega)}$ given by this theorem is important because it can be used to construct superconvergent approximations (i.e., approximations which converge at a higher order than p_h) of p (see for example [6]).

Theorem 3.1.2 *If the spaces V_h and Q_h satisfy (3.5), (3.6) and Π_h satisfies (3.12) then, there exists a constant C such that*

$$\|p - p_h\|_{L^2(\Omega)} \leq C\{\|p - P_h p\|_{L^2(\Omega)} + \|\mathbf{u} - \Pi_h \mathbf{u}\|_{L^2(\Omega)}\} \tag{3.13}$$

If moreover, the equation (3.1) satisfies the a priori estimate (3.9), and (3.11) and (3.10) hold, then, there exists a constant $C > 0$ such that

$$\|P_h p - p_h\|_{L^2(\Omega)} \leq C\{h\|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)} + h^2\|div(\mathbf{u} - \mathbf{u}_h)\|_{L^2(\Omega)}\} \tag{3.14}$$

Proof. First we observe that (3.6) together with (3.12) imply that for any $q \in Q_h$ there exists $\mathbf{v}_h \in V_h$ such that $div\,\mathbf{v}_h = q$ and, $\|\mathbf{v}_h\|_{L^2(\Omega)} \leq C\|q\|_{L^2(\Omega)}$. Indeed, take $\mathbf{v} \in H^1(\Omega)$ such that $div\,\mathbf{v} = q$. Such a $\mathbf{v}$ can be obtained by solving the equation

$$\begin{cases} \Delta\phi = q & \text{in } B \\ \phi = 0 & \text{on } \partial B \end{cases}$$

where B is a ball containing Ω and taking $\mathbf{v} = \nabla\phi$. Then, from the a priori estimate (2.5) on B we know that $\|\mathbf{v}\|_{H^1(\Omega)} \leq C\|q\|_{L^2(\Omega)}$. Now, we take $\mathbf{v}_h = \Pi_h \mathbf{v}$ and it follows from (3.6) and (3.12) that it satisfies the required conditions.

Now, from the error equation (3.7) and (3.5) we have

$$\int_\Omega (P_h p - p_h)\, div\, \mathbf{v}\, dx = \int_\Omega (\mathbf{u} - \mathbf{u}_h)\mathbf{v}\, dx$$

and so, taking $\mathbf{v} \in V_h$ such that $div\, \mathbf{v} = (P_h p - p_h)$ and

$$\|\mathbf{v}\|_{L^2(\Omega)} \leq C\|(P_h p - p_h)\|_{L^2(\Omega)}$$

we obtain

$$\|(P_h p - p_h)\|^2_{L^2(\Omega)} \leq C\|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)}\|(P_h p - p_h)\|_{L^2(\Omega)}$$

which combined with Theorem 3.1.1 and the triangular inequality yields (3.13).

In order to prove (3.14) we use a duality argument. Let ϕ be the solution of

$$\begin{cases} div\,(a\nabla\phi) = P_h p - p_h & \text{in } \Omega \\ \phi = 0 & \text{on } \partial\Omega \end{cases}$$

Using (3.6), (3.5), (3.7), (3.8), (3.10) and (3.11) we have,

$$\|P_h p - p_h\|^2_{L^2(\Omega)} = \int_\Omega (P_h p - p_h) div\,(a\nabla\phi)\, dx = \int_\Omega (P_h p - p_h)\, div\, \Pi_h(a\nabla\phi)\, dx$$

$$= \int_\Omega (p - p_h)\, div\, \Pi_h(a\nabla\phi)\, dx = \int_\Omega \mu(\mathbf{u} - \mathbf{u}_h)(\Pi_h(a\nabla\phi) - a\nabla\phi)\, dx$$

$$+ \int_\Omega (\mathbf{u} - \mathbf{u}_h)\nabla\phi\, dx = \int_\Omega \mu(\mathbf{u} - \mathbf{u}_h)(\Pi_h(a\nabla\phi) - a\nabla\phi)\, dx - \int_\Omega div\,(\mathbf{u} - \mathbf{u}_h)(\phi - P_h\phi)\, dx$$

$$\leq C\|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)} h\|\phi\|_{H^2(\Omega)} + C\|div\,(\mathbf{u} - \mathbf{u}_h)\|_{L^2(\Omega)} h^2\|\phi\|_{H^2(\Omega)}$$

where for the last inequality we have used that a is smooth (for example C^1). The proof concludes by using the a priori estimate (3.9). □

3.2 Examples of mixed finite element spaces

There are several possible choices of spaces satisfying the conditions required for the convergence results proved above. The main question is how to construct V_h, which has to be a subspace of $H(div, \Omega)$, and the associated operator Π_h. In this section we recall some of the known spaces V_h with the corresponding Q_h. We refer the reader to the book [13] for a more complete review of this kind of spaces as well as for other interesting applications of them.

We consider the 2-d case and our first example are the Raviart-Thomas spaces introduced in [29]. Consider first the case of triangular elements. With the notation of Chapter 2 we assume that we have a regular family of triangulations $\{\mathcal{T}_h\}$ of Ω. Given an integer number $k \geq 0$ we define

$$RT_k(T) = \mathcal{P}_k^2 + (x_1, x_2)\mathcal{P}_k \tag{3.15}$$

and

$$V_h = \{\mathbf{v} \in H(div, \Omega) : \mathbf{v}|_T \in RT_k(T) \quad \forall T \in \mathcal{T}_h\} \tag{3.16}$$

In the following lemma we give some elementary but very useful properties of the spaces $RT_k(T)$. We denote with ℓ_i $i = 1, 2, 3$, the sides of a triangle T and with $\mathbf{n}_i$ its corresponding exterior normal.

Lemma 3.2.1 *a)* $dim\, RT_k(T) = (k+1)(k+3)$

b) If $\mathbf{v} \in RT_k(T)$ *then,* $\mathbf{v} \cdot \mathbf{n}_i \in \mathcal{P}_k(\ell_i)$ *for* $i = 1, 2, 3$

c) If $\mathbf{v} \in RT_k(T)$ *is such that* $div\, \mathbf{v} = 0$ *then,* $\mathbf{v} \in \mathcal{P}_k^2$

Proof. Any $\mathbf{v} \in RT_k(T)$ can be written as

$$\mathbf{v} = \mathbf{w} + (\sum_{i+j=k} a_{ij} x_1^{i+1} x_2^j, \sum_{i+j=k} a_{ij} x_1^i x_2^{j+1}) \tag{3.17}$$

with $\mathbf{w} \in \mathcal{P}_k^2$. Then, a) follows from the fact that $dim\, \mathcal{P}_k^2 = (k+2)(k+1)$ and that there are $k+1$ coefficients a_{ij} in the definition of $\mathbf{v}$ above.

Now, if a side is on a line of equation $rx_1 + sx_2 = t$, its normal direction is given by $\mathbf{n} = (r, s)$ and, if $\mathbf{v} = (w_1 + x_1 w + w_2 + x_2 w)$ with $w_1, w_2, w \in \mathcal{P}_k$ we have

$$\mathbf{v} \cdot \mathbf{n} = rw_1 + sw_2 + tw \in \mathcal{P}_k$$

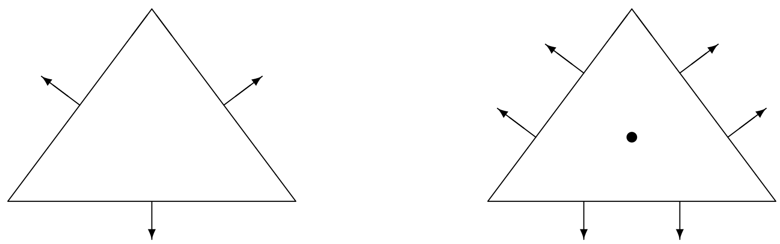

Figure 3.1: Degrees of freedom for RT_0 and RT_1

Finally, if $div\, \mathbf{v} = 0$ we take the divergence in the expression (3.17) and conclude easily that $a_{ij} = 0$ for all i, j and therefore c) holds.□

The approximation space for the scalar variable p is chosen as

$$Q_h = \{q \in L^2(\Omega) : q|_T \in \mathcal{P}_k : \quad \forall T \in \mathcal{T}_h\} \tag{3.18}$$

Note that we do not require any continuity for $q \in Q_h$, since this only needs to be a subspace of $L^2(\Omega)$. With these definitions we see immediately that $div\, V_h \subset Q_h$. The other inclusion, and thus (3.5), will be a consequence of the existence of the operator Π_h satisfying (3.6) as was shown in the proof of Theorem 3.1.2.

In order to construct the operator Π_h we proceed as follows. First we observe that a piecewise polynomial vector function will be in $H(div, \Omega)$ if and only if it has continuous normal component (this can be verified by applying the divergence theorem). Therefore, we can take the normal components at $(k+1)$ points on each side as degrees of freedom in order to ensure continuity. Figure 3.1 shows the degrees of freedom for $k = 0$ and $k = 1$. The arrows indicate normal components values and the filled circle, values of $\mathbf{v}$ (and so it corresponds to two degrees of freedom).

To define the operator $\Pi_h : H^1(\Omega)^2 \to V_h$, the degrees of freedom are taken as averages instead of point values, in order to satisfy condition (3.6). This operator is defined locally in the following lemma.

Lemma 3.2.2 *Given a triangle T and $\mathbf{v} \in H^1(T)^2$ there exists a unique $\Pi_T \mathbf{v} \in RT_k(T)$ such that*

$$\int_{\ell_i} \Pi_T \mathbf{v} \cdot \mathbf{n}_i p_k \, d\ell = \int_{\ell_i} \mathbf{v} \cdot \mathbf{n}_i p_k \, d\ell \quad \forall p_k \in \mathcal{P}_k(\ell_i)\,, \quad i = 1, 2, 3 \tag{3.19}$$

and

$$\int_T \Pi_T \mathbf{v} \cdot \mathbf{p}_{k-1} \, dx = \int_T \mathbf{v} \cdot \mathbf{p}_{k-1} \, dx \quad \forall \mathbf{p}_{k-1} \in \mathcal{P}_{k-1}^2 \tag{3.20}$$

Proof. The number of conditions defining $\Pi_T \mathbf{v}$, $(k+1)(k+3)$, equals the dimension of $RT_k(T)$. Therefore, it is enough to verify uniqueness. So, take $\mathbf{v} \in RT_k(T)$ such that

$$\int_{\ell_i} \mathbf{v} \cdot \mathbf{n}_i p_k \, d\ell = 0 \quad \forall p_k \in \mathcal{P}_k(\ell_i)\,, \quad i = 1, 2, 3 \tag{3.21}$$

and

$$\int_T \mathbf{v} \cdot \mathbf{p}_{k-1} \, dx = 0 \quad \forall \mathbf{p}_{k-1} \in \mathcal{P}_{k-1}^2 \tag{3.22}$$

From $b)$ of Lemma 3.2.1 and (3.21) it follows that $\mathbf{v} \cdot \mathbf{n}_i = 0$. On the other hand, using (3.21) and (3.22) we have

$$\int_T (div\, \mathbf{v})^2 \, dx = -\int_T \mathbf{v} \cdot \nabla(div\, \mathbf{v}) \, dx + \int_{\partial T} \mathbf{v} \cdot \mathbf{n} \, div\, \mathbf{v} \, d\ell = 0$$

because $\nabla(div\, \mathbf{v}) \in \mathcal{P}_{k-1}^2$ and $div\, \mathbf{v}|_{\ell_i} \in \mathcal{P}_k(\ell_i)$. Consequently $div\, \mathbf{v} = 0$ which together with $c)$ of Lemma 3.2.1 implies that there exists $\psi \in \mathcal{P}_{k+1}$ such that $\mathbf{v} = curl\, \psi = (-\frac{\partial \psi}{\partial x_2}, \frac{\partial \psi}{\partial x_1})$. But, since $\mathbf{v} \cdot \mathbf{n}_i = 0$, the tangential derivatives of ψ vanish on the three sides. Therefore ψ is constant on ∂T and, since it is defined up to a constant, we can take $\psi = 0$ on ∂T and then, $\psi = b_T p_{k-2}$ where b_T is a bubble function on T (i.e., a polynomial of degree 3 vanishing on ∂T) and $p_{k-2} \in \mathcal{P}_{k-2}$.

Now, using again (3.22) we have that, for any $\mathbf{p} = (p_1, p_2) \in \mathcal{P}_{k-1}^2$

$$0 = \int_T curl\, \psi \cdot \mathbf{p} \, dx = \int_T \psi(\frac{\partial p_1}{\partial x_2} - \frac{\partial p_2}{\partial x_1}) \, dx = \int_T b_T p_{k-2}(\frac{\partial p_1}{\partial x_2} - \frac{\partial p_2}{\partial x_1}) \, dx$$

and taking $\mathbf{p}$ such that $(\frac{\partial p_1}{\partial x_2} - \frac{\partial p_2}{\partial x_1}) = p_{k-2}$ we conclude that $p_{k-2} = 0$ and then $\mathbf{v} = 0$ as we wanted to see.□

In view of Lemma 3.2.2 we can define the operator $\Pi_h : H^1(\Omega)^2 \to V_h$ by $\Pi_h \mathbf{v}|_T = \Pi_T \mathbf{v}$. Observe that $\Pi_h \mathbf{v} \in V_h$ because the degrees of freedom defining Π_T enforce the continuity of the normal component between two neigbour elements. On the other hand it is easy to see that Π_h satisfies the fundamental property (3.6). Indeed, by using (3.19) and(3.20) it follows that for any $\mathbf{v} \in H^1(T)^2$ and any $q \in \mathcal{P}_k$

$$\int_T div\,(\mathbf{v} - \Pi_T \mathbf{v}) q\, dx = -\int_T (\mathbf{v} - \Pi_T \mathbf{v}) \cdot \nabla q\, dx + \int_{\partial T} (\mathbf{v} - \Pi_T \mathbf{v}) \cdot \mathbf{n} q = 0$$

In order to prove convergence by using the general results obtained in Section 3.1, we need to analyze the approximation properties of the operator Π_h. The following lemma gives error estimates for $\mathbf{v} - \Pi_T \mathbf{v}$ on each T. We omit the proof, which uses general standard arguments for polynomial preserving operators (see [14]). The main difference with the proof for Lagrange interpolation is that here we have to use an appropriate transformation which preserves the degrees of freedom defining $\Pi_T \mathbf{v}$. It is known as the Piola transform and is defined in the following way. Given the affine map F which transform $\widehat{T}$ into T we define for $\widehat{\mathbf{v}} \in L^2(\widehat{T})^2$

$$\mathbf{v}(x) = \frac{1}{J(\hat{x})} DF(\hat{x}) \widehat{\mathbf{v}}(\hat{x})$$

where $x = F(\hat{x})$, DF is the Jacobian matrix of F and, $J = |det\, DF|$. We refer to [29, 33] for details.

Lemma 3.2.3 *There exists a constant $C > 0$ depending on the constant σ in (2.2) such that for any $\mathbf{v} \in H^m(T)^2$ and $1 \leq m \leq k+1$*

$$\|\mathbf{v} - \Pi_T \mathbf{v}\|_{L^2(T)} \leq C h_T^m \|\mathbf{v}\|_{H^m(T)} \tag{3.23}$$

Now we can apply the results of Section 3.1 together with (3.23) to obtain the following error estimates for the mixed finite element approximation of problem (3.1) obtained with the Raviart-Thomas space of order k.

Theorem 3.2.4 *If the family of triangulations $\{\mathcal{T}_h\}$ is regular and $\mathbf{u} \in H^{k+1}(\Omega)$ and $p \in H^{k+1}(\Omega)$, then the mixed finite element approximation $(\mathbf{u}_h, p_h) \in V_h \times Q_h$ satisfies*

$$\|\mathbf{u} - \mathbf{u}_h\|_{L^2(\Omega)} \leq C h^{k+1} \|\mathbf{u}\|_{H^{k+1}(\Omega)} \tag{3.24}$$

and

$$\|p - p_h\|_{L^2(\Omega)} \leq C h^{k+1} \{\|\mathbf{u}\|_{H^{k+1}(\Omega)} + \|p\|_{H^{k+1}(\Omega)}\} \tag{3.25}$$

and when Ω is convex, $k \geq 1$ and $p \in H^{k+2}(\Omega)$

$$\|P_h p - p_h\|_{L^2(\Omega)} \leq C h^{k+2} \{\|\mathbf{u}\|_{H^{k+1}(\Omega)} + \|p\|_{H^{k+2}(\Omega)}\} \tag{3.26}$$

Proof. The result follows immediately from Theorems 3.1.1 and 3.1.2, (3.23) and standard error estimates for the L^2 projection. $\square$

The Raviart-Thomas spaces defined above were the first introduced for the mixed approximation of second order elliptic problems. They were constructed in order to approximate both vector and scalar variables with the same order. However, if one is most interested in the approximation of the vector variable $\mathbf{u}$ one can try to use different order approximations for each variable in order to reduce the degrees of freedom (thus, reducing the computational cost) while preserving the same order of convergence for $\mathbf{u}$ as the one provided by the RT_k spaces. This is the main idea to define the following spaces which were introduced by Brezzi, Douglas and Marini [12]. Although with this choice the order of convergence for p is reduced, estimate (3.26) allows to improve it by a post processing of the computed solution [12]. As for all the examples below, we will define the local spaces for each variable. Clearly, the global spaces V_h and Q_h are defined as in (3.16) and (3.18) replacing RT_k and $\mathcal{P}_k$ by the corresponding local spaces.

For $k \geq 1$ and T a triangle, the $BDM_k(T)$ is defined in the following way:

$$BDM_k(T) = \mathcal{P}_k^2 \tag{3.27}$$

and the corresponding space for the scalar variable is $\mathcal{P}_{k-1}$.

Observe that $dim\, BDM_k(T) = (k+1)(k+2)$. For example, $dim\, BDM_1(T) = 6$ and $dim\, BDM_2(T) = 12$. Figure 3.2 shows the degrees of freedom for these two spaces. The arrows correspond to normal component degrees of freedom while the circles indicate the internal degrees of freedom corresponding to the second and third conditions in the definition of Π_T below.

The operator Π_T for this case is defined by the following degrees of freedom:

$$\int_{\ell_i} \Pi_T \mathbf{v} \cdot \mathbf{n}_i p_k \, d\ell = \int_{\ell_i} \mathbf{v} \cdot \mathbf{n}_i p_k \, d\ell \quad \forall p_k \in \mathcal{P}_k(\ell_i)\,, \quad i = 1, 2, 3$$

$$\int_T \Pi_T \mathbf{v} \cdot \nabla p_{k-1} \, dx = \int_T \mathbf{v} \cdot \nabla p_{k-1} \, dx \quad \forall p_{k-1} \in \mathcal{P}_{k-1}$$

and, when $k \geq 2$

$$\int_T \Pi_T \mathbf{v} \cdot curl\, b_T p_{k-2} \, dx = \int_T \mathbf{v} \cdot curl\, b_T p_{k-2} \, dx \quad \forall p_{k-2} \in \mathcal{P}_{k-2}$$

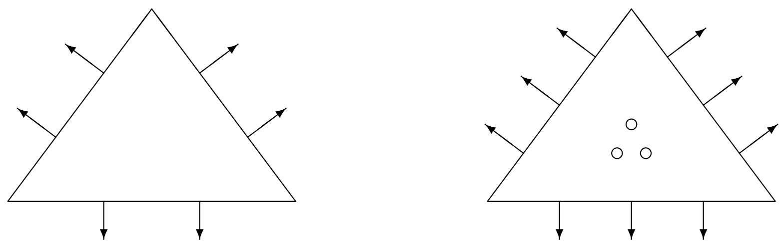

Figure 3.2: Degrees of freedom for BDM_1 and BDM_2

The reader can check that all the conditions for convergence are satisfied in this case. Property (3.6) follows from the definition of Π_T and the proof of its existence is similar to that of Lemma 3.2.2. Consequently, the general analysis provides the same error estimate for $\mathbf{u}$ as that in Theorem 3.2.4 while for p the order of convergence is reduced in one with respect to the estimate in that theorem, i.e.,

$$\|p - p_h\|_{L^2(\Omega)} \leq Ch^k\{\|\mathbf{u}\|_{H^k(\Omega)} + \|p\|_{H^k(\Omega)}\}$$

and the estimate for $\|P_h p - p_h\|_{L^2(\Omega)}$ is the same as that in Theorem 3.2.4 with the restriction $k \geq 2$.

Several rectangular elements have been introduced for mixed approximations also. We recall some of them (and again refer to [13] for a more complete review).

First we define the spaces introduced by Raviart and Thomas [29]. For nonnegative integers j, k we call

$$\mathcal{Q}_{k,m} = \{q \in C^0 : q(x_1, x_2) = \sum_{i=0}^{k}\sum_{j=0}^{m} a_{ij} x_1^i x_2^j\}$$

then, the $RT_k(R)$ space on a rectangle R is given by

$$RT_k(R) = \mathcal{Q}_{k+1,k} \times \mathcal{Q}_{k,k+1}$$

and the space for the scalar variable is $\mathcal{Q}_k$. It can be checked that $dim\, RT_k(R) = 2(k+1)(k+2)$. Figure 3.3 shows the degrees of freedom for $k = 0$ and $k = 1$.

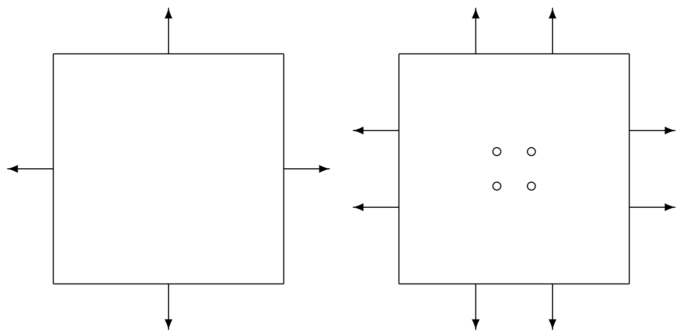

Figure 3.3: Degrees of freedom for RT_0 and RT_1

Denoting with ℓ_i, $i = 1, 2, 3, 4$ the four sides of R, the degrees of freedom defining the operator Π_T for this case are

$$\int_{\ell_i} \Pi_T \mathbf{v} \cdot \mathbf{n}_i p_k \, d\ell = \int_{\ell_i} \mathbf{v} \cdot \mathbf{n}_i p_k \, d\ell \quad \forall p_k \in \mathcal{P}_k(\ell_i) \,, \quad i = 1, 2, 3, 4$$

and (for $k \geq 1$)

$$\int_R \Pi_T \mathbf{v} \cdot \phi_k \, dx = \int_R \mathbf{v} \cdot \phi_k \, dx \quad \forall \phi_k \in \mathcal{Q}_{k-1,k} \times \mathcal{Q}_{k,k-1}$$

Our last example in the 2-d case are the spaces introduced by Brezzi, Douglas and Marini on rectangular elements. They are defined for $k \geq 1$ as

$$BDM_k(R) = \mathcal{P}_k^2 + \langle curl\,(x^{k+1}y) \rangle + \langle curl\,(xy^{k+1}) \rangle$$

and the associated scalar space is $\mathcal{P}_{k-1}$. It is easy to see that $dim\, BDM_k(R) = (k+1)(k+2) + 2$. The degrees of freedom for $k = 1$ and $k = 2$ are shown in Figure 3.4.

The operator Π_T is defined by

$$\int_{\ell_i} \Pi_T \mathbf{v} \cdot \mathbf{n}_i p_k \, d\ell = \int_{\ell_i} \mathbf{v} \cdot \mathbf{n}_i p_k \, d\ell \quad \forall p_k \in \mathcal{P}_k(\ell_i) \,, \quad i = 1, 2, 3, 4$$

and (for $k \geq 2$)

$$\int_R \Pi_T \mathbf{v} \cdot \mathbf{p}_{k-2} \, dx = \int_R \mathbf{v} \cdot \mathbf{p}_{k-2} \, dx \quad \forall \mathbf{p}_{k-2} \in \mathcal{P}_{k-2}^2$$

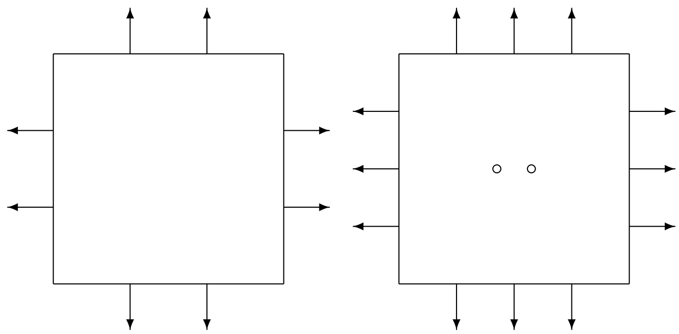

Figure 3.4: Degrees of freedom for BDM_1 and BDM_2

The RT_k as well as the BDM_k spaces on rectangles have analogous properties to those on triangles. Therefore the same error estimates obtained for triangular elements are valid in both cases.

3-d extensions of the spaces defined above have been introduced by Nedelec [27, 28] and by Brezzi, Douglas, Durán and Fortin [10]. For tetrahedral elements the spaces are defined in an analogous way, although the construction of the operator Π_T requires a different analysis (we refer to [27] for the extension of the RT_k spaces and to [28, 10] for the extension of the BDM_k spaces). In the case of 3-d rectangular elements, the extensions of RT_k are again defined in an analogous way [27] and the extensions of BDM_k [10] can be defined for a 3-d rectangle R by

$$BDDF_k(R) = \mathcal{P}_k^3 + \langle\{curl\,(0,0,xy^{i+1}z^{k-i}),\quad i=0,\dots,k\}\rangle$$
$$+\langle\{curl\,(0,x^{k-i}yz^{i+1},0),\quad i=0,\dots,k\}\rangle$$
$$+\langle\{curl\,(x^{i+1}y^{k-i}z,0,0),\quad i=0,\dots,k\}\rangle$$

All the convergence results obtain in 2-d can be extended for the 3-d spaces mentioned here. Other families of spaces, in both 2 and 3 dimensions which are intermediate between the RT and the BDM spaces were introduced and analized by Brezzi, Douglas, Fortin and Marini [11].

3.3 The general abstract setting

The problem considered in the previous section is a particular case of a general class of problems that we are going to analize in this section. Let V

and Q be two Hilbert spaces and suppose that $a(\ ,\ )$ and $b(\ ,\ )$ are continuous bilinear forms on $V \times V$ and $V \times Q$ respectively, i.e.,

$$|a(u, v)| \leq \|a\|\|u\|_V\|v\|_V \qquad \forall u \in V, \forall v \in V$$

and

$$|b(v, q)| \leq \|b\|\|v\|_V\|q\|_Q \qquad \forall v \in V,\ \forall q \in Q$$

We can introduce the continuous operators $A : V \to V'$, $B : V \to Q'$ and its adjoint $B^* : Q \to V'$ defined by,

$$\langle Au, v\rangle_{V' \times V} = a(u, v)$$

and

$$\langle Bv, q\rangle_{Q' \times Q} = b(v, q) = \langle v, B^* q\rangle_{V \times V'}$$

Consider the following problem: given $f \in V'$ and $g \in Q'$ find $(u, p) \in V \times Q$ solution of

$$\begin{cases} a(u, v) + b(v, p) = \langle f, v\rangle & \forall v \in V \\ b(u, q) = \langle g, q\rangle & \forall q \in Q \end{cases} \tag{3.28}$$

which can also be written as

$$\begin{cases} Au + B^* p = f & \text{in } V' \\ Bu = g & \text{in } Q' \end{cases} \tag{3.29}$$

This is a particular (but very important!) case of the general problem (1.6) analyzed in Chapter 1. Indeed, equations (3.28) can be written as

$$c((u, p), (v, q)) = \langle f, v\rangle + \langle g, q\rangle \ \ \forall (v, q) \in V \times Q \tag{3.30}$$

where c is the continuous bilinear form on $V \times Q$ defined by

$$c((u, p), (v, q)) = a(u, v) + b(v, p) + b(u, q)$$

The form c is not coercive and so, in order to apply the theory one would have to show that it satisfies the inf-sup conditions (1.9) and (1.10). We will give sufficient conditions (indeed they are also necessary although we are not going to prove it here, we refer to [13, 23]) on the forms a and b for the existence and uniqueness of a solution of problem (3.28). Below, we will also show that their discrete version ensures the stability condition (i.e., the inf-sup condition (1.9) for the bilinear form c) and therefore, optimal order error estimates for the Galerkin approximations. These results were obtained by Brezzi [9] (see also [13] where more general results are proven).

Let us introduce $W = KerB \subset V$ and for $g \in Q'$, $W(g) = \{v \in V : Bv = g\}$. Now, if $(u, p) \in V \times Q$ is a solution of (3.28) then, it is easy to see that $u \in W(g)$ is a solution of the following problem,

$$a(u, v) = \langle f, v \rangle \qquad \forall v \in W \tag{3.31}$$

We will find conditions under which problems (3.28) and (3.31) are equivalent, in the sense that given a solution $u \in W(g)$ of (3.31), there exists a unique $p \in Q$ such that (u, p) is a solution of (3.28).

Lemma 3.3.1 *The following properties are equivalent:*

a) There exists $\beta > 0$ such that

$$\sup_{v \in V} \frac{b(v, q)}{\|v\|_V} \geq \beta \|q\|_Q \qquad \forall q \in Q \tag{3.32}$$

b) B^* is an isomorphism from Q onto W^0 and,

$$\|B^* q\|_{V'} \geq \beta \|q\|_Q \qquad \forall q \in Q \tag{3.33}$$

c) B is an isomorphism from $W^\perp$ onto Q' and,

$$\|Bv\|_{Q'} \geq \beta \|v\|_V \qquad \forall v \in W^\perp \tag{3.34}$$

Proof. Assume that $a)$ holds. Then, (3.33) is satisfied and so B^* is injective and ImB^* is a closed subspace of V' (this follows easily from (3.33) as was shown in the proof of Theorem 1.2.1). Consequently, using (1.13) we obtain that $ImB^* = W^0$ and therefore $b)$ holds.

Now, we observe that W^0 can be isometrically identified with $(W^\perp)'$. Indeed, denoting with $P^\perp : V \to W^\perp$ the orthogonal projection, for any $g \in (W^\perp)'$ we define $\tilde{g} \in W^0$ by $\tilde{g} = g \circ P^\perp$ and it is easy to check that $g \to \tilde{g}$ is an isometric bijection from $(W^\perp)'$ onto W^0 and then, we can identify these two spaces. Therefore $b)$ and $c)$ are equivalent.□

Corollary 3.3.2 *If the form b satisfies (3.32) then, problems (3.28) and (3.31) are equivalent, that is, there exists a unique solution of (3.28) if and only if there exists a unique solution of (3.31).*

Proof. If (u,p) is a solution of (3.28) we know that $u \in W(g)$ and that it is a solution of (3.31). It rests only to check that for a solution $u \in W(g)$ of (3.31) there exists a unique $p \in Q$ such that $B^*p = f - Au$, but this follows from b) of the previous lemma since, as it is easy to check, $f - Au \in W^0$.□

Now we can prove the fundamental existence and uniqueness theorem for problem (3.28).

Theorem 3.3.3 *If b satisfies the inf-sup condition (3.32) and there exists $\alpha > 0$ such that a satisfies*

$$\sup_{v\in W} \frac{a(u,v)}{\|v\|_V} \geq \alpha \|u\|_V \qquad \forall u \in W \tag{3.35}$$

$$\sup_{u\in W} \frac{a(u,v)}{\|u\|_V} \geq \alpha \|v\|_V \qquad \forall v \in W \tag{3.36}$$

then there exists a unique solution $(u,p) \in V \times Q$ of problem (3.28) and moreover,

$$\|u\|_V \leq \frac{1}{\alpha}\|f\|_{V'} + \frac{1}{\beta}(1 + \frac{\|a\|}{\alpha})\|g\|_{Q'} \tag{3.37}$$

and

$$\|p\|_Q \leq \frac{1}{\beta}(1 + \frac{\|a\|}{\alpha})\|f\|_{V'} + \frac{\|a\|}{\beta^2}(1 + \frac{\|a\|}{\alpha})\|g\|_{Q'} \tag{3.38}$$

Proof. First we show that there exists a solution $u \in W(g)$ of problem (3.31). Since (3.32) holds, we know from Lemma 3.3.1 that there exists a unique $u_0 \in W^\perp$ such that $Bu_0 = g$ and

$$\|u_0\|_V \leq \frac{1}{\beta}\|g\|_{Q'} \tag{3.39}$$

then, the existence of a solution $u \in W(g)$ of (3.31) is equivalent to the existence of $w = u - u_0 \in W$ such that

$$a(w,v) = \langle f, v\rangle - a(u_0,v) \quad \forall v \in W$$

but, from (3.35), (3.36) and Theorem 1.2.1, it follows that such a w exists and moreover,

$$\|w\|_V \leq \frac{1}{\alpha}\{\|f\|_{V'} + \|a\|\|u_0\|_V\} \leq \frac{1}{\alpha}\{\|f\|_{V'} + \frac{\|a\|}{\beta}\|g\|_{Q'}\}$$

where we have used (3.39).

Therefore, $u = w + u_0$ is a solution of (3.31) and satisfies (3.37).

Now, from Corollary (3.3.2) it follows that there exists a unique $p \in Q$ such that (u, p) is a solution of (3.28). On the other hand, from Lemma 3.3.1 it follows that (3.33) holds and by using it, it is easy to check that

$$\|p\|_Q \leq \frac{1}{\beta}\{\|f\|_{V'} + \|a\|\|u\|_V\}$$

which combined with (3.37) yields (3.38). Finally, the uniqueness of the solution follows from (3.37) and (3.38).□

Assume now that we have two families of subspaces $V_h \subset V$ and $Q_h \subset Q$. We can define the Galerkin approximation $(u_h, p_h) \in V_h \times Q_h$ to be the solution $(u, p) \in V \times Q$ of problem (3.28), i.e., (u_h, p_h) satisfies,

$$\left\{ \begin{aligned} a(u_h, v) + b(v, p_h) &= \langle f, v\rangle \qquad \forall v \in V_h \\ b(u_h, q) &= \langle g, q\rangle \qquad \forall q \in Q_h \end{aligned} \right. \tag{3.40}$$

For the error analysis it is convenient to introduce the associated operator $B_h : V_h \to Q'_h$ defined by

$$\langle B_h v, q\rangle_{Q'_h \times Q_h} = b(v, q)$$

and the subsets of V_h, $W_h = Ker\, B_h$ and

$$W_h(g) = \{v \in V_h \ : B_h v = g \ \text{ in } \ Q'_h\}$$

where g is restricted to Q_h.

In order to have a well-defined Galerkin approximation we need to know that there exists a unique solution $(u_h, p_h) \in V_h \times Q_h$ of problem (3.40). In view of Theorem 3.3.3, this will be true if there exist $\alpha^* > 0$ and $\beta^* > 0$ such that

$$\sup_{v \in W_h} \frac{a(u, v)}{\|v\|_V} \geq \alpha^* \|u\|_V \qquad \forall u \in W_h \tag{3.41}$$

$$\sup_{u \in W_h} \frac{a(u, v)}{\|u\|_V} \geq \alpha^* \|v\|_V \qquad \forall v \in W_h \tag{3.42}$$

and

$$\sup_{v \in V_h} \frac{b(v, q)}{\|v\|_V} \geq \beta^* \|q\|_Q \qquad \forall q \in Q_h \tag{3.43}$$

In fact, as we have mentioned in Chapter 1, (3.42) follows from (3.41) since W_h is finite dimensional.

Now, we can prove the fundamental general error estimates due to Brezzi [9].

Theorem 3.3.4 *If the forms a and b satisfy (3.41), (3.42) and (3.43), there exists $C > 0$, depending only on α^*, β^*, $\|a\|$ and $\|b\|$ such that the following estimates hold. In particular, if the constants α^* and β^* are independent of h, then C is independent of h.*

$$\|u - u_h\|_V + \|p - p_h\|_Q \leq C\{\inf_{v \in V_h} \|u - v\|_V + \inf_{q \in Q_h} \|p - q\|_Q\} \tag{3.44}$$

and, when $Ker\, B_h \subset Ker\, B$,

$$\|u - u_h\|_V \leq C \inf_{v \in V_h} \|u - v\|_V \tag{3.45}$$

Proof. From Theorem 3.3.3 we know that, under these assumptions, there exists a unique solution $(u_h, p_h) \in V_h \times Q_h$ of (3.40) and that it satisfies

$$\|u_h\|_V + \|p_h\|_Q \leq C\{\|f\|_{V'} + \|g\|_{Q'}\}$$

with $C = C(\alpha^*, \beta^*, \|a\|, \|b\|)$. Therefore, the form c defined in (3.30) satisfies the condition (1.19) on the space $V_h \times Q_h$ with the inverse of this constant C (see Remark 1.4.1). Therefore, we can apply Lemma 1.4.1 to obtain the estimate (3.44).

On the other hand, we know that $u_h \in W_h(g)$ is the solution of

$$a(u_h, v) = \langle f, v \rangle \quad \forall v \in W_h \tag{3.46}$$

and, since $W_h \subset W$, subtracting (3.46) from (3.31) we have,

$$a(u - u_h, v) = 0 \quad \forall v \in W_h$$

Now, since a satisfies (3.41), given $w \in W_h(g)$ we can proceed as in Lemma 1.4.1 to show that

$$\|w - u_h\|_V \leq \frac{\|a\|}{\alpha^*}\|u - w\|_V$$

and therefore,

$$\|u - u_h\|_V \leq (1 + \frac{\|a\|}{\alpha^*}) \inf_{w \in W_h(g)} \|u - w\|_V$$

To conclude the proof we will see that, if (3.43) holds, then

$$\inf_{w \in W_h(g)} \|u - w\|_V \leq (1 + \frac{\|b\|}{\beta^*}) \inf_{v \in V_h} \|u - v\|_V \tag{3.47}$$

Given $v \in V_h$, from Lemma 3.3.1 we know that there exists a unique $z \in W_h^\perp$ such that

$$b(z, q) = b(u - v, q) \quad \forall q \in Q_h$$

and

$$\|z\|_V \leq \frac{\|b\|}{\beta^*} \|u - v\|_V$$

thus, $w = z + v \in V_h$ satisfies $B_h w = g$, that is, $w \in W_h(g)$. But

$$\|u - w\|_V \leq \|u - v\|_V + \|z\|_V \leq (1 + \frac{\|b\|}{\beta^*}) \|u - v\|_V$$

and so (3.47) holds. □

In the applications, a very useful criterion to check the inf-sup condition (3.43) is the following result due to Fortin [21].

Theorem 3.3.5 *Assume that (3.32) holds. Then, the discrete inf-sup condition (3.43) holds with a constant $\beta^* > 0$ independent of h, if and only if, there exists an operator*

$$\Pi_h : V \to V_h$$

such that

$$b(v - \Pi_h v, q) = 0 \quad \forall v \in V\,,\ \forall q \in Q_h \tag{3.48}$$

and

$$\|\Pi_h v\|_V \leq C \|v\|_V \quad \forall v \in V \tag{3.49}$$

with a constant $C > 0$ independent of h.

Proof. Assume that such an operator Π_h exists. Then, from (3.48), (3.49) and (3.32) we have, for $q \in Q_h$,

$$\beta \|q\|_Q \leq \sup_{v \in V} \frac{b(v, q)}{\|v\|_V} = \sup_{v \in V} \frac{b(\Pi_h v, q)}{\|v\|_V} \leq C \sup_{v \in V} \frac{b(\Pi_h v, q)}{\|\Pi_h v\|_V}$$

and therefore, (3.43) holds with $\beta^* = \beta / C$.

Conversely, suppose that (3.43) holds with β^* independent of h. Then, from (3.34) we know that, for any $v \in V$, there exists a unique $v_h \in W_h^\perp$ such that

$$b(v_h, q) = b(v, q) \quad \forall q \in Q_h$$

and

$$\|v_h\|_V \leq \frac{\|b\|}{\beta^*}\|v\|_V$$

and therefore, $\Pi_h v = v_h$ defines the required operator. □

Remark 3.3.1 *In practice, it is sometimes enough to show the existence of the operator Π_h verifying (3.48) and (3.49) for $v \in S$, where $S \subset V$ is a subspace where the exact solution belongs, and the norm on the right hand side of (3.49) is replaced by a strongest norm (that of the space S). This is in some cases easier because the explicit construction of the operator Π_h requires regularity assumptions which do not hold for a general function in V. For example, in the problem analyzed in the previous section we have constructed this operator on a subspace of $V = H(div, \Omega)$ because the degrees of freedom defining the operator do not make sense in $H(div, T)$. Indeed, we need more regularity for $\mathbf{v}$ (for example $\mathbf{v} \in H^1(T)^2$) in order to have the integral of the normal component of $\mathbf{v}$ against a polynomial on a side ℓ of T well defined. It is possible to show the existence of Π_h defined on $H(div, \Omega)$ satisfying (3.48) and (3.49) (see [21]). However, as we have seen, this is not really necessary to obtain error estimates.*

Bibliography

[1] Agmon, S., Douglis, A., Nirenberg, L., Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions I, *Comm. Pure Appl. Math.*, **12**, 623-727, 1959.

[2] Babuska, I., Error bounds for finite element method, *Numer. Math.*, **16**, 322-333, 1971.

[3] Babuska, I., The finite element method with lagrangian multipliers, *Numer. Math.*, **20**, 179-192, 1973.

[4] Babuska, I., Suri, M., The p and h-p versions of the finite element method, basic principles and properties, *SIAM Review*, **36**, 578-632, 1994.

[5] Babuska, I., Azis, A. K., On the angle condition in the finite element method, *SIAM J. Numer. Anal.*, **13**, 214-226, 1976.

[6] Bramble, J.H., Xu, J.M., Local post-processing technique for improving the accuracy in mixed finite element approximations, *SIAM J. Numer. Anal.*, **26**, 1267-1275, 1989.

[7] Brenner, S., Scott, L.R., *The Mathematical Analysis of Finite Element Methods*, Springer Verlag, 1994.

[8] Brezis, H., *Analyse Fonctionnelle, Théorie et Applications*, Masson, 1983.

[9] Brezzi, F., On the existence, uniqueness and approximation of saddle point problems arising from lagrangian multipliers, *R.A.I.R.O. Anal. Numer.*, **8**, 129-151, 1974.

[10] Brezzi, F., Douglas, J., Durán, R., Fortin, M., Mixed finite elements for second order elliptic problems in three variables, *Numer. Math.*, **51**, 237-250, 1987.

[11] Brezzi, F., Douglas, J., Fortin, M., Marini, L.D., Efficient rectangular mixed finite elements in two and three space variables, *Math. Model. Numer. Anal.*, **21**, 581-604, 1987.

[12] Brezzi, F., Douglas, J., Marini, L.D., Two families of mixed finite elements for second order elliptic problems, *Numer. Math.*, **47**, 217-235, 1985.

[13] Brezzi, F., Fortin, M., *Mixed and Hybrid Finite Element Methods*, Springer Verlag, 1991.

[14] Ciarlet, P. G., *The Finite Element Method for Elliptic Problems*, North Holland, 1978.

[15] Ciarlet, P. G., *Mathematical Elasticity, Volume 1. Three-Dimensional Elasticity*, North Holland, 1988.

[16] Crouzeix, M., Raviart, P.A., Conforming and non-conforming finite element methods for solving the stationary Stokes equations,*R.A.I.R.O. Anal. Numer.*, **7**, 33-76, 1973.

[17] Douglas, J., Roberts, J.E., Global estimates for mixed methods for second order elliptic equations, *Math. Comp.*, **44**, 39-52, 1985.

[18] Durán, R.G., Error estimates for narrow 3-d finite elements, preprint.

[19] Durán, R.G., Error Analysis in L^p for Mixed Finite Element Methods for Linear and quasilinear elliptic problems, *R.A.I.R.O. Anal. Numér*, **22**, 371-387, 1988.

[20] Falk, R.S., Osborn, J., Error estimates for mixed methods, *R.A.I.R.O. Anal. Numer.*, **4**, 249-277, 1980.

[21] Fortin, M., An analysis of the convergence of mixed finite element methods, *R.A.I.R.O. Anal. Numer.*, **11**, 341-354, 1977.

[22] Gilbarg, D., Trudinger, N.S., *Elliptic Partial Differential Equations of Second Order*, Springer Verlag, 1983.

[23] Girault, V., Raviart, P.A., *Finite Element Methods for Navier-Stokes Equations, Theory and Algorithms*, Springer Verlag, 1986.

[24] Grisvard, P., *Elliptic Problems in Non-Smooth Domains*, Pitman, 1985.

[25] Jamet, P., Estimations d'erreur pour des éléments finis droits presque dégénérés, *RAIRO Anal. Numér*, **10**, 46-61, 1976.

[26] Krízek, M., On the maximum angle condition for linear tetrahedral elements, *SIAM J. Numer. Anal.*, **29**, 513-520, 1992.

[27] Nedelec, J.C., Mixed finite elements in $\mathbb{R}^3$, *Numer. Math.*, **35**, 315-341, 1980.

[28] Nedelec, J.C., A new family of mixed finite elements in $\mathbb{R}^3$, *Numer. Math.*, **50**, 57-81, 1986.

[29] Raviart, P.A., Thomas, J.M., A mixed finite element method for second order elliptic problems, *Mathematical Aspects of the Finite Element Method* (I. Galligani, E. Magenes, eds.), Lectures Notes in Math. 606, Springer Verlag, 1977.

[30] Roberts, J.E.,Thomas, J.M., Mixed and Hybrid Methods, in *Handbook of Numerical Analysis*, (P.G. Ciarlet and J.L. Lions,eds.), *Vol.II, Finite Element Methods (Part 1)*, North Holland, 1989.

[31] Shenk, N. Al., Uniform error estimates for certain narrow Lagrange finite elements, *Math. Comp.*, **63**, 105-119, 1994.

[32] Strang, G., Fix, G.J., *An Analysis of The Finite Element Method*, Prentice Hall, 1973.

[33] Thomas, J.M., Sur l'analyse numérique des méthodes d'éléments finis hybrides et mixtes, Thèse d'Etat, Université Pierre et Marie Curie, Paris, 1977.

[34] Yosida, K., *Functional Analysis*, Spinger Verlag, 1966.