AN ANALYSIS OF NEW MIXED FINITE ELEMENTS FOR THE APPROXIMATION OF WAVE PROPAGATION PROBLEMS

E. BÉCACHE[†¶], P. JOLY[‡¶], AND C. TSOGKA[§¶]

Abstract. We construct and analyse a new family of rectangular (2D) or cubic (3D) mixed finite elements for the approximation of the acoustic wave equations. The main advantage of this element is that it permits us to obtain through mass lumping an explicit scheme even in an anisotropic medium. Non-classical error estimates are given for this new element.

Key words. mixed finite elements, mass lumping, anisotropic waves

AMS subject classifications. 65M60, 65M12, 65M15, 65C20, 73D25

1. Introduction. In this paper we develop and analyze a new family of mixed finite elements for the scalar anisotropic wave equation. It is essential, for the applications we have in mind, to use the mixed formulation involving both velocity and pressure and not the traditional displacement formulation. Let us first explain our motivations for introducing such mixed formulation.

This work falls within the more general framework of developing efficient numerical methods for approximating wave propagation in complex media such as anisotropic heterogeneous media with cracks or obstacles of arbitrary shapes. As a first step towards the more complicated case of elastic waves with a free boundary condition (on free surface or crack), we consider the scalar anisotropic wave equation, with a Neumann boundary condition on the boundary Γ of an obstacle (or crack).

The characteristics of wave propagation problems (large scale computations, complex geometries for the cracks or the obstacles, unbounded domains) together with the need to construct an efficient method lead us to adopt the following criteria.

To facilitate the implementation and reduce the calculation time, we use regular meshes, squares in 2D and cubes in 3D. The question with such a choice is how to take into account in an accurate way complex geometries. A staircase approximation of the geometry introduces artificial singularities on the boundary which produce spurious diffractions (see [12] for a numerical illustration of this phenomenon). To reduce these spurious diffractions a very fine mesh in space is required. Moreover, because of the CFL stability condition, the time discretisation step is then very small for an explicit scheme.

To avoid this, we intend to use the fictitious domain method (cf. [12, 18, 22]). The basic principle of this method is to extend the solution to a domain of simple geometry that ignores the obstacle or the crack (typically a rectangle in 2D), the boundary condition being seen as an essential condition, i.e. an equality constraint in a given functional space. To apply this method to the wave equation with a Neumann boundary condition, we need to rewrite the Neumann condition as an essential one. This can be achieved by introducing the unknown $p = \nabla u$, where u is the solution of the scalar wave equation. The Neumann condition on u can then be written as an essential condition on p. This justifies our choice to consider the wave equation as a first order system in $(v = \partial_t u, p)$. The boundary condition is then taken into account via the introduction of a Lagrange multiplier, which can be interpreted as the jump of v through the boundary. We have in that case two computational domains, a domain of simple geometry in which a uniform mesh can be used for the unknowns (v, p) and the boundary in which a conforming mesh is considered for the Lagrange multiplier.

[†](eliane.becache@inria.fr)
[‡](patrick.joly@inria.fr)
[§](chrysoula.tsogka@inria.fr)
[¶]INRIA, Domaine de Voluceau-Rocquencourt, BP 105, F-78153 Le Chesnay Cédex

This is our main reason to use a mixed variational formulation. However in this paper, we will not consider the presence of the crack, and we refer for that to a future paper.

A second advantage of the first order system is that it permits the use of a new absorbing layer model, inspired by the Perfectly Matched Layer (P.M.L) introduced by Bérenger for the 2D Maxwell problem (c.f [4, 11, 13]). This model is very efficient compared with the traditional absorbing boundary conditions and can be naturally extended to any first order system, [14].

Finally in order to get an efficient method we wish to obtain an explicit time discretisation scheme through mass lumping. This is why we consider in this paper non-standard mixed finite elements (inspired by Nédélec's second family [19]) which we shall show to be compatible with mass lumping. These elements are in fact the scalar version of a new family of mixed finite elements we have introduced in [2] for the velocity-stress formulation of elastodynamics.

We construct arbitrary higher order elements which are of great importance in the case of large scale problems. In this sense, this work can be considered as the continuation of many articles about higher order methods for wave propagation problems (c.f [23, 8, 9, 10, 26, 24]).

The main objective of the present paper is the analysis of these new elements. We shall show that they do not satisfy the classical Ladyzhenskaya-Babuska-Brezzi approximation condition (c.f [6, 5, 1]) (defect of coerciveness). Similar difficulties have been raised by Nédélec in [19] (p 80). Our main contribution is to present a method for overcoming this difficulty. The abstract results concerning the elliptic problem have been announced in a short note [3]. Let us mention that the generalization of this finite element analysis to the case of elastic waves presents a second major difficulty linked to the symmetry condition of the stress tensor. We intend to treat this difficulty in a forthcoming paper.

The present paper is organized as follows. In §2 we recall why the classical $RT_{[k]}$ mixed finite elements introduced by Raviart and Thomas in [21] do not lead to an explicit scheme in the case of an anisotropic medium and we propose instead the use of new mixed finite elements. The lowest order element is presented in §2.2, the higher orders in §2.3. Section 3 is the main part of the paper and is concerned with the mixed approximation of the elliptic problem which is nothing but the stationary problem associated to the evolution problem presented in §2. This analysis will be used to study an elliptic projection operator that will be useful for the analysis of the approximation of the evolution problem. We shall show in §3.1 why the analysis of the new element does not fit the classical theory. That is why we develop in §3.2 a novel abstract theory leading to non-classical error estimates. In §3.3, we show that the mixed finite elements that we have constructed enter this framework and error estimates are given. Section 4 is devoted to relating the error estimates on the time domain solution to the error estimates obtained in the previous section on the elliptic problem. This part of the analysis is inspired by [15] and essentially relies on energy estimates.

Although in this paper we restrict ourselves to the 2D case, the extension of this work to 3D is straightforward.

2. Presentation of the new mixed finite elements.

2.1. The model problem : the anisotropic wave equation. Let Ω be a bounded domain of $\mathbb{R}^2$ and let $A(x)$ be a positive definite symmetric matrix satisfying

$$A(x)\,\xi\cdot\xi \geq \alpha\,|\xi|^2\,,\ \alpha>0,\ \forall\xi\in\mathbb{R}^2,\ \text{a.e. } x\in\Omega. \tag{2.1}$$

In practice, Ω will be the simple geometrical domain (let us say a rectangle) we referred to in the introduction. We consider the scalar evolution problem

$$\begin{cases} \text{Find } u\ : [0,T]\longmapsto H_0^1(\Omega) \text{ such that} \\ \\ \dfrac{\partial^2 u}{\partial t^2} - \text{div }(A^{-1}(x)\nabla u) = f,\ f\in C^0(0,T;L^2(\Omega)), \end{cases} \tag{2.2}$$

subject to the initial conditions

$$u(t=0)=u_0\in H^1(\Omega)\ ;\ \frac{\partial u}{\partial t}(t=0)=u_1\in L^2(\Omega).$$

Remark 2.1. *In this section we are not interested in the regularity of u with respect to time variable t. This is why we simply write $u\ : [0,T]\longmapsto H_0^1(\Omega)$.*

Remark 2.2. *The homogeneous Dirichlet condition on $\partial\Omega$ has just been considered for simplicity. This has nothing to do with a Neumann condition on the boundary of an obstacle interior to Ω that will be taken into account with the ficitious domain method as explained in the introduction.*

Let

$$p = A^{-1}(x)\nabla u \ \text{ and } v=\frac{\partial u}{\partial t}.$$

Substituting into (2.2) yields

$$\begin{cases} \dfrac{\partial v}{\partial t} - \text{div } p = f, \\ \\ A\dfrac{\partial p}{\partial t} - \nabla v = 0, \end{cases} \tag{2.3}$$

with initial conditions

$$p(0)=p_0=A^{-1}(x)\nabla u_0\ ;\ v(0)=v_0=u_1. \tag{2.4}$$

Remark 2.3. *Throughout this paper, the generic point of $\mathbb{R}^2$ will be indifferently denoted by the letter x or the couple (x,y) (depending on the context).*

A mixed formulation associated to equations (2.3) is given by the following problem

$$\begin{cases} \text{Find } (p,v) : [0,T]\longmapsto X\times M\equiv H(\text{div }; \Omega)\times L^2(\Omega) \text{ such that} \\ \\ \dfrac{d}{dt}a(p,q)+b(v,q) \quad = 0, \qquad \forall q\in X, \\ \\ \dfrac{d}{dt}(v,w)-b(w,p) \quad = (f,w), \quad \forall w\in M, \end{cases} \tag{2.5}$$

where

$$
\left\{
\begin{aligned}
& a(p,q) = \int_\Omega A(x)\ p\cdot q\ dx, && \forall (p,q)\in X\times X,\\
& b(w,q) = \int_\Omega w\ \ \mathrm{div}\ q\ dx, && \forall (w,q)\in M\times X,\\
& (f,w) = \int_\Omega f\ w\ dx, && \forall w\in M.
\end{aligned}
\right.
\tag{2.6}
$$

The bilinear form $a(\cdot,\cdot)$ (resp. $b(\cdot,\cdot)$) is continuous on $H\times H$ ($H=(L^2(\Omega))^2$) (resp. on $X\times M$). The bilinear form $a(\cdot,\cdot)$ (resp. $b(\cdot,\cdot)$) thus defines a linear continuous operator $\mathcal{A}: H\to H'$ by $\langle \mathcal{A}p,q\rangle_{H'\times H} = a(p,q)$ (resp. $B: X\to M'$ by $\langle Bq,w\rangle_{M'\times M} = b(w,q)$). They satisfy the following properties (see for instance [6]) :

$$
\left|
\begin{array}{ll}
(i) & \text{The continuous inf-sup condition:}\\[1ex]
& \exists c>0\,/\,\forall\ w\in M,\ \exists\ q\in X\ /\ b(w,q)\geq c\,\|w\|_M\,\|q\|_X\,.\\[1ex]
(ii) & \text{The coercivity of the form } a(\cdot,\cdot) \text{ on } V\equiv \mathrm{Ker}B:\\[1ex]
& \exists \alpha>0\,/\,\forall\ p\in V,\quad a(p,p)\geq \alpha\,\|p\|_X^2\,.
\end{array}
\right.
\tag{2.7}
$$

In the following, we consider only the semi-discretisation in space of this problem, but with the specific objective of obtaining a discretisation for which mass lumping is possible.

2.2. Presentation of the $Q_1^{div}-Q_0$ finite element in the lowest order. We suppose now that Ω is a union of rectangles in such a way that we can consider a regular mesh $(\mathcal{T}_h)$ with square elements K of edge $h>0$. We introduce the following approximation spaces

$$
\left\{
\begin{aligned}
& X_h = \left\{ q_h\in X\ /\ \forall K\in\mathcal{T}_h,\ q_h|_K\in\hat{X}\right\},\\
& M_h = \left\{ w_h\in M\ /\ \ \forall K\in\mathcal{T}_h,\ w_h|_K\in\hat{M}\right\},
\end{aligned}
\right.
\tag{2.8}
$$

where $\hat{X}$ (resp. $\hat{M}$) denotes a finite dimensional space of vector (resp. scalar) functions. The discrete problem associated to (2.5) and (2.4) is

$$
\left\{
\begin{array}{lll}
\text{Find } (p_h,v_h):[0,T]\to X_h\times M_h \text{ such that} & & \\[1ex]
\dfrac{d}{dt}a(p_h,q_h)+b(v_h,q_h) & =0, & \forall q_h\in X_h,\\[1ex]
\dfrac{d}{dt}(v_h,w_h)-b(w_h,p_h) & =(f,w_h), & \forall w_h\in M_h,
\end{array}
\right.
\tag{2.9}
$$

subject to the initial conditions

$$
p_h(0)=p_{0,h}\ ;\ v_h(0)=v_{1,h}.
$$

The usual choice consists in taking for $\hat{X}$ the lowest order Raviart-Thomas element, see [21]:

$$
\hat{X} = RT_{[0]} = P_{10}\times P_{01},
\tag{2.10}
$$

and for $\hat{M}$ the piecewise constants:

$$\hat{M} = Q_0. \tag{2.11}$$

Here, P_k is the space of polynomials of degree not greater than k and P_{kl} is defined by

$$P_{kl} = \left\{ p(x,y) \;\mid\; p(x,y) = \sum_{i \le k, j \le l} a_{ij} x^i y^j \right\}.$$

This choice however does not lead to an explicit scheme when one considers the evolution problem of anisotropic waves corresponding to (2.9). We introduce here $B_{N_1} = \{\tau_i\}_{i=1}^{N_1}$, $B_{N_2} = \{\phi_i\}_{i=1}^{N_2}$ the bases of X_h and M_h respectively, where $N_1 = \dim X_h$ and $N_2 = \dim M_h$. We denote then by $[P] = (P_1, ..., P_{N_1})$ and $[V] = (V_1, ..., V_{N_2})$ the coordinates of p_h and v_h in the bases B_{N_1} and B_{N_2}. In these bases, problem (2.9) can be written in the following form

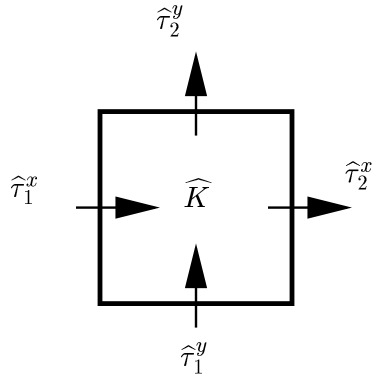
$$\left\{ \begin{array}{l} \text{Find } (P,V) : [0,T] \longmapsto \mathbb{R}^{N_1} \times \mathbb{R}^{N_2} \text{ such that} \\[2ex] M_p \dfrac{dP}{dt} + C^T V = 0, \\[2ex] M_v \dfrac{dV}{dt} - CP = F, \\[2ex] + \text{ initial conditions,} \end{array} \right.$$

with

$$\begin{array}{lll} (i) & (M_p)_{i,j} = (A\tau_i, \tau_j)_{(L^2(\Omega))^2}\,, & 1 \le i,j \le N_1, \\[1ex] (ii) & (M_v)_{i,j} = (\phi_i, \phi_j)_{L^2(\Omega)}\,, & 1 \le i,j \le N_2, \\[1ex] (iii) & (C)_{i,j} = (\phi_i, \operatorname{div} \tau_j)_{L^2(\Omega)}\,, & 1 \le i \le N_2,\ 1 \le j \le N_1, \\[1ex] (iv) & (F)_j = (f, \phi_j)_{L^2(\Omega)} & 1 \le j \le N_2\ ; \end{array} \tag{2.12}$$

C^T denotes the transpose of C. If we use a centered finite difference approximation for the time discretisation, the solution at each time step is obtained by inverting the mass matrices M_v and M_p. Although they are symmetric and sparse, this inversion can become costly (for large systems). To avoid that, we want to reduce them to diagonal (or block diagonal) matrices by using a mass lumping technique (see [10, 25]). This consists in using adequate quadrature formulas to approximate the integrals in (2.12 (i) and (ii)). One can remark that, M_h being chosen discontinuous, the matrix M_v does not really need to be mass lumped (it is already diagonal here); so, we focus our attention on the mass lumping of M_p.

Let now (τ_I) be the $RT_{[0]}$ global basis functions. In Fig. 2.1, we represent the 4 local basis vector functions in $RT_{[0]}(\widehat{K})$, $\widehat{K}$ being the reference element. Each global basis function τ_I is associated to an edge of the mesh and is supported by the two elements adjacent to this edge, see Fig. 2.2. The restriction of τ_I on one element K of its support, corresponds to a local basis function, denoted $\widehat{\tau}_i$, on the reference element $\widehat{K}$. In fact $\widehat{\tau}_i$ corresponds to $\widehat{\tau}_i^x$ (resp. $\widehat{\tau}_i^y$) when τ_I is associated to a vertical (resp. horizontal) edge of the mesh. Now consider the integral



$$\widehat{\tau}_1^y = \begin{pmatrix} 0 \\ 1-y \end{pmatrix}, \; \widehat{\tau}_2^y = \begin{pmatrix} 0 \\ y \end{pmatrix}$$

$$\widehat{\tau}_1^x = \begin{pmatrix} 1-x \\ 0 \end{pmatrix}, \; \widehat{\tau}_2^x = \begin{pmatrix} x \\ 0 \end{pmatrix}$$

FIG. 2.1. *Local vector basis functions in* $RT_{[0]}(\widehat{K})$

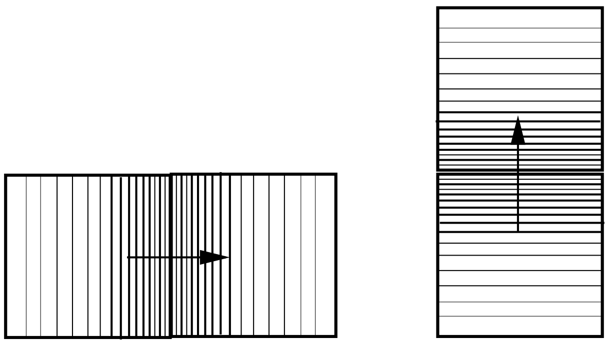


FIG. 2.2. *Support of the basis functions in* $RT_{[0]}$

$$a(\tau_I, \tau_J) = \sum_{K \in \mathcal{T}_h} \int_K A\tau_I|_K \cdot \tau_J|_K dx = \sum_{K \in \mathcal{T}_h} \text{mes}(K) \int_{\widehat{K}} A\widehat{\tau}_i \cdot \widehat{\tau}_j d\widehat{x} \tag{2.13}$$

If we use the quadrature formula

$$\int_{\widehat{K}} \widehat{f}(\widehat{x}) d\widehat{x} = \sum_{q=1}^{N} \omega_q \widehat{f}(\widehat{M}_q),$$

with $(\widehat{M}_q)_{q=1,\ldots,N}$ being the quadrature points and $(\omega_q)_{q=1,\ldots,N}$ the associated weights to approximate (2.13), we obtain

$$\int_{\widehat{K}} A\widehat{\tau}_i \cdot \widehat{\tau}_j d\widehat{x} \approx \sum_q \omega_q A\widehat{\tau}_i(\widehat{M}_q) \cdot \widehat{\tau}_j(\widehat{M}_q).$$

This would lead to a diagonal matrix if

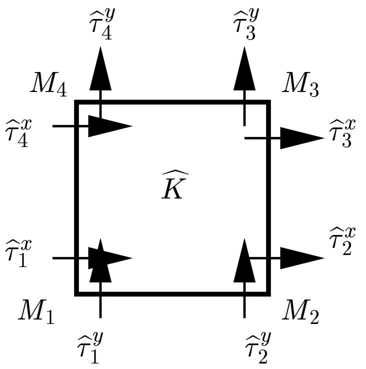
$$A\widehat{\tau}_i(\widehat{M}_q) \cdot \widehat{\tau}_j(\widehat{M}_q) = 0, \quad \forall\, i \neq j. \tag{2.14}$$

For the isotropic wave equation ($A = Id$), it is easy to find a quadrature formula satisfying (2.14), take for instance the lowest order Gauss-Lobatto quadrature formula using the vertices of the element $\widehat{K}$ as quadrature points. But for the anisotropic wave equation, there is no usual quadrature formula that satisfies (2.14). The difficulty is that for two orthogonal edges of the element $\widehat{K}$, the functions $A\widehat{\tau}_1^y$ and $\widehat{\tau}_1^x$ are no longer orthogonal.

The alternative solution that we propose consists in changing the approximation space $\hat{X}$. Let

$$\hat{X} = Q_1 \times Q_1 \tag{2.15}$$

where Q_1 is the space of piecewise bilinear functions. This element was initially introduced by Nédélec in [19]. In Fig. (2.3) we represent the 8 local basis vector functions on the reference element $\widehat{K}$. We will call this choice (2.15) combined with (2.11) the $Q_1^{div} - Q_0$ element.



$$\widehat{\tau}_4^y = (0,\ (1-x)y)^t\,,\ \widehat{\tau}_3^y = (0,\ xy)^t$$

$$\widehat{\tau}_4^x = ((1-x)y,\ 0)^t\,,\ \widehat{\tau}_3^x = (xy,\ 0)^t$$

$$\widehat{\tau}_1^x = ((1-x)(1-y),\ 0)^t\,,\ \widehat{\tau}_2^x = (x(1-y), 0)^t$$

$$\widehat{\tau}_1^y = (0,\ (1-x)(1-y))^t\,,\ \widehat{\tau}_2^y = (0,\ x(1-y))^t$$

FIG. 2.3. *The local basis functions in* $\hat{X} = Q_1 \times Q_1$

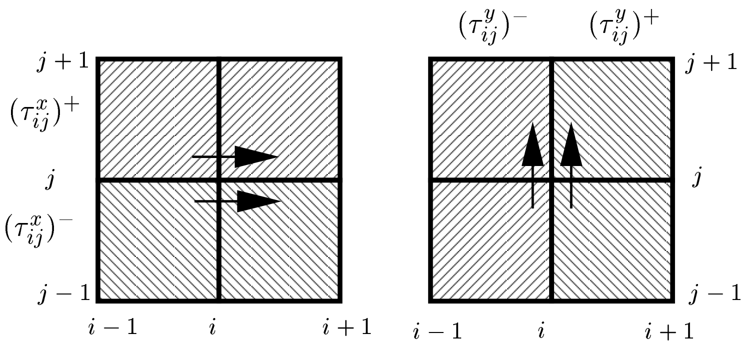


FIG. 2.4. *The* $Q_1 \times Q_1$ *element*

The lowest order Gauss-Lobatto quadrature formula satisfies now (2.14) with the new basis functions; the key point is that the 4 quadrature points M_j coincide with the locations in space of the degrees of freedom and the new basis functions satisfy

$$\tau_i^x(M_j) = \delta_{ij}(1,\ 0)^t, \quad \tau_i^y(M_j) = \delta_{ij}(0,\ 1)^t.$$

Note that, over each element, there are 4 quadrature points but 8 degrees of freedom. For more details on quadrature formulae and mass lumping techniques we refer to the work of G. Cohen, P. Joly and N. Tordjman for the acoustic wave equation ([25, 9, 8]) and to P. Monk, G. Cohen [10] and A. Elmkies [16] for Maxwell's equations.

REMARK 2.4. *With Gauss-Lobatto's rule, and the $Q_1 \times Q_1$ element presented above, the mass matrix $a(p_h, q_h)$ is block diagonal, the size of each block being equal to the number of degrees of freedom at each vertex, which is 4, see Fig. 2.4, thus after an inversion of local 4×4 matrices, we obtain an explicit scheme. In the particular case of the isotropic wave equation, $A = Id$, the mass matrix becomes 2×2 block diagonal, because the basis functions τ_i^x and τ_j^y are orthogonals.*

REMARK 2.5. *From the physical point of view, having 4 degrees of freedom for p at each vertex is coherent with the actual discontinuities of p_x or p_y that occur when A is dicontinuous accross the edges of the mesh.*

Before analyzing this element, let us show how it can easily be extended to higher order.

2.3. Extension to higher order and mass lumping. The natural generalization of the lowest order element, presented in the previous section, consists in taking

$$\begin{cases} X_h = \{q_h \in X \;/\; \forall K \in \mathcal{T}_h,\; q_h|_K \in Q_{k+1} \times Q_{k+1}\}, \\ \\ M_h = \{w_h \in M \;/\; \forall K \in \mathcal{T}_h,\; w_h|_K \in Q_k\}, \end{cases} \tag{2.16}$$

and we will call it the $Q_{k+1}^{div} - Q_k$ element. This choice still satisfies our requirement with respect to mass lumping. We represent in figures 2.5 and 2.6 the degrees of freedom for $k = 1$ and $k = 2$, the exact locations of the quadrature points and the associated weights are given in appendix A. We also indicate in these figures the number of degrees of freedom per node. The locations of the degrees of freedom correspond to tensor products of 1D quadrature points associated to Gauss-Lobatto (for p_h) or Gauss-Legendre (for v_h) quadrature formulas.

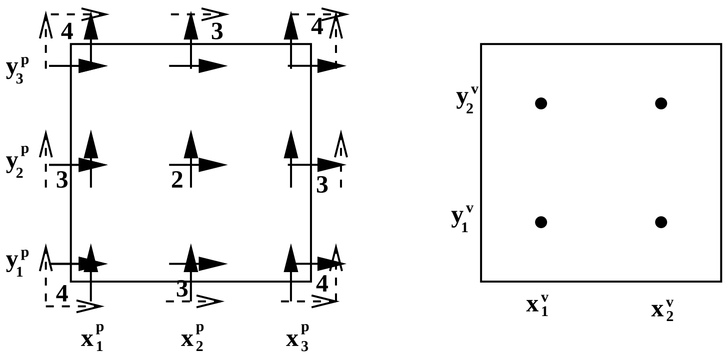


FIG. 2.5. *Degrees of freedom in the $Q_2^{div} - Q_1$ element*

Indeed, following the approach of P. Monk, G. Cohen [10] and N. Tordjman [25], we obtain mass-lumping by approximating the integrals in (2.12(i), (ii)) with adequate quadrature formulas, for which the position of the quadrature points coincide with the position of the degrees of freedom. For our elements they consist in

- using the Gauss-Lobatto quadrature formulas for the approximation of the M_p matrix. The resulting matrix is now block diagonal. Each block is associated to one quadrature point and its dimension is equal to the number of degrees of

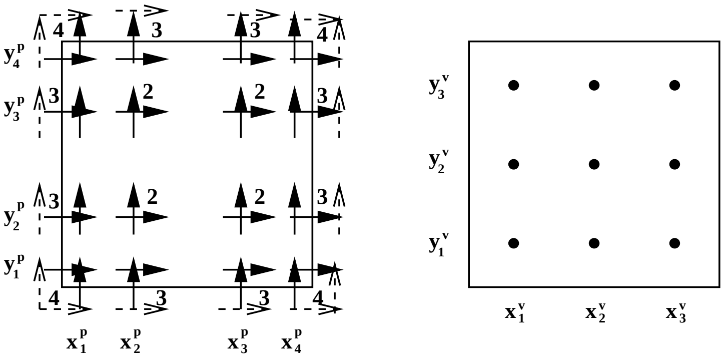


FIG. 2.6. *Degrees of freedom in the $Q_3^{div} - Q_2$ element*

freedom at this point (the "worst" case concerns the vertices of the elements K, where the dimension of the local block is 4×4).

- using the Gauss-Legendre quadrature formula for the approximation of the M_v matrix, so that the resulting matrix is diagonal.

REMARK 2.6. *The Gauss-Legendre quadrature formula is exact for the integration of the product $v_h w_h$ which is in Q_{2k}. On the other hand, considering for simplicity the case where A is constant, the product of $Ap_h.q_h$ belongs to Q_{2k+2}. However, following [7, 25, 17] one can prove that one does not lose any order of accuracy provided that the quadrature formula is exact in Q_k which is the case with the Gauss-Lobatto formula we use.*

REMARK 2.7. *The Raviart-Thomas k-th order approximation consists in the choice*

$$\hat{X}^{RT} = RT_{[k]} = P_{k+1,k} \times P_{k,k+1} \quad and \quad \hat{M} = Q_k.$$

It is clear that the new approximate space X_h contains the space X_h^{RT}. Thus, we have enriched the p-approximation, while keeping the same approximation space for v.

3. Analysis of the new mixed finite element for an elliptic problem.

Following [5], we will study in this section the mixed approximation of the elliptic problem which is in fact the stationary problem associated to the evolution problem (2.5). Actually, we give in §3.2 an abstract result for a class of elliptic problems posed in a more general framework and show in §3.3 that the model problem enters this framework. This analysis will be used in §3.4 to study an elliptic projection operator that will be useful for the analysis of the approximation of the evolution problem done in §4.

3.1. The elliptic problem.

The elliptic problem we consider here is

$$\left\{ \begin{array}{l} \text{Find } u \in H_0^1(\Omega) \text{ such that} \\ \\ -\text{div } (A^{-1}(x)\nabla u) = f, \;\; f \in L^2(\Omega), \end{array} \right. \tag{3.1}$$

We know that (3.1) admits a unique solution $u \in H_0^1(\Omega)$ and that there exists $c > 0$ such that

$$\|u\|_{H^1(\Omega)} \leq \; c \; \|f\|_{L^2(\Omega)} . \tag{3.2}$$

As for the time dependent problem, we set

$$p = A^{-1}(x)\nabla u, \tag{3.3}$$

and this gives

$$-\text{div } p = f. \tag{3.4}$$

The mixed formulation associated to equations (3.3) and (3.4) is

$$\left\{ \begin{array}{lll} \text{Find } (p,u) \in X \times M = H(\text{div }; \Omega) \times L^2(\Omega) \text{ such that} & & \\ a(p,q) + \quad b(u,q) & = 0, & \forall q \in X, \\ \qquad\qquad b(w,p) & = -(f,w), & \forall w \in M, \end{array} \right. \tag{3.5}$$

where $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ are defined by (2.6) and satisfy properties (2.7). As proven in [6], there exists a unique solution (p,u) in $X \times M$ to problem (3.5), where u is also the solution of the initial problem (3.1) (in fact the abstract result yields the uniqueness of u only in $M/$ Ker B^t, but it is easy to check that for the divergence operator Ker $B^t = \{0\}$). For the approximation of this problem, we again consider the finite dimensional spaces X_h and M_h defined by (2.8). The discrete problem associated to (3.5) is

$$\left\{ \begin{array}{lll} \text{Find } (p_h,u_h) \in X_h \times M_h \text{ such that} & & \\ a(p_h,q_h) + \quad b(u_h,q_h) & = 0, & \forall q_h \in X_h, \\ \qquad\qquad b(w_h,p_h) & = -(f,w_h), & \forall w_h \in M_h. \end{array} \right. \tag{3.6}$$

The elliptic problem (3.5) and its approximation (3.6) have been studied by several authors, see for example [21, 6], and we know that it admits a unique solution (p_h,u_h) in $X_h \times M_h$ with the convergence property

$$(p_h,u_h) \longrightarrow (p,u) \quad \in \quad X \times M, \tag{3.7}$$

when the following assumptions are satisfied

$$\left| \begin{array}{ll} (i) & \text{The uniform discrete inf-sup condition:} \\ & \exists c > 0 \text{ independent of } h \text{ such that} \\ & \forall\ w_h \in M_h,\ \exists\ q_h \in X_h\ /\ b(w_h,q_h) \geq c\,\|w_h\|_M\,\|q_h\|_X\,; \\ (ii) & \text{The coercivity of the form } a(\cdot,\cdot) \text{ on } V_h: \\ & \exists \alpha > 0 \text{ independent of } h \text{ such that} \\ & \forall\ p_h \in V_h, \quad a(p_h,p_h) \geq \alpha\,\|p_h\|_X^2\ , \\ & \text{where } V_h = \text{Ker}B_h = \{q_h \in X_h\ /\ b(w_h,q_h) = 0,\ \forall\ w_h \in M_h\} \end{array} \right. \tag{3.8}$$

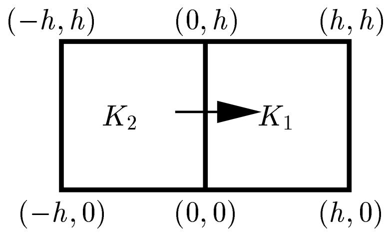


FIG. 3.1. *The function* $\vec{f}_h$

These assumptions are satisfied by the couple of spaces (X_h, M_h) defined in (2.8), $\hat{X}$ being the lowest order Raviart Thomas element (2.10) and $\hat{M}$ the piecewise constants (2.11), see [21]. With our new choice for $\hat{X}$ it is easy to verify that property (3.8-i) is still true but we no longer satisfy (3.8-ii). Indeed, consider the function $\vec{f}_h$ given by

$$\vec{f}_h\,|_{K_1} = \begin{pmatrix} (1-\dfrac{x}{h})(1-2\dfrac{y}{h}) \\ 0 \end{pmatrix} \text{ and } \vec{f}_h\,|_{K_2} = \begin{pmatrix} (1+\dfrac{x}{h})(1-2\dfrac{y}{h}) \\ 0 \end{pmatrix}.$$

We can easily see that $\vec{f}_h \in V_h$ (if $\hat{M} = Q_0$) and

$$a(f_h, f_h) \underset{h\to 0}{\longrightarrow} 0\ ,$$

$$\|f_h\|_X \underset{h\to 0}{\longrightarrow} \frac{2}{3},$$

so that we can not expect to verify (3.8-ii). However, we will prove (see §3) that this choice gives a good approximate solution and we will show a new convergence result.

REMARK 3.1. *In order to preserve (3.8-ii) one could change the approximation space $\hat{M}$. The natural choice is*

$$\hat{M} = P_1,$$

the key point being that the divergence operator sends $Q_1 \times Q_1$ into P_1. We have eliminated this choice because it is rather expensive in terms of computational time and memory requirements.*

3.2. An abstract result. The first point in our new theory is that we need to introduce a third Hilbert space. More precisely, let M, X, H be three Hilbert spaces with

$$X \subset H, \quad X \text{ dense in } H \quad \text{and} \quad |\cdot|_H \leq \|\cdot\|_X\,. \tag{3.9}$$

The reader can have in mind that for our application we shall have

$$H = (L^2(\Omega))^2,\ X = H(\mathrm{div};\Omega) \text{ and } M = L^2(\Omega).$$

Let $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ be two continuous bilinear forms in $H \times H$ and $M \times X$. In the same way as in §2.1, the bilinear form $a(\cdot,\cdot)$ defines an operator $\mathcal{A}$ in $\mathcal{L}(H)$, such

*which implies $V_h = \{q_h \in X_h\ /\ \mathrm{div}\ q_h = 0\} \subset \mathrm{Ker}\ B$, so that (3.8-ii) is a consequence of (2.7-ii) that $a(p,q) = (\mathcal{A}p,q)_H$ for all $(p,q) \in H \times H$ and the bilinear form $b(\cdot,\cdot)$ defines an operator $B : X \to M'$ (and its transpose $B^t : M \to X'$) such that

$$\langle Bp, w\rangle_{M' \times M} = \langle p, B^t w\rangle_{X \times X'} = b(w,p), \quad \forall (p,w) \in X \times M.$$

The kernels of B and B^t are defined as follows

$$\left\{ \begin{array}{l} V \equiv \text{ Ker } B = \{p \in X \ / \ b(w,p) = 0, \ \forall w \in M\}, \\ \\ \text{Ker } B^t = \{w \in M \ / \ b(w,p) = 0, \ \forall p \in X\}. \end{array} \right.$$

We make the assumptions:

$$(3.10) \quad \left| \begin{array}{ll} (i) & \exists\, c > 0 \ / \ \forall\, w \in M, \ \exists\, q \in X \ / \ b(w,q) \geq c\, \|w\|_{M/\text{Ker } B^t} \|q\|_X, \\ \\ (ii) & \exists\, \alpha > 0 \ / \ \forall\, p \in V, \quad a(p,p) \geq \alpha\, \|p\|_X^2. \end{array} \right.$$

where the norm in the quotient space is defined by $\|w\|_{M/\text{Ker } B^t} = \inf_{w_0 \in \text{Ker } B^t} \|w + w_0\|_M$. We shall identify the quotient space $M/\text{Ker } B^t$ with the orthogonal complement of Ker B^t

$$M/\text{Ker } B^t \equiv (\text{Ker } B^t)^\perp \equiv \left\{ w \in M \ / \ (v,w)_M = 0, \forall\, v \in \text{Ker } B^t \right\}.$$

We are interested in the numerical approximation of the solution (p,u) to the following problem

$$(3.11) \qquad \left\{ \begin{array}{llll} \text{Find } (p,u) \in X \times M \text{ such that} & & & \\ \\ a(p,q) + & b(u,q) & = 0, & \forall q \in X, \\ \\ & b(w,p) & = -\langle f, w\rangle, & \forall w \in M, \end{array} \right.$$

with $f \in M'$, the dual space of M. Under these assumptions, we have the following classical result (see [6])

THEOREM 3.1. *For all $f \in \text{Im } B$, problem (3.11) has a unique solution (p,u) in $X \times (M/\text{Ker } B^t)$. Moreover,*

$$\|u\|_{M/\text{Ker } B^t} + \|p\|_X \leq C\, \|f\|_{M'}.$$

Suppose $X_h \subset X$ and $M_h \subset M$ to be finite dimensional approximation spaces. We consider then the approximate problem

$$(3.12) \qquad \left\{ \begin{array}{llll} \text{Find } (p_h,u_h) \in X_h \times M_h \text{ such that} & & & \\ \\ a(p_h,q_h) + & b(u_h,q_h) & = 0, & \forall q_h \in X_h, \\ \\ & b(w_h,p_h) & = -\langle f, w_h\rangle, & \forall w_h \in M_h. \end{array} \right.$$

We set

$$(3.13) \quad \left\{ \begin{array}{l} V_h(f) = \{q_h \in X_h \ / \ b(w_h,q_h) = -\langle f, w_h\rangle, \ \forall w_h \in M_h\} = B_h^{-1}(f), \\ \\ V_h = V_h(0) = \text{Ker } B_h, \ \text{(see below for the meaning of } B_h) \end{array} \right.$$

and make the following hypotheses

(H0) $\forall f \in \text{Im } B,\ V_h(f) \neq \emptyset$.

(H1) Orthogonal decomposition of X_h:

$$\left| \begin{array}{l} X_h = X_h^s \oplus X_h^r \quad (p_h = p_h^s + p_h^r)\ , \quad X_h^r \subset V_h; \\[2mm] \forall (p_h^s, p_h^r) \in X_h^s \times X_h^r \quad (p_h^s, p_h^r)_H = 0. \end{array} \right.$$

(H2) "Strong" discrete uniform inf-sup condition:

there exists a constant $c > 0$, independent of h, such that

$$\forall\, w_h \in M_h,\ \ \exists\ q_h^s \in X_h^s \quad / \quad b(w_h, q_h^s) \geq c\, \|w_h\|_{M/\text{Ker } B_h^t}\, \|q_h^s\|_X\,.$$

(H3) "Weak" coercivity:

there exists a constant $\alpha > 0$, independent of h, such that

$$\forall\ p_h \in V_h, \quad a(p_h, p_h) \geq \alpha \left(\|p_h^s\|_X^2 + |p_h^r|_H^2 \right).$$

(H4) Approximation properties

$$\left\{ \begin{array}{ll} \lim\limits_{h\to 0} \inf\limits_{q_h^s \in X_h^s} \|p - q_h^s\|_X = 0, & \forall p \in X, \\[2mm] \lim\limits_{h\to 0} \inf\limits_{w_h \in M_h} \|u - w_h\|_M = 0, & \forall u \in M. \end{array} \right.$$

REMARK 3.2. *Note that, since X is dense in H, (H4) implies that:*

$$\lim_{h\to 0} \inf_{q_h^s \in X_h^s} |p - q_h^s|_H = 0, \quad \forall p \in H.$$

We can introduce, as for the continuous problem, an operator B_h from X_h to M_h defined by

$$(B_h p_h, w_h)_{M'\times M} = \left(p_h, B_h^t w_h\right)_{X\times X'} = b(p_h, w_h), \quad \forall p_h \in X_h,\ \forall w_h \in M_h,$$

and the kernels of B_h (see (3.13)) and of its transpose

$$\text{Ker} B_h^t = \{ w_h \in M_h \ /\ b(p_h, w_h) = 0,\ \forall p_h \in X_h \}\,.$$

REMARK 3.3. *It may be more convenient to characterize hypothesis (H0) with one of the following equivalent statements*

$(H0)-(ii)$ $\quad \forall p \in X,\ \exists p_h \in X_h$ such that $b(p - p_h, w_h) = 0,\ \ \forall w_h \in M_h$;

$(H0)-(iii)$ $\quad \text{Ker } B_h^t = \text{Ker } B^t \cap M_h \subset \text{Ker } B^t$.

REMARK 3.4. *We call hypothesis (H2) "strong" as we suppose the existence of a q_h^s in X_h^s instead of X_h, which would be the classical assumption; more precisely $X_h^s \subsetneq X_h$. On the contrary, hypothesis (H3) is "weaker" in the sense that for all $p_h \in X_h^s$ or $p_h \in X_h^r$,*

$$\|p_h^s\|_X^2 + |p_h^r|_H^2 \leq \|p_h\|_X^2\,.$$

THEOREM 3.2. *Under hypotheses (H0) - (H4), problem (3.12) admits a unique solution such that*

$$(p_h = p_h^s + p_h^r, u_h) \in X_h \times (M_h/\mathrm{Ker}\ B_h^t),$$

and the following convergence result holds

- $(p_h^s, u_h) \quad \to (p,u) \quad$ *in* $X \times M$,
- $p_h^r \quad \to 0 \quad$ *in* H.

More precisely:

$$|p_h^r|_H + \|p - p_h^s\|_X + \|u - u_h\|_{M/\mathrm{Ker}\ B_h^t} \leq \ C\left(\inf_{q_h^s \in X_h^s} \|p - q_h^s\|_X + \inf_{w_h \in M_h} \|u - w_h\|_M \right.$$
$$\left. + \inf_{z_h^s \in X_h^s} |\mathcal{A}p - z_h^s|_H \right).$$

Proof. First note that hypothesis (H0) and nonuniform discrete coercivity on the kernel Ker B_h, i.e.,

$$\exists \alpha_h > 0,\ \forall p_h \in V_h \ \ a(p_h, p_h) \geq \alpha_h \|p_h\|_X^2\,, \tag{3.14}$$

ensure the existence and uniqueness of the solution (p_h, u_h) of the discrete problem (3.12) in $X_h \times (M_h/\mathrm{Ker}\ B_h^t)$. Since in finite dimensions all norms are equivalent, the nonuniform discrete coercivity (3.14) on the kernel Ker B_h is a consequence of (H3).

It remains to prove the error estimate. The convergence result will follow from assumption (H4) and remark 3.2.

The second equation of (3.12) shows that $p_h \in V_h(f)$ (cf (3.13)). If we also take $q_h \in V_h(f)$, the difference is in the kernel, i.e. $p_h - q_h \in V_h$.We can write

$$a(p_h - q_h, p_h - q_h) = a(p - q_h, p_h - q_h) + a(p_h - p, p_h - q_h). \tag{3.15}$$

Taking the difference between the first equation of the continuous (3.11) and the discrete (3.12) problem we have

$$a(p_h - p, p_h - q_h) = b(u - w_h, p_h - q_h), \quad \forall w_h \in M_h. \tag{3.16}$$

Using (H1) in (3.16), we can write, with obvious notation

$$a(p_h - p, p_h - q_h) = b(u - w_h, p_h^s - q_h^s) + b(u, p_h^r - q_h^r). \tag{3.17}$$

We have used the fact that $b(w_h, p_h^r - q_h^r) = 0$, since $X_h^r \subset V_h$. Now, since we only have a "weak" coercivity, we want to keep only terms leading to H-norm on the X_h^r part, thus we want to eliminate $b(u, p_h^r - q_h^r)$ (which would give $\|p_h^r - q_h^r\|_X$from the continuity of b). Then, by (3.11),

$$\begin{aligned} a(p_h - p, p_h - q_h) &= b(u - w_h, p_h^s - q_h^s) - a(p, p_h^r - q_h^r) \\ &= b(u - w_h, p_h^s - q_h^s) - (\mathcal{A}p, p_h^r - q_h^r)_H. \end{aligned} \tag{3.18}$$

We choose now

$$\text{(3.19)}\qquad \left|\begin{array}{l} q_h = q_h^s \in V_h^s(f) = V_h(f) \cap X_h^s, \\[2ex] q_h^r = 0. \end{array}\right.$$

From (3.15), (3.18) and (3.19), we obtain

$$a(p_h - q_h^s, p_h - q_h^s) = a(p - q_h^s, p_h - q_h^s) + b(u - w_h, p_h^s - q_h^s) - (\mathcal{A}p, p_h^r),$$

or, by using the orthogonality of X_h^r and X_h^s from (H1),

$$\text{(3.20)}\qquad \left|\begin{array}{ll} a(p_h - q_h^s, p_h - q_h^s) = & a(p - q_h^s, p_h - q_h^s) + b(u - w_h, p_h^s - q_h^s) \\[2ex] & -(\mathcal{A}p - z_h^s, p_h^r)_H, \qquad \forall\ z_h^s \in X_h^s. \end{array}\right.$$

Furthermore, from the inequality

$$|p_h - q_h^s|_H \le \|p_h^s - q_h^s\|_X + |p_h^r|_H$$

and from hypothesis (H1) and (H3), (3.20) leads to

$$\left|\begin{array}{ll} \alpha(\|p_h^s - q_h^s\|_X^2 + |p_h^r|_H^2) & \le \|a\|\, |p - q_h^s|_H \,(\|p_h^s - q_h^s\|_X + |p_h^r|_H) \\[2ex] & \quad + \|b\|\, \|u - w_h\|_M \, \|p_h^s - q_h^s\|_X + |\mathcal{A}p - z_h^s|_H \, |p_h^r|_H \,. \end{array}\right.$$

We deduce the existence of a constant C depending on f, p, $\|a\|$, $\|b\|$, and α such that, for all $(q_h^s, w_h, z_h^s) \in V_h^s(f) \times M_h \times X_h^s$,

$$\|p_h^s - q_h^s\|_X + |p_h^r|_H \le C\,(|p - q_h^s|_H + \|u - w_h\|_M + |\mathcal{A}p - z_h^s|_H)\,.$$

This gives, using (3.9)

$$\left|\begin{array}{ll} \|p - p_h^s\|_X & \le \|p - q_h^s\|_X + \|p_h^s - q_h^s\|_X\,, \quad \forall q_h^s \in V_h^s(f) \\[2ex] & \le (1 + C)\left(\inf_{q_h^s \in V_h^s(f)} \|p - q_h^s\|_X + \inf_{w_h \in M_h} \|u - w_h\|_M + \inf_{z_h^s \in X_h^s} |\mathcal{A}p - z_h^s|_H\right), \\[2ex] |p_h^r|_H & \le C\left(\inf_{q_h^s \in V_h^s(f)} \|p - q_h^s\|_X + \inf_{w_h \in M_h} \|u - w_h\|_M + \inf_{z_h^s \in X_h^s} |\mathcal{A}p - z_h^s|_H\right). \end{array}\right.$$

To conclude, let us recall that the inf-sup condition (H2) implies (cf. [6]) that

$$\inf_{q_h^s \in V_h^s(f)} \|p - q_h^s\|_X \le c_1 \inf_{q_h^s \in X_h^s} \|p - q_h^s\|_X\,.$$

Finally, it remains to prove estimates for $\|u - u_h\|_{M/\mathrm{Ker}\ B_h^t}$. Let us subtract the first equation of (3.12) from the first equation of (3.11). We get

$$a(p - p_h, q_h) + b(u - u_h, q_h) = 0, \ \forall q_h \in X_h,$$

so that, for any w_h in M_h,

$$b(u_h - w_h, q_h) = a(p - p_h, q_h) + b(u - w_h, q_h), \ \forall q_h \in X_h.$$

Choosing now $q_h = q_h^s$ implies that

$$b(u_h - w_h, q_h^s) = a(p - p_h, q_h^s) + b(u - w_h, q_h^s),\ \forall q_h^s \in X_h^s.$$

Using this and the inf-sup condition, we have

$$\begin{aligned}\|u_h - w_h\|_{M/\mathrm{Ker}\ B_h^t} &\le \frac{1}{c} \sup_{q_h^s \in X_h^s} \frac{b(u_h - w_h, q_h^s)}{\|q_h^s\|_X} \\ &\le \frac{1}{c} \sup_{q_h^s \in X_h^s} \frac{a(p - p_h, q_h^s) + b(u - w_h, q_h^s)}{\|q_h^s\|_X} \\ &\le \frac{1}{c} \left(\|a\|\,|p - p_h|_H + \|b\|\,\|u - w_h\|_M\right).\end{aligned}$$

It follows from the triangle inequality that

$$\|u - u_h\|_{M/\mathrm{Ker}\ B_h^t} \le C' \left\{ \inf_{w_h \in M_h} \|u - w_h\|_M + |p_h^r|_H + |p - p_h^s|_H \right\}. \qquad \square$$

REMARK 3.5. *One could easily remark that the space X_h^s satisfies all assumptions (H0) to (H4). Why not using X_h^s as approximate space instead of X_h? In fact, we should imagine the case in which either we cannot characterize the space X_h^s or we prefer X_h for other reasons. This occurs in particular for the evolution problem, where we prefer X_h in order to achieve mass lumping.*

3.3. Application to the model problem. For problem (3.5) we have,

$$H = (L^2(\Omega))^2\ ;\ X = H(\mathrm{div}\ ;\Omega)\ ;\ M = L^2(\Omega)$$

The operator B is the divergence operator, which is surjective from X onto M, so that Ker $B^t = \{0\}$. From Remark 3.3, it follows that Hypothesis (H0) is equivalent to the surjectivity of B_h from X_h onto M_h. This is true for the classical Raviart-Thomas approximations [21]; i.e. B_h is surjective from X_h^{RT} onto M_h. Since the new space X_h contains the space X_h^{RT}, as noted in remark 2.7, it is straightforward to check (H0); therefore, Ker $B_h^t = \{0\}$.

3.3.1. Approximation with the lowest order finite element $Q_1^{div} - Q_0$. Let us begin by checking Hypothesis (H1). We take as approximation spaces

$$\begin{cases} X_h = \{q_h \in X \ /\ \forall K \in \mathcal{T}_h,\ q_h|_K \in Q_1 \times Q_1\}, \\ M_h = \{w_h \in M \ /\ \ \forall K \in \mathcal{T}_h,\ w_h|_K \in Q_0\}. \end{cases}$$

We define X_h^s as the lowest order Raviart-Thomas element $RT_{[0]}$ defined in (2.10)

$$X_h^s = \left\{q_h \in X \ /\ \forall K \in \mathcal{T}_h,\ q_h|_K \in RT_{[0]} = P_{10} \times P_{01}\right\} \subset X_h,\ \ \dim RT_{[0]} = 4$$

In order to describe its orthogonal complement X_h^r, denote by

- (i, j), the node of $\mathcal{T}_h$ with coordinates (ih, jh),
- $(i + \frac{1}{2}, j)$, the horizontal side joining the nodes (i, j) and $(i + 1, j)$,
- $(i, j + \frac{1}{2})$, the vertical side joining the nodes (i, j) and $(i, j + 1)$.

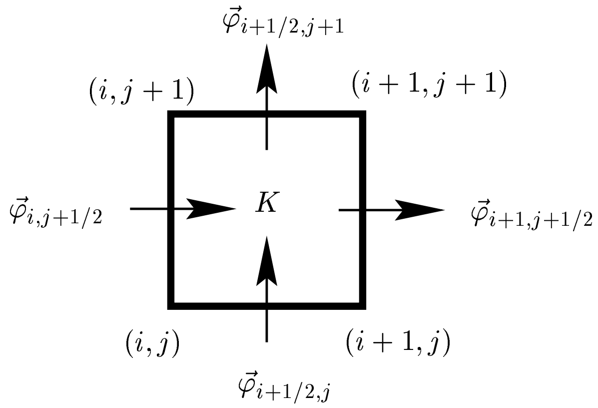


FIG. 3.2. *The $RT_{[0]}$ element*

Over the square K, the 4 basis functions of $RT_{[0]}$ can be written as (see figure 3.2 for an illustration)

$$\left\{\begin{array}{ll} \vec{\varphi}_{i,j+\frac{1}{2}} = \begin{pmatrix} \varphi_{i,j+\frac{1}{2}}(x) \\ 0 \end{pmatrix}, & \vec{\varphi}_{i+\frac{1}{2},j} = \begin{pmatrix} 0 \\ \varphi_{i+\frac{1}{2},j}(y) \end{pmatrix}, \\ \vec{\varphi}_{i+1,j+\frac{1}{2}} = \begin{pmatrix} \varphi_{i+1,j+\frac{1}{2}}(x) \\ 0 \end{pmatrix}, & \vec{\varphi}_{i+\frac{1}{2},j+1} = \begin{pmatrix} 0 \\ \varphi_{i+\frac{1}{2},j+1}(y) \end{pmatrix}, \end{array}\right.$$

where

$$\left\{\begin{array}{ll} \varphi_{i,j+\frac{1}{2}}(x) = \dfrac{(x_{i+1}-x)}{h} \in P_{10}, & \varphi_{i+\frac{1}{2},j}(y) = \dfrac{(y_{j+1}-y)}{h} \in P_{01}, \\ \varphi_{i+1,j+\frac{1}{2}}(x) = \dfrac{(x-x_i)}{h} \in P_{10}, & \varphi_{i+\frac{1}{2},j+1}(y) = \dfrac{(y-y_j)}{h} \in P_{01} \ . \end{array}\right.$$

It is then easy to prove the following Lemma

LEMMA 3.3. *If $X_h^r = (X_h^s)^\perp$, the space $X_h^r|_K$ of the restrictions to K of elements of X_h^r can be generated from the 4 functions:*

$$\vec{\psi}_{i+\frac{1}{2},j} = \begin{pmatrix} 0 \\ \left(x - x_{i+\frac{1}{2}}\right)\varphi_{i+\frac{1}{2},j}(y) \end{pmatrix}, \qquad \vec{\psi}_{i,j+\frac{1}{2}} = \begin{pmatrix} \left(y - y_{j+\frac{1}{2}}\right)\varphi_{i,j+\frac{1}{2}}(x) \\ 0 \end{pmatrix},$$

$$\vec{\psi}_{i+\frac{1}{2},j+1} = \begin{pmatrix} 0 \\ \left(x - x_{i+\frac{1}{2}}\right)\varphi_{i+\frac{1}{2},j+1}(y) \end{pmatrix}, \quad \vec{\psi}_{i+1,j+\frac{1}{2}} = \begin{pmatrix} \left(y - y_{j+\frac{1}{2}}\right)\varphi_{i+1,j+\frac{1}{2}}(x) \\ 0 \end{pmatrix},$$

where $x_{i+\frac{1}{2}} = (i+\frac{1}{2})h$ and $y_{j+\frac{1}{2}} = (j+\frac{1}{2})h$. Moreover, $X_h^r \subset V_h$.

REMARK 3.6. *Note that the dimension of $Q_1 \times Q_1$ (=8) is the sum of the dimensions of $X_h^s|_K$ (= dim $RT_{[0]} = 4$) and $X_h^r|_K$ (=4).*

Using now the well known properties of X_h^s and M_h [21, 20, 6], hypothesis (H2) and (H4) are trivial; only (H3) needs to be checked.

LEMMA 3.4. *For all p_h in V_h,*

$$a(p_h,p_h) \geq \alpha \left(\|p_h^s\|_X^2 + |p_h^r|_H^2 \right),$$

α being defined by (2.1).

Proof. From (2.1),

$$a(p_h,p_h) \geq \alpha \, \|p_h\|_0^2 = \alpha \left(\|p_h^s\|_0^2 + \|p_h^r\|_0^2 \right).$$

Note now that $p_h \in V_h$ and $X_h^r \subset V_h$ imply $p_h^s \in V_h \cap X_h^s$. Hence it follows (cf. [6]) that

$$\begin{cases} (i) \quad \text{div } p_h^s \in M_h, \\ (ii) \quad (\text{ div } p_h^s, w_h) = 0, \quad \forall w_h \in M_h. \end{cases}$$

Recalling (i), we can take

$$w_h = \text{ div } p_h^s$$

in (ii), which gives

$$\text{div } p_h^s = 0.$$

The claim is thus proved. $\square$

We can apply now Theorem 3.2 to the approximation problem (3.6), and by using the usual interpolation results (cf. [20]) we obtain: (here we have Ker $B_h^t = \{0\}$)

THEOREM 3.5. *The problem (3.6) admits a unique solution*

$$(p_h, u_h) \in X_h \times M_h,$$

which satisfies

- $(p_h^s, u_h) \longrightarrow (p,u)$ *in* $H(\text{div };\Omega) \times L^2(\Omega)$,
- $p_h^r \longrightarrow 0$ *in* $(L^2(\Omega))^2$.

Furthermore, if the solution is assumed to be sufficiently regular, i.e., $(p,u) \in (H^1(\Omega))^2 \times H^1(\Omega)$, div $p \in L^2(\Omega)$ *and* $\mathcal{A}p \in (H^1(\Omega))^2$, *then*

$$\|p_h^r\|_{(L^2(\Omega))^2} + \|p - p_h^s\|_{H(\text{div};\Omega)} + \|u-u_h\|_{L^2(\Omega)} \leq$$
$$Ch \left(|u|_{H^1} + |p|_{(H^1)^2} + |\text{div } p|_{H^1} + |\mathcal{A}p|_{(H^1)^2} \right),$$

where $|.|_{H^1}$ *(resp.* $|.|_{(H^1)^2}$*) denotes the usual semi-norm in* $H^1(\Omega)$ *(resp.* $(H^1(\Omega))^2$*).*

REMARK 3.7. *Let us now consider the case of the isotropic wave equation, which corresponds in taking $A(x)$ in (3.1) as a diagonal matrix ($A_d(x)$). For the discretisation we assume that $A_d(x)$ is approximated by piecewise constant functions per element. It is then easy to prove that the approximate problem*

$$\text{(3.21)} \quad \begin{cases} \text{Find } (p_h,u_h) \in X_h \times M_h \text{ such that} \\ a_d(p_h,q_h) + \quad b(u_h,q_h) \quad = 0, \qquad \forall q_h \in X_h, \\ \qquad\qquad\qquad b(w_h,p_h) \quad = -\langle f, w_h \rangle, \quad \forall w_h \in M_h, \end{cases}$$

admits a unique solution $(p_h = p_h^s + p_h^r, u_h)$ *with* $p_h^r \equiv 0$ *and* (p_h^s, u_h) *being the solution of the following problem :*

$$\left\{\begin{array}{lll} \textit{Find } (p_h^s, u_h) \in X_h^s \times M_h \textit{ such that} & & \\ a_d(p_h^s, q_h^s) + \ b(u_h, q_h^s) & = 0, & \forall q_h^s \in X_h^s, \\ \qquad\qquad b(w_h, p_h^s) & = -\langle f, w_h \rangle\,, & \forall w_h \in M_h, \end{array}\right.$$

where $a_d(\cdot,\cdot)$ *is given by (2.6) after replacing* $A(x)$ *by* $A_d(x)$*. Indeed, in this particular case,*

$$\left\{\begin{array}{l} a_d(p_h^r, q_h^s) = 0, \quad \forall q_h^s \in X_h^s, \\ a_d(p_h^s, q_h^r) = 0, \quad \forall q_h^r \in X_h^r. \end{array}\right.$$

By using the fact that $X_h^r \subset V_h$*, we can decompose (3.21) in two independent problems in* X_h^s *and* X_h^r

$$\left\{\begin{array}{lll} \textit{Find } (p_h^s, u_h) \in X_h^s \times M_h \textit{ such that} & & \\ a_d(p_h^s, q_h^s) + \ b(u_h, q_h^s) & = 0, & \forall q_h^s \in X_h^s, \\ \qquad\qquad b(w_h, p_h^s) & = -\langle f, w_h \rangle\,, & \forall w_h \in M_h, \end{array}\right.$$

and

$$\left\{\begin{array}{l} \textit{Find } p_h^r \in X_h^r \textit{ such that} \\ a_d(p_h^r, q_h^r) \ = 0, \quad \forall q_h^r \in X_h^r, \end{array}\right. \tag{3.22}$$

It is obvious then from (3.22) that $p_h^r \equiv 0$*. This remark is no longer true when* $A(x)$ *is not diagonal.*

3.3.2. Approximation with higher order finite elements, $Q_{k+1}^{div} - Q_k$**.** We have seen in §2.3 that the natural generalization of the lowest order element consists in taking

$$\hat{X} = X_k = Q_{k+1} \times Q_{k+1} \quad \text{and} \quad \hat{M} = M_k = Q_k;$$

and thus to introduce the spaces

$$\left\{\begin{array}{l} X_h = \{q_h \in X \,/\, \forall K \in \mathcal{T}_h,\ q_h|_K \in X_k\}\,, \\ M_h = \{w_h \in M \,/\ \forall K \in \mathcal{T}_h,\ w_h|_K \in M_k\}\,. \end{array}\right. \tag{2.16}$$

We are going to show that we can apply the abstract theory of §3.2 to these spaces, with the main difficulty lying in the construction of an orthogonal decomposition of X_h such that assumptions (H0) to (H4) are satisfied. In fact we shall deduce such a global decomposition from a local decomposition of the space $\hat{X}$, considered as a subspace of $L^2(K)$, where K is a single element. First, we recall that

$$RT_{[k]} = P_{k+1,k} \times P_{k,k+1}\,,$$

which obviously satisfies the inclusion

$$RT_{[k]} \subset X_k.$$

Let us define Ψ_k as the orthogonal complement in X_k of $RT_{[k]}$ (for the inner product in $L^2(K)$):

$$\Psi_k = \left\{ \psi \in \hat{X} \;/\; \int_K \psi \; \phi dx = 0, \;\forall \phi \in RT_{[k]} \right\}.$$

Note that

$$\dim \Psi_k = 2(k+2).$$

The main property of the space Ψ_k is the following.

LEMMA 3.6. *For any ψ in Ψ_k and v in M_k,*

$$\int_K \text{div} \;\psi \; v \; dx = 0. \tag{3.23}$$

Proof. For simplicity we consider the reference element $K = [0,1] \times [0,1]$.

(i) We begin by a characterization of Ψ_k. Let σ_k be the polynomial in one variable of degree $k+1$ such that

$$P_{k+1} = P_k \oplus [\sigma_k].$$

Equivalently, σ_k is defined, up to a multiplicative constant by

$$\begin{cases} \displaystyle\int_0^1 \sigma_k(x) p(x) dx = 0, \quad \forall p \in P_k, \\ \sigma_k \in P_{k+1}, \;\; \sigma_k \not\equiv 0 \end{cases}$$

Then, we claim that

$$\Psi_k = \widetilde{\Psi}_k \equiv \left\{ \begin{pmatrix} p_x(x)\sigma_k(y) \\ \\ p_y(y)\sigma_k(x) \end{pmatrix}, \; (p_x, p_y) \in P_{k+1} \times P_{k+1} \right\}.$$

Indeed, as $\dim \widetilde{\Psi}_k = 2(k+2)$, it is sufficient to check that $\widetilde{\Psi}_k$ is orthogonal to $RT_{[k]}$. To see that, we first remark that (see [21]) $RT_{[k]}$ is generated by vector fields of the form

$$\phi(x,y) = \begin{bmatrix} \tau_x(x) q_x(y) \\ \\ \tau_y(y) q_y(x) \end{bmatrix} \quad \text{with} \quad (\tau_x, \tau_y) \in P_{k+1} \times P_{k+1}, \; (q_x, q_y) \in P_k \times P_k.$$

Let us consider

$$\psi(x,y) = \begin{bmatrix} p_x(x)\sigma_k(y) \\ \\ p_y(y)\sigma_k(x) \end{bmatrix} \quad \text{with} \quad (p_x, p_y) \in P_{k+1} \times P_{k+1}. \tag{3.24}$$

We have

$$
\begin{aligned}
(\phi,\psi)_{L^2(K)} &= \left(\int_0^1 \tau_x(x)p_x(x)dx\right)\left(\int_0^1 \sigma_k(y)q_x(y)dy\right)\\
&+ \left(\int_0^1 \sigma_k(x)q_y(x)dx\right)\left(\int_0^1 \tau_y(y)p_y(y)dy\right).
\end{aligned}
$$

As q_x and q_y are in P_k, by the definition of σ_k,

$$\int_0^1 \sigma_k(y)q_x(y)dy = \int_0^1 \sigma_k(x)q_y(x)dx = 0 \Rightarrow (\phi,\psi)_{L^2(K)} = 0.$$

(ii) To prove (3.23) we first use Green's formula

$$\int_K \text{div } \psi \; v dx = -\int_K \psi \cdot \nabla v dx + \int_{\partial K} (\psi \cdot n) \; v \; d\gamma.$$

If v belongs to Q_{k+1}, ∇v belongs to $RT_{[k]}$. Therefore, since ψ belongs to Ψ_k,

$$\int_K \text{div } \psi \; v \; dx = \int_{\partial K} (\psi \cdot n) \; v \; d\gamma.$$

Let us decompose ∂K as

$$\partial K = T_1 \cup T_2 \cup T_3 \cup T_4,$$

as shown in Figure 3.3. Then, we check that

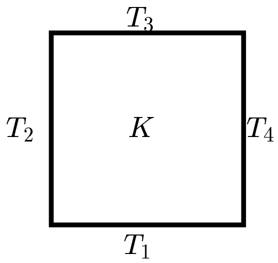


FIG. 3.3. *The edges of K.*

$$\text{for } 1 \le j \le 4 \text{ and } \forall q \in R_k(T_j), \; \int_{T_j} (\psi \cdot n) q d\gamma = 0,$$

where $R_k(T_j)$ is the set of polynomials of degree k with respect to the abscissa along T_j. Let us consider, for instance, $j = 1$ and assume that ψ is given by (3.24). Then

$$\int_{T_1} (\psi \cdot n) q d\gamma = -\left(\int_0^1 \sigma_k(x)q(x)\right) p_y(0) = 0.$$

To conclude, it remains to notice that, if v belongs to Q_{k+1}, then $v|_{T_j}$ belongs to $R_k(T_j)$. $\square$

This result suggests that we define X_h^s and X_h^r by

$$\left\{ \begin{array}{l} X_h^s = \left\{ p_h \in H(\mathrm{div};\Omega) \,/\, \forall K \in \mathcal{T}_h,\ p_h|_K \in RT_{[k]} \right\}, \\[2ex] X_h^r = \left\{ p_h \in H(\mathrm{div};\Omega) \,/\, \forall K \in \mathcal{T}_h,\ p_h|_K \in \Psi_k \right\}. \end{array} \right. \tag{3.25}$$

Let us now state the main result of this section,

THEOREM 3.7. *The spaces (X_h, M_h), defined in (2.16), satisfy assumptions (H0) to (H4). The spaces X_h^s and X_h^r that achieve the orthogonal decomposition (H1) of X_h are the ones defined in (3.25).*

Proof. (H0), (H3) and (H4) are classical properties of the Raviart-Thomas approximation spaces [21]. (H1) and (H2) are straightforward consequences (decomposing the integrals over Ω as the sum of the integrals over elements K) of the definition of Ψ_k and Lemma 3.6. ◻

REMARK 3.8. *For $k = 0$ we obtain the orthogonal decomposition of X_h described in §3.3.1 that we summarize in figure 3.4. In the same way we can summarize the*

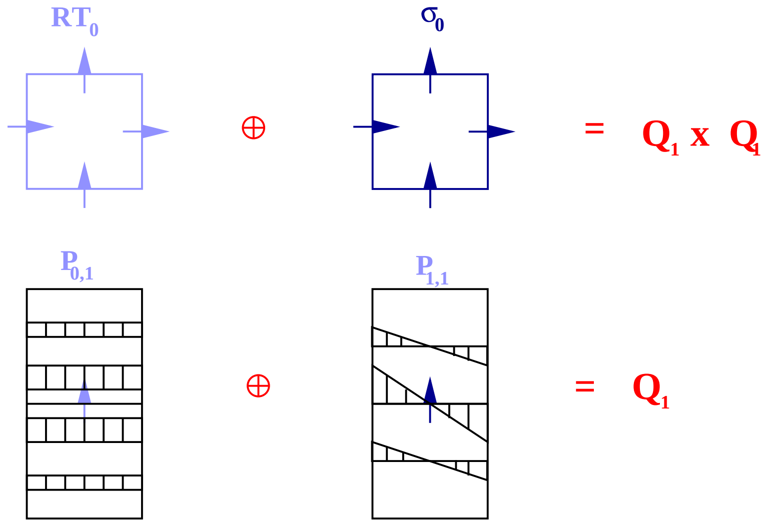


FIG. 3.4. *Orthogonal decomposition of $Q_1 \times Q_1$*

orthogonal decomposition of X_h for any k in figure 3.5.

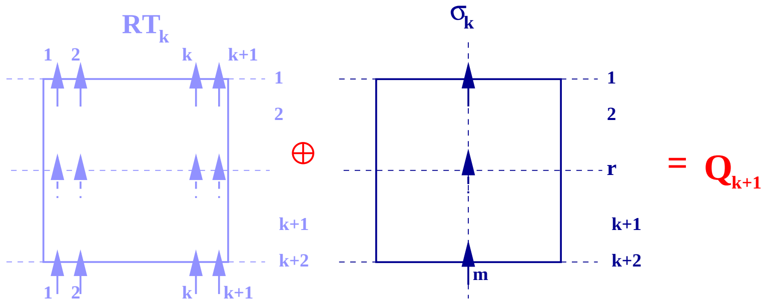


FIG. 3.5. *Orthogonal decomposition of Q_{k+1}*

We can now apply Theorem 3.1 to the approximation problem (3.6), and by using the usual interpolation results (cf. [20]) in X_h^s and M_h defined by (2.16), (3.25) we obtain the following result.

THEOREM 3.8. *Problem (3.6) admits a unique solution* $(p_h, u_h) \in X_h \times M_h$, *which satisfies*

- $(p_h^s, u_h) \quad \to (p,u) \quad$ *in* $H(\text{div}\,;\Omega) \times L^2(\Omega)$,
- $p_h^r \quad \to 0 \quad$ *in* $(L^2(\Omega))^2$.

Furthermore, if we assume that the solution (p,u) *of (3.5) is such that* $(p,u) \in (H^m(\Omega))^2 \times (H^m(\Omega))$, $\text{div}\ p \in H^m(\Omega)$ *and* $\mathcal{A}p \in (H^m(\Omega))^2$ *for* $m = k+1$, *then*

$$
(3.26)\quad \left\{ \begin{aligned} & \|p_h^r\|_{(L^2)^2} + \|p - p_h^s\|_{H(\text{div};\Omega)} + \|u - u_h\|_{L^2} \le \\ & \qquad C\ h^m \left(|u|_{H^m} + |p|_{(H^m)^2} + |\text{div}\ p|_{H^m} + |\mathcal{A}p|_{(H^m)^2} \right), \end{aligned} \right.
$$

where $|.|_{H^m}$ *(resp.* $|.|_{(H^m)^2}$*) denotes the usual semi-norm in* $H^m(\Omega)$ *(resp.* $(H^m(\Omega))^2$*).*

3.4. Application to the elliptic projection operator. We come back to the abstract framework described in §3.2. We consider the problem of finding $(\widehat{p}_h, \widehat{u}_h) \equiv \Pi_h(p,u)$ such that

$$
(3.27)\quad \left\{ \begin{aligned} a(p - \widehat{p}_h, q_h) + \quad b(u - \widehat{u}_h, q_h) &= 0, \quad \forall q_h \in X_h, \\ b(w_h, p - \widehat{p}_h) &= 0, \quad \forall w_h \in M_h. \end{aligned} \right.
$$

We set

$$
D(B^t) = \{w \in M \ / \ b(w,q) \le C(w)\, |q|_H\,,\ \forall q \in X\} = \left\{ w \in M \ / \ B^t w \in H \right\}.
$$

For $w \in D(B^t)$, we have

$$
b(w,q) = (B^t w, q)_H, \ \ \forall\ q \in X.
$$

Let us introduce the notation

$$
(3.28)\quad \left\{ \begin{aligned} & |||p - \widehat{p}_h||| = \|p - \widehat{p}_h{}^s\|_X + |\widehat{p}_h{}^r|_H, \\ & |||(p,u) - \Pi_h(p,u)|||_C = |||p - \widehat{p}_h||| + \|u - \widehat{u}_h\|_M. \end{aligned} \right.
$$

It is straightforward, by application of the abstract result, to get interpolation estimates

THEOREM 3.9. *Assume (H0) - (H4).*

(i) For all $(p,u) \in X \times M$, *problem (3.27) admits a unique solution*

$$
\Pi_h(p,u) = (\widehat{p}_h{}^s + \widehat{p}_h{}^r, \widehat{u}_h) \in X_h \times (M_h/\text{Ker}\ B_h^t).
$$

(ii) For all $(p,u) \in X \times D(B^t)$, *there exists a constant* C, *independent of* h, *such that*

$$
|||(p,u) - \Pi_h(p,u)|||_C \le C\ \mathcal{E}_h(p,u),
$$

where

$$\mathcal{E}_h(p,u) = \inf_{q_h^s \in X_h^s} \|p - q_h^s\|_X + \inf_{w_h \in M_h} \|u - w_h\|_M + \inf_{z_h^s \in X_h^s} |B^t u - z_h^s|_H.$$

In particular, $|||p - \widehat{p}_h|||$ *and* $\|u - \widehat{u}_h\|_M$ *tend to zero.*

Proof. (i) The existence and uniqueness still comes from hypotheses (H0) and (H3) (with $f = Bp$ which is, of course, in Im B).

(ii) The proof of this point is a minor modification of the proof of Theorem 3.2. The only difference concerns the treatment of the term $b(u, p_h^r - q_h^r)$ in (3.17). In order to get the error estimate, it is necessary to relate this term to the H-scalar product. In the proof of Theorem 3.2, this was done using the first continuous equation which implies that $b(u, p_h^r - q_h^r) = -a(p, p_h^r - q_h^r) \equiv -(\mathcal{A}p, p_h^r - q_h^r)_H$. Now we assume that $u \in D(B^t)$, so that we have

$$b(u, p_h^r - q_h^r) = (B^t u, p_h^r - q_h^r)_H.$$

The result is thus obtained by replacing $-\mathcal{A}p$ with $B^t u$. $\square$

4. Error estimates for the evolution problem. Let us now come back to the initial evolution problem (2.5), (2.4) and see how we can relate the error estimates to the one obtained for the elliptic problem (3.5). This part of the analysis is inspired from [15] developed for the approximation of second order hyperbolic problems with conforming finite elements. Although we have constructed this element in order to be able to do mass lumping, we analyze the error for the discrete problem without mass lumping. Of course, if mass lumping is applied, one should add to this error the quadrature error due to the numerical integration (see [7, 25, 17] and Remark 2.6).

4.1. From error estimates for the elliptic problem to error estimates for the evolution problem: the abstract case. In this part, we use the same notation and hypotheses as in §3.2, and we consider the evolution problem

$$(4.1)\qquad \begin{cases} \text{Find } (p,v) : [0,T] \to X \times M \text{ such that} \\[2mm] \dfrac{d}{dt} a(p,q) + b(v,q) = 0, & \forall q \in X, \\[2mm] \dfrac{d}{dt}(v,w) - b(w,p) = (f,w), & \forall w \in M, \\[2mm] p(0) = p_0 \ ; \ v(0) = v_0. \end{cases}$$

Or, equivalently, in an operator form, if we assume enough regularity on the solution in time

$$(4.2)\qquad \begin{cases} \text{Find } (p,v) \in (C^1(0,T;H) \cap C^0(0,T;X)) \times C^1(0,T;M) \text{ such that} \\[2mm] \mathcal{A}\dfrac{dp}{dt} + B^t v = 0 \quad \text{in } X', \\[2mm] \dfrac{dv}{dt} - Bp = f \quad \text{in } M', \\[2mm] p(0) = p_0 \ ; \ v(0) = v_0. \end{cases}$$

In the following, we use the notation $\mathcal{C}^{m,r} = C^m(0,T;H) \cap C^r(0,T;X)$. X_h and M_h are, as in §3.2, finite dimensional spaces satisfying hypothesis (H0) to (H4). We consider the approximate problem

$$
\left\{
\begin{array}{lll}
\text{Find } (p_h, v_h) : [0,T] \to X_h \times M_h \text{ such that} & & \\[2mm]
\dfrac{d}{dt} a(p_h, q_h) + b(v_h, q_h) & = 0, & \forall q_h \in X_h, \\[2mm]
\dfrac{d}{dt}(v_h, w_h) - b(w_h, p_h) & = (f, w_h), & \forall w_h \in M_h, \\[2mm]
p_h(0) = p_{0,h} \ ; \ v_h(0) = v_{0,h}. & &
\end{array}
\right.
\tag{4.3}
$$

From the classical theory of ODE's, we have the following result.

THEOREM 4.1. *If $f \in C^0(0,T;M_h)$, then problem (4.3) has a unique solution $(p_h, v_h) \in C^1(0,T;X_h) \times C^1(0,T;M_h)$.*

Following [12, 15] we introduce the elliptic operator defined in (3.27). By application of Theorem 3.9, we get the following interpolation results

LEMMA 4.2. *Let (p,v) be the solution of (4.2) and assume that $(p,v) \in \mathcal{C}^{1,0} \times C^1(0,T;M)$, then*

(i) there exists a primitive of v, $u \in C^1(0,T;M)$, satisfying

$$
\left\{
\begin{array}{l}
\dfrac{du}{dt} = v, \\[2mm]
\mathcal{A}p_0 + B^t u(0) = 0.
\end{array}
\right.
\tag{4.4}
$$

This primitive is unique up to a constant element of Ker B^t.

(ii) $\forall t \in [0,T]$, problem (3.27) admits a unique solution $\Pi_h(p,u)(t) = (\widehat{p}_h, \widehat{u}_h)(t) \in X_h \times (M_h/\text{Ker } B_h^t)$ and there exists a constant C independent of h such that

$$
|||(p,u) - \Pi_h(p,u)|||_C\,(t) \le C\ \mathcal{E}_h(p,u)(t),
\tag{4.5}
$$

where $\mathcal{E}_h$ is defined by,

$$
\mathcal{E}_h(p,u)(t) = \inf_{q_h^s \in X_h^s} \|p(t) - q_h^s\|_X + \inf_{w_h \in M_h} \|u(t) - w_h\|_M + \inf_{z_h^s \in X_h^s} |\mathcal{A}p(t) - z_h^s|_H .
$$

In particular, $|||p - \widehat{p}_h|||$ and $\|u - \widehat{u}_h\|_M$ tend to zero uniformly in time ($t \in [0,T]$) ($|||.|||_C$ and $|||.|||$ are defined in (3.28)).

(iii) In the same way, if $(p,u) \in C^k(0,T;X) \times C^k(0,T;M)$, $k \ge 1$, there exists a constant C independent of h such that

$$
\left|\left|\left|(\partial_t^k p, \partial_t^k u) - \Pi_h\left(\partial_t^k p, \partial_t^k u\right)\right|\right|\right|(t) \le C\ \mathcal{E}_h\left(\partial_t^k p, \partial_t^k u\right)(t),
\tag{4.6}
$$

where $\partial_t^k g = \dfrac{d^k g}{dt^k}$.

REMARK 4.1. *Operators Π_h and ∂_t^k commute, and we set*

$$
\widehat{v}_h = \partial_t\left(\widehat{u}_h\right) = \left(\widehat{\partial_t u}\right)_h .
\tag{4.7}
$$

REMARK 4.2. *It is quite astonishing that the error estimate does not require the same regularity in time on v and on p, as one could expect ($p \in C^k$ while $v \in C^{k-1}$). This is a specificity of the new element and is due to the fact that in order to eliminate the terms involving an X-norm of elements in X_h^r, we used the first equation satisfied by the solution, which relates the embarrassing term $b(v, p_h^r - q_h^r)$ to a term which is like an H-scalar product, or more precisely, for the evolution problem, which is the time derivative of something like an H-scalar product, say $\frac{d}{dt}a(p, p_h^r - q_h^r)$ (see proof of Theorem 3.2).*

Proof. (of Lemma 4.2). (i) We set $f_0 = -Bp_0 \in \mathrm{Im}\ B$. From hypothesis (3.10), we know that there is a unique $(p_0, u_0) \in X \times (M/\mathrm{Ker}\ B^t)$ such that

$$\begin{cases} a(p_0, q) + \quad b(u_0, q) &= 0, \qquad \forall q \in X, \\ \qquad\qquad b(w, p_0) &= -(f_0, w), \quad \forall w \in M, \end{cases}$$

which means that, p_0 being fixed, there is a unique $u_0 \in (M/\mathrm{Ker}\ B^t)$ such that $\mathcal{A}p_0 + B^t u_0 = 0$. Now we define u as

$$u(t) = u_0 + \int_0^t v(s)ds.$$

It is clear that $u \in C^1(0, T; M)$ and is the unique solution of (4.4).

(ii) Let $u \in C^1(0, T; M)$ be the primitive of v, substituting this into the first equation of (4.2) gives

$$\frac{d}{dt}(\mathcal{A}p + B^t u) = 0 \Longrightarrow (\mathcal{A}p + B^t u)(t) = \mathcal{A}p_0 + B^t u(0) = 0;$$

thus, $(p, u) \in C^1(0, T; X) \times C^1(0, T; M)$ satisfies

$$a(p, q) + b(u, q) = 0, \quad \forall q \in X,$$

so that $u \in D(B^t)$ and $B^t u = -\mathcal{A}p$. Applying Theorem 3.9, we get the existence and uniqueness of the elliptic projection, for t fixed, and also the error estimate (4.5).
(iii) If (p, u) is sufficiently regular in time, we can differentiate with respect to t and get

$$\mathcal{A}\partial_t^k p + B^t \partial_t^k u = 0.$$

We apply again Theorem 3.9 to get (4.6). $\square$

We now give the main result.

THEOREM 4.3. *Assume (H0) – (H4). Let (p, v) be the solution of (4.2) and (p_h, v_h) the solution of the approximate problem (4.3) with the initial conditions*

$$(p_{0,h}, v_{0,h}) = \Pi_h(p_0, v_0). \tag{4.8}$$

- *If $(p, v) \in C^2(0, T; X) \times C^1(0, T; M)$, then, $\forall t \in [0, T]$,*

$$|p - p_h^s|_H(t) \to 0 \ ; \ |p_h^r|_H(t) \to 0 \ ; \quad \|v - v_h\|_M(t) \to 0.$$

More precisely,

$$
(4.9)\quad \left\{ \begin{array}{l} |p - p_h^s|_H(t) + |p_h^r|_H(t) \leq C \left(\mathcal{E}_h(p,u)(t) + \displaystyle\int_0^t \mathcal{E}_h \left(\partial_t^2 p, \partial_t v \right)(s) ds \right), \\ \|v - v_h\|_M(t) \leq C \left(\mathcal{E}_h \left(\partial_t p, v \right)(t) + \displaystyle\int_0^t \mathcal{E}_h \left(\partial_t^2 p, \partial_t v \right)(s) ds \right). \end{array} \right.
$$

• *If, in addition,* $(p,v) \in C^3(0,T;X) \times C^2(0,T;M)$ *and* $(p_h, v_h) \in C^2(0,T;X_h) \times C^2(0,T;M_h)$,

$$
\forall t \in [0,T], \ \|p - p_h^s\|_X(t) \to 0.
$$

More precisely,

$$
(4.10)\quad \left\{ \begin{array}{ll} \|p - p_h^s\|_X(t) \leq & C(\mathcal{E}_h(p,u)(t) + \mathcal{E}_h \left(\partial_t^2 p, \partial_t v \right)(t) \\ & \qquad + \displaystyle\int_0^t \left(\mathcal{E}_h \left(\partial_t^2 p, \partial_t v \right)(s) + \mathcal{E}_h \left(\partial_t^3 p, \partial_t^2 v \right)(s) \right) ds). \end{array} \right.
$$

In order to prove Theorem 4.3 we need the following lemma

LEMMA 4.4. *Let (p,v) be the solution of (4.2) and (p_h, v_h) the solution of the approximate problem (4.3) with the initial conditions (4.8). Let $\Pi_h(p,u) = (\widehat{p}_h{}^s + \widehat{p}_h{}^r, \widehat{u}_h)$ be the elliptic projection defined in Lemma 4.2, and let $\widehat{v}_h$ be defined by (4.7). Set $\delta_h = v - \widehat{v}_h$.*

(i) If $(p,v) \in \mathcal{C}^{1,0} \times C^1(0,T;M)$, there exists a constant C_1, independent of h such that, $\forall t \in [0,T]$,

$$
(4.11)\quad |\widehat{p}_h{}^s - p_h^s|_H(t) + |\widehat{p}_h{}^r - p_h^r|_H(t) + \|\widehat{v}_h - v_h\|_M(t) \leq C_1 \int_0^t \|\partial_t \delta_h\|_M(s)\ ds.
$$

(ii) Moreover, if $(p,v) \in \mathcal{C}^{2,1} \times C^2(0,T;M)$ and $(p_h, v_h) \in C^2(0,T;X_h) \times C^2(0,T;M_h)$, there exists a constant C_2, independent of h such that, $\forall t \in [0,T]$,

$$
(4.12)\quad \|\widehat{p}_h{}^s - p_h^s\|_X(t) \leq C_2 \left\{ \|\partial_t \delta_h\|_M(t) + \int_0^t \left(\|\partial_t \delta_h\|_M(s) + \|\partial_t^2 \delta_h\|_M(s) \right) ds \right\}.
$$

Proof. (of Lemma 4.4). (i) To prove estimation (4.11), we begin by rewriting equations (4.1) with test functions $q = q_h \in X_h \subset X$ and $w = w_h \in M_h \subset M$ and by subtracting it from (4.3)

$$
(4.13)\quad \left\{ \begin{array}{ll} \dfrac{d}{dt} a(p - p_h, q_h) + b(v - v_h, q_h) = 0, & \forall q_h \in X_h, \\ \dfrac{d}{dt}(v - v_h, w_h) - b(w_h, p - p_h) = 0, & \forall w_h \in M_h, \\ (p - p_h)(0) = p_0 - p_{0,h} \ \ ; \ \ (v - v_h)(0) = v_0 - v_{0,h}. & \end{array} \right.
$$

Introducing the elliptic projection $\Pi_h(p,u) = (\widehat{p}_h, \widehat{u}_h)$, we split the error between the approximate solution and the exact solution into two parts

$$
(4.14)\quad \left\{ \begin{array}{ll} (p - p_h)(t) = & (p - \widehat{p}_h)(t) + (\widehat{p}_h - p_h)(t), \\ (v - v_h)(t) = & (v - \widehat{v}_h)(t) + (\widehat{v}_h - v_h)(t), \end{array} \right.
$$

and we choose as approximate initial conditions the elliptic projection of the exact initial condition, (4.8), so that at time $t = 0$

$$(\widehat{p}_h - p_h)(0) = 0 \;\; ; \;\; \widehat{v}_h - v_h(0) = 0.$$

By (4.13) and (4.14), for any $(q_h, w_h) \in X_h \times M_h$,

$$\text{(4.15)} \quad \left\{ \begin{array}{l} a(\partial_t(\widehat{p}_h - p_h), q_h) + b(\widehat{v}_h - v_h, q_h) = -a(\partial_t(p - \widehat{p}_h), q_h) - b(v - \widehat{v}_h, q_h), \\ \\ (\partial_t(\widehat{v}_h - v_h), w_h) - b(w_h, \widehat{p}_h - p_h) = -(\partial_t(v - \widehat{v}_h), w_h) + b(w_h, p - \widehat{p}_h). \end{array} \right.$$

By differentiating the first equation of (3.27) (written for (p, u)) with respect to t, we see that

$$\left\{ \begin{array}{rll} a(\partial_t(p - \widehat{p}_h), q_h) + \quad b(v - \widehat{v}_h, q_h) & = 0, & \forall q_h \in X_h, \\ \\ b(w_h, p - \widehat{p}_h) & = 0, & \forall w_h \in M_h. \end{array} \right.$$

Substituting into (4.15) gives, for any $(q_h, w_h) \in X_h \times M_h$,

$$\text{(4.16)} \qquad \left\{ \begin{array}{ll} a(\partial_t(\widehat{p}_h - p_h), q_h) + b(\widehat{v}_h - v_h, q_h) & = 0, \\ \\ (\partial_t(\widehat{v}_h - v_h), w_h) - b(w_h, \widehat{p}_h - p_h) & = -(\partial_t(v - \widehat{v}_h), w_h). \end{array} \right.$$

Furthermore, by taking $q_h = \widehat{p}_h - p_h$ and $w_h = \widehat{v}_h - v_h$ in (4.15) and by adding the two equations, we get

$$\text{(4.17)} \quad a(\partial_t(\widehat{p}_h - p_h), \widehat{p}_h - p_h) + (\partial_t(\widehat{v}_h - v_h), \widehat{v}_h - v_h) = -(\partial_t(v - \widehat{v}_h), \widehat{v}_h - v_h).$$

Next, set

$$E_h(t) = \frac{1}{2} \left(a(\widehat{p}_h - p_h, \widehat{p}_h - p_h) + (\widehat{v}_h - v_h, \widehat{v}_h - v_h) \right)(t).$$

Since for some constant $C > 0$,

$$E_h^{1/2}(t) \geq C \left(|\widehat{p}_h - p_h|_H^2(t) + \|\widehat{v}_h - v_h\|_M^2(t) \right)^{1/2},$$

it follows that

$$\text{(4.18)} \qquad \frac{dE_h^{1/2}}{dt}(t) \leq C' \|\partial_t(v - \widehat{v}_h)\|_M(t) \equiv C' \|\partial_t \delta_h\|_M(t),$$

where we set $\delta_h = v - \widehat{v}_h$. From the choice (4.8) for the initial conditions,

$$E_h(0) = 0.$$

It is then easy to see that (4.18) leads to (4.11) (from the orthogonality $X_h^s \perp X_h^r$, $|q_h|_H^2 = |q_h^s|_H^2 + |q_h^r|_H^2$, $\forall q_h \in X_h$) .

(ii) To get the estimate (4.12) in the X-norm, we recall that, for all $\eta_h^s \in X_h^s$, we have $\eta_h^s = \eta_1 + \eta_2$ with $\eta_1 \in \text{Ker } B_h$ and $\eta_2 \in (\text{Ker } B_h)^\perp$, so that

$$\|\eta_h^s\|_X^2 = |\eta_1|_H^2 + \|\eta_2\|_X^2$$

(recalling that $\|\eta_1\|_X = |\eta_1|_H, \ \ \forall \eta_1 \in \mathrm{Ker}\ B_h$). We set $\eta_h = \widehat{p}_h - p_h$. The term $|\eta_1|_H$ is already estimated from the inequality (4.11); therefore, in order to get the second inequality (4.12), we only need to estimate $\|\eta_2\|_X$. To do so, we start by recalling that the inf-sup condition (Hypothesis (H2)) is equivalent to

$$\text{(4.19)} \qquad \left\{ \begin{array}{l} \text{there exists a constant } C>0, \text{ independent of } h, \text{ such that} \\ \\ \forall\ q_h^s \in X_h^s, \quad \sup\limits_{w_h \in M_h} \dfrac{b(w_h, q_h^s)}{\|w_h\|_M} \geq C\ \|q_h^s\|_{X/\mathrm{Ker} B_h}. \end{array} \right.$$

We also know that $\|\eta_2\|_X = \|\eta_h^s\|_{X/\mathrm{Ker}\ B_h}$. Thus, by taking $q_h^s = \eta_2$ in (4.19),

$$\sup_{w_h \in M_h} \frac{b(w_h, \eta_2)}{\|w_h\|_M} \geq C\ \|\eta_2\|_X.$$

Now use the second equation of (4.16) to obtain

$$\text{(4.20)} \qquad \|\eta_2\|_X \leq C' \left\{ \|\partial_t (v - \widehat{v}_h)\|_M + \|\partial_t (\widehat{v}_h - v_h)\|_M \right\}.$$

Till now, we only have used the C^1 regularity of the solution. In order to bound $\|\partial_t(\widehat{v}_h - v_h)\|_M$, we need C^2. Indeed, we want to apply (4.11) with v_h replaced by $\partial_t v_h$, $\widehat{v}_h$ by $\partial_t \widehat{v}_h$, and so on... More precisely, we have

$$\text{(4.21)} \qquad \|\partial_t (\widehat{v}_h - v_h)\|_M(t) \leq C \int_0^t \|\partial_t^2 \delta_h\|_M(s)\ ds.$$

Finally, combining (4.20) and (4.21), we get

$$\|\eta_2\|_X(t) \leq C' \left\{ \|\partial_t \delta_h\|_M(t) + \int_0^t \|\partial_t^2 \delta_h\|_M(s)\ ds \right\},$$

and the proof is achieved. □

Proof. (of Theorem 4.3). We combine results given in Lemma 4.2 and in Lemma 4.4.

- Estimates in H and M-norms. We have

$$\text{(4.22)} \qquad |p - p_h^s|_H + |p_h^r|_H \leq |p - \widehat{p}_h^s|_H + |\widehat{p}_h^r|_H + |\widehat{p}_h^s - p_h^s|_H + |\widehat{p}_h^r - p_h^r|_H.$$

The first term on the right hand side is bounded by $|||p - \widehat{p}_h|||$ and thus by $|||(p,u) - \Pi_h(p,u)|||_C$. Assuming that $(p,v) \in \mathcal{C}^{1,0} \times C^1(0,T;M)$, by (4.5) it can be estimated:

$$\text{(4.23)} \qquad |||p - \widehat{p}_h|||\,(t) \leq C\ \mathcal{E}_h(p,u)(t).$$

The second term is estimated using (4.11) of Lemma 4.4. This requires $\|\partial_t \delta_h\|_M$ to be estimated. For this, we use (4.6), for $k=2$, which requires $(p,v) \in C^2(0,T;X) \times C^1(0,T;M)$. We get

$$\text{(4.24)} \qquad \|\partial_t (v - \widehat{v}_h)\|_M(s) \leq C\ \mathcal{E}_h\left(\partial_t^2 p, \partial_t v\right)(s);$$

then (4.22), (4.23) and (4.24) lead to the first inequality of (4.9). Now, for v, we write

$$\|v - v_h\|_M \leq \|v - \widehat{v}_h\|_M + \|\widehat{v}_h - v_h\|_M.$$

We apply (4.6) for $k=1$ and get

$$\|v-\widehat{v}_h\|_M \leq C\,\mathcal{E}_h\left(\partial_t p, v\right).$$

Using again estimate (4.11) to bound $\|\widehat{v}_h - v_h\|_M$, we easily get the second inequality of (4.9).

• Estimate in X-norm. We write

$$\|p-p_h^s\|_X \leq \|p-\widehat{p}_h^s\|_X + \|\widehat{p}_h^s - p_h^s\|_X.$$

The first term is again bounded by $|||p-\widehat{p}_h|||$ and can be estimated from (4.5). For the second term, we now use (4.12) of Lemma 4.4 which requires $(p,v)\in \mathcal{C}^{2,1}\times C^2(0,T;M)$ and $(p_h,v_h)\in C^2(0,T;X_h)\times C^2(0,T;M_h)$. In the right hand side of (4.12) appears the second derivative of $v-\widehat{v}_h$. We thus use estimate (4.6) for $k=3$, which requires $(p,v)\in C^3(0,T;X)\times C^2(0,T;M)$, and we get (4.10). ☐

4.2. Application to the approximation of the anisotropic wave equation with the new finite element, $Q_{k+1}^{div}-Q_k$. We come back to the original problem described in §2. We consider the approximate spaces given in (2.16). From §3.3 we know that the $Q_{k+1}^{div}-Q_k$ element is covered by the abstract framework. It is then straightforward to apply Theorem 4.3 and the following theorem results.

THEOREM 4.5. *Let (p,v) be the solution of (2.5), (2.4) and $u\in C^2(0,T;M)$ the primitive of v, defined by*

$$\frac{du}{dt}=v \;\;;\;\; B^t u(0) = -\mathcal{A}p_0.$$

Let (p_h,u_h) be the solution of the approximate problem (2.9) with initial conditions

$$(p_{0,h}, v_{0,h}) = \Pi_h(p_0\;,\;v_0).$$

• Convergence in the L^2-norm :

(i) If $(p,v)\in C^2(0,T;X)\times C^1(0,T;M)$, for all $t\in[0,T]$,

$$|p-p_h^s|_H(t)\to 0\;;\;|p_h^r|_H(t)\to 0\;;\;\;\|v-v_h\|_M(t)\to 0.$$

(ii) Furthermore, if $(p,u)\in C^2(0,T;H^{k+1}(\mathrm{div};\Omega))\times C^2(0,T;H^{k+1}(\Omega))$ and $\mathcal{A}p\in C^2(0,t;(H^{k+1}(\Omega))^2)$, then

$$|p-p_h^s|_H(t)+|p_h^r|_H(t)+\|v-v_h\|_M(t)\leq C_1(t)h^k,$$

with $C_1(t)=O(\|p\|_{C^2(0,t;H^{k+1}(\mathrm{div};\Omega))}+\|\mathcal{A}p\|_{C^2(0,t;(H^{k+1}(\Omega))^2)}+\|u\|_{C^2(0,t;H^{k+1}(\Omega))})$.

• Convergence in the $H(\mathrm{div})$-norm :

(iii) If $(p,v)\in C^3(0,T;X)\times C^2(0,T;M)$, for all $t\in[0,T]$,

$$\|p-p_h^s\|_X(t)\to 0.$$

(iv) Furthermore, if $(p,u)\in C^3(0,T;H^{k+1}(\mathrm{div};\Omega))\times C^3(0,T;H^{k+1}(\Omega))$ and $\mathcal{A}p\in C^3(0,t;(H^{k+1}(\Omega))^2)$, then

$$\|p-p_h^s\|_X(t)\leq C_2(t)h^k,$$

with $C_2(t) = O(\|p\|_{C^3(0,t;H^{k+1}(\mathrm{div};\Omega))} + \|\mathcal{A}p\|_{C^3(0,t;(H^{k+1}(\Omega))^2)} + \|u\|_{C^3(0,t;H^{k+1}(\Omega))})$.

Proof. We apply Theorem 4.3 which relates the errors to quantities such as $\mathcal{E}_h(\partial_t^m p, \partial_t^l u)$; i.e., to the error due to the approximation of $H(\mathrm{div};\Omega)$ with the Raviart-Thomas $RT_{[k]}$ space and to the error due to the approximation of $L^2(\Omega)$ with the Q_k discontinuous elements. ☐

Appendix. Gauss-Lobatto and Gauss-Legendre quadrature formulae in 2-D. We give the locations of the quadrature points and the associated weights used respectively for the $Q_2^{div} - Q_1$ element ($k = 1$, Fig. 2.5) and the $Q_3^{div} - Q_2$ element ($k = 2$, Fig. 2.6).

For $k = 1$ (Fig. 2.5)

- The Gauss-Lobatto points and weights:

$$\begin{cases} x_1^p = y_1^p = 0,\ x_2^p = y_2^p = \dfrac{1}{2},\ x_3^p = y_3^p = 1; \\ \omega_{11}^p = \omega_{31}^p = \omega_{33}^p = \omega_{13}^p = \dfrac{1}{36} \\ \omega_{21}^p = \omega_{32}^p = \omega_{23}^p = \omega_{12}^p = \dfrac{1}{9} \\ \omega_{22}^p = \dfrac{4}{9} \end{cases}$$

- The Gauss-Legendre points and weights:

$$\begin{cases} x_1^v = y_1^v = \dfrac{3-\sqrt{3}}{6},\ x_2^v = y_2^v = \dfrac{3+\sqrt{3}}{6} \\ \omega_{11}^v = \omega_{21}^v = \omega_{22}^v = \omega_{12}^v = \dfrac{1}{4} \end{cases}$$

For $k = 2$ (Fig. 2.6)

- The Gauss-Lobatto points and weights:

$$\begin{cases} x_1^p = y_1^p = 0,\ x_2^p = y_2^p = \dfrac{5-\sqrt{5}}{10},\ x_3^p = y_3^p = \dfrac{5+\sqrt{5}}{10},\ x_4^p = y_4^p = 1 \\ \omega_{11}^p = \omega_{41}^p = \omega_{44}^p = \omega_{14}^p = \dfrac{1}{144} \\ \omega_{21}^p = \omega_{31}^p = \omega_{42}^p = \omega_{43}^p = \omega_{34}^p = \omega_{24^p} = \omega_{13}^p = \omega_{12}^p = \dfrac{5}{144} \\ \omega_{22}^p = \omega_{32}^p = \omega_{33}^p = \omega_{23}^p = \dfrac{25}{144} \end{cases}$$

- The Gauss-Legendre points and weights:

$$\begin{cases} x_1^v = y_1^v = \dfrac{5-\sqrt{15}}{10},\ x_2^v = y_2^v = \dfrac{1}{2},\ x_3^v = y_3^v = \dfrac{5+\sqrt{15}}{10}. \\ \omega_{11}^v = \omega_{31}^v = \omega_{33}^v = \omega_{13}^v = \dfrac{25}{324} \\ \omega_{21}^v = \omega_{32}^v = \omega_{23}^v = \omega_{12}^v = \dfrac{10}{81} \\ \omega_{22}^v = \dfrac{16}{81} \end{cases}$$

REFERENCES

[1] I. Babuska, *The Finite Element Method with Lagrangian Multipliers*, Numer. Math., 20 (1973), pp. 179–192.

[2] E. Bécache, P. Joly, and C. Tsogka, *Eléments finis mixtes et condensation de masse en élastodynamique linéaire. (I) Construction*, C.R. Acad. Sci. Paris, t. 325, Série I (1997), pp. 545–550.

[3] ———, *Etude d'un nouvel élément fini mixte permettant la condensation de masse*, C.R. Acad. Sci. Paris, t. 324, Série I (1997), pp. 1281–1286.

[4] J. Bérenger, *A perfectly matched layer for the absorption of electromagnetic waves*, Journal of Comp. Physics., **114** (1994), pp. 185–200.

[5] F. Brezzi, *On the existence uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers*, Revue Française d'Automatique, Informatique et Recherche Opérationnelle, R-2 (1974), pp. 129–151.

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

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

[8] G. Cohen and P. Joly, *Construction and Analysis of Fourth-Order Finite Difference Schemes for the Acoustic Wave Equation in Nonhomogeneous Media* , SIAM J. on Num. Anal., 33 (1996), pp. 1266–1302.

[9] G. Cohen, P. Joly, and N. Tordjman, *Construction and analysis of higher order finite elements with mass lumping for the wave equation*, in Proc. of the 2nd Int. Conf. on Math. and Num. Aspects of Wave Propagation, SIAM, ed., Delaware, USA, 1993, pp. 152–160.

[10] G. Cohen and P. Monk, *Gauss point mass lumping schemes for Maxwell's equations*, Numer. Methods Partial Differential Equations, 14 (1998), pp. 63–88.

[11] F. Collino, *Perfectly matched absorbing layers for the paraxial equations*, J. Comput. Phys., 131 (1996), pp. 164–180.

[12] F. Collino, P. Joly, and F. Millot, *Fictitious domain method for unsteady problems: Application to electromagnetic scattering*, Tech. Report 2963, INRIA, Août 1996.

[13] F. Collino and P. Monk, *Conditions et couches absorbantes pour les équations de Maxwell*, in Aspects récents en méthodes numériques pour les équations de Maxwell, G. Cohen and P. Joly, eds., Ecole des Ondes, INRIA, Rocquencourt, 1998, ch. 4.

[14] F. Collino and C. Tsogka, *Application of the PML Absorbing Layer Model to the Linear Elastodynamic Problem in Anisotropic Heteregeneous Media*, Rapport de Recherche 3471, INRIA., 1998. submitted to Geophysics.

[15] T. Dupont, *L^2-estimates for Galerkin methods for second order hyperbolic equations*, SIAM J. Numer. Anal., 10 (1973), pp. 880–889.

[16] A. Elmkies and P. Joly, *Éléments finis et condensation de masse pour les équations de Maxwell: le cas 2D*, Tech. Report 3035, INRIA, Novembre 1996.

[17] G.A. Baker and V.A. Dougalis , *The effect of quadrature errors on finite element approximations for the second-order hyperbolic equations*, SIAM Jour. Numer. Anal., 13 (1976), pp. 577–598.

[18] V. Girault and R. Glowinski, *Error analysis of a fictitious domain method applied to a Dirichlet problem*, Japan J. Indust. Appl. Math., 12 (1995), pp. 487–514.

[19] J. C. Nédélec, *A new family of mixed finite elements in* $\mathbb{R}^3$, Numer. Math., 50 (1986), pp. 57–81.

[20] J.E. Roberts and J. M. Thomas, *Mixed and Hybrid Methods, in Handbook of Numerical Analysis*, vol. 2, Finite Element Methods(Part 1), P.G. Ciarlet and J.L. Lions eds., 1989.

[21] P. A. Raviart and J. M. Thomas, *A mixed finite element method for 2nd order elliptic problems*, in Lecture Notes in Mathematics, no. 606, Proc. of math. aspects on the finite element method, Berlin Heidelberg New York: Springer, 1977, pp. 292–315.

[22] R. Glowinski and T.W. Pan and J. Periaux, *A fictitious domain method for Dirichlet problem and applications*, Comp. Meth. in Appl. Mech. and Engin., 111 (1994), pp. 283–304.

[23] R.M. Alford and K.R. Kelly and D.M. Boore, *Accuracy of finite difference modeling of the acoustic wave equation*, Geophysics, 39 (1974), pp. 834–842.

[24] G. Seriani and E. Priolo, *Spectral element method for acoustic wave simulation in heterogeneous media*, in Proc. of ICOSAHOM'92, Montpellier, France, 1992.

[25] N. Tordjman, *Eléments finis d'ordre élevé avec condensation de masse pour l'équation des ondes*, PhD thesis, Univ. Paris IX, 1995.

[26] W. A. Mulder, *A comparison between higher-order finite elements and finite differences for solving the wave equation*, Num. Meth. in Eng., (1996), pp. 344–350.