Numerical Solution of the Pressure Equation in a Simple Model of Aluminium DC-casting

Thesis
for the Cand. Scient. Degree
in Informatics

Wen Shen
Institute of Informatics
University of Oslo
and
Section for Applied Mathematics, SINTEF

April 1994, Oslo, Norway

Acknowledgement

This paper is my thesis for the Cand. Scient. degree in Informatics. The presented work began at Autumn of 1993, and was undertaken at Section for Applied Mathematics, SINTEF. The programming implementation is carried out in DIFFPACK.

First of all, my gratitude should be dedicated to my supervisor, Professor ***Aslak Tveito*** at Institute of Informatics, University of Oslo / SINTEF. I want to take this opportunity to thank him for offering me this great chance to work my thesis at SINTEF. I am grateful for the inspiration, the encouragement and the supervision through out the last year.

My boyfriend ***Xing Cai*** has really been a great help to me. While working on his own Cand. Scient. thesis at the same time, he has not only shared the daily work with me, but also used his intelligence to inspire me with my study. Those nice plots of streamlines are produced with his assistance.

I would also like to thank all those people who have cared and helped me during my work. Without them, things would have be different. To ***Are Magnus Bruaset*** at SINTEF, I thank him for all the help I got from him. He also read and modified the Chapter on mixed methods. Thanks a lot. To ***Torgeir Rusten*** at NR, I thank him for his theoretical direction on the implementation of mixed FEM. To ***Hans Petter Langtangen*** at Institute of Mechanics, UIO / SINTEF, I really admire his enthusiasm for work, and I feel so lucky that I could have the chance to use such a great software package like DIFFPACK. I thank him a lot for guiding me with the implementations. To ***Marit Larsen***, who worked at SINTEF and works at IFE now, I thank her for getting me started with DIFFPACK. I thank also ***Håvard Thevik*** at SINTEF for defining the aluminium DC-casting model, and ***Bjørn Fredrik Nielsen*** for all those nice chats.

The last year at Section of Applied Mathematics, SINTEF is very wonderful for me. I hereby thank the personnel at this Section for being so kind to me and making it one of the most memorable years in my life.

And last, but certainly not least, I want to thank my family in China — my parents and my younger brother, who support and encourage me throughout my life, who enjoy and be proud of every little success of mine. I love you all so much!

Wen Shen

Oslo, Norway

April, 1994

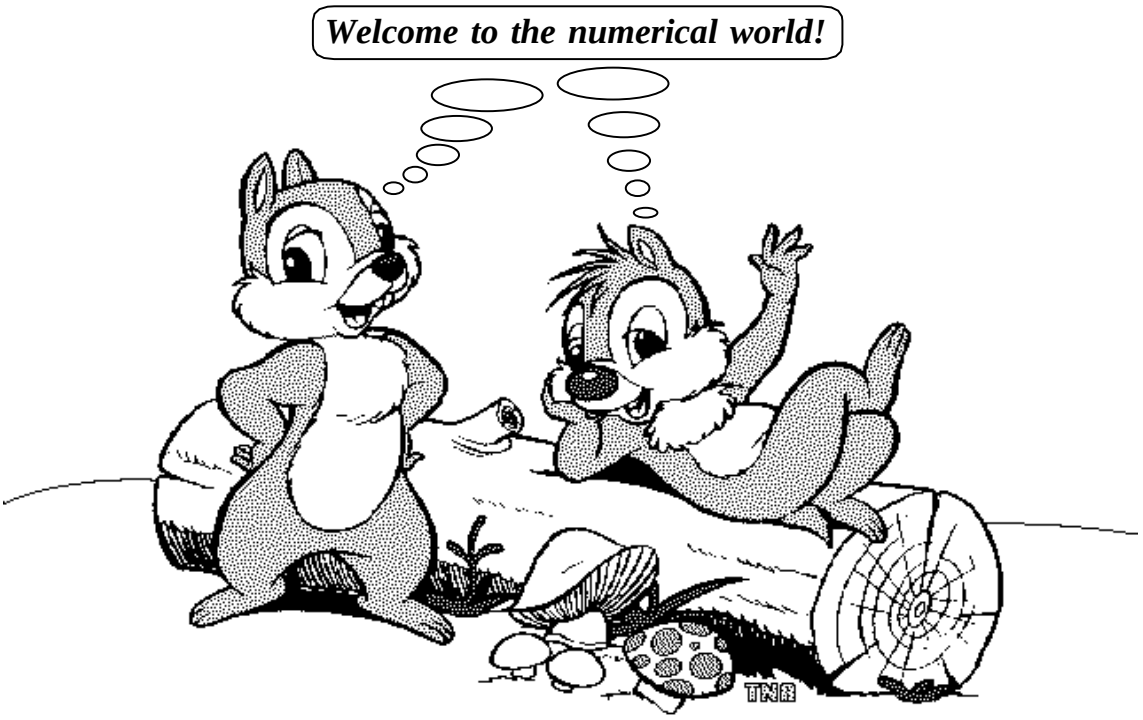
Contents

List of Figures

List of Tables

Chapter 1

Introduction



The main objective of this thesis is to study the numerical solutions of the second-order elliptic boundary value problems. Two different methods will be used, namely the conforming finite element element method and the mixed finite element method. Throughout the thesis, the abbreviation "FEM" stands for "finite element method".

Consider the elliptic model problem

$$-\nabla\cdot(\lambda(\nabla p+\mathbf{E}))=f \qquad \text{in } \Omega\in\mathbb{R}^2, \tag{1.1}$$

subject to the boundary conditions

$$p = g_p \qquad \text{on } \partial\Omega_1, \tag{1.2}$$

$$\mathbf{v}\cdot\mathbf{n} = g_v \qquad \text{on } \partial\Omega_2, \tag{1.3}$$

where $\mathbf{v}=-\lambda\left(\nabla p+\mathbf{E}\right)$. Here $\partial\Omega=\partial\Omega_1\cup\partial\Omega_2$ is the boundary of Ω with $\partial\Omega_1\cap\partial\Omega_2=\emptyset$, and $\partial\Omega_1$ and $\partial\Omega_2$ can be unions of disjointed boundary segments, not necessarily connected. Moreover, $\mathbf{E}$ is a vector-valued function. The solution p can be regarded as a model of pressure, for which λ is the mobility (or permeability). To ensure that the equation remains elliptic, we assume that λ has the same sign throughout Ω and satisfies the condition $|\lambda|\geq e$ for some positive value e at every point in Ω. The uniqueness of solution for the problem is dependent on the boundary conditions. In other words, the problem has a unique solution when $\partial\Omega_1$ is nonempty; otherwise, the uniqueness is lost since $p+c$ is a solution for any constant c provided p is a solution. Note also that with $\lambda\equiv 1$ and $\mathbf{E}\equiv 0$, equation (1.1) is reduced to a Poisson equation.

Elliptic equations can be solved with both conforming and mixed FEM. These two methods are discussed in detail in the Thesis. We compare these two methods in the context of the formulations, the rates of convergence in different norms and the implementations. Numerous case studies are carried out especially for the study of the rates of convergence. We also apply the numerical software to a highly non-trivial model problem related to the aluminium DC-casting surface segregation process, with the computing results being presented graphically.

For elliptic problems, singularities in the solutions can be caused by special λ functions. The main problem in the DC-casting model from a numerical point of view is that the permeability function is singular, i.e., it is zero on one part of the boundary and infinite on another part of the boundary. This phenomenon is studied in numerical detail. Moreover, a new method for estimating the rates of convergence in complicated problems is introduced and applied to the DC-casting model.

The numerical software is implemented in `DIFFPACK`[1], using the object-oriented programming language C++. For the conforming FEM, the class hierarchy is already established in `DIFFPACK`, cf. [6], [7], [8] and [9]. Based on the implementation for the conforming FEM, I implement the mixed FEM in a similar style.[2] However, this version is preliminary in the `DIFFPACK` library, and might be replaced by a more general code in the future.

The content of the thesis can be summarized as the following:

[1] The development of `DIFFPACK` is supported by The Research Council of Norway through the research program no. STP 28402: *Toolkits in Industrial Mathematics at SINTEF.*

[2] The work is supported by The Research Council of Norway through program no. STP.29643, at Section for Applied Mathematics, SINTEF, Oslo, Norway.

⋆ **Chapter 2:**

The aluminium DC-casting surface segregation model is described in this Chapter. The derivation of the differential equations is briefly discussed, together with the parameters and the corresponding boundary conditions. The basis of the numerical solution is obtained at the end of the Chapter.

⋆ **Chapter 3:**

Conforming FEM is described in this Chapter by proceeding the variational formulation of a general elliptic problem. The implementation of the method in `DIFFPACK` is also explained. We demonstrate the method with some case studies, and conclude the Chapter by presenting the graphical solutions of the aluminium DC-casting problem.

⋆ **Chapter 4:**

The other numerical method, which is called the mixed FEM, is described and applied to the elliptic problems. The basic theory of the method is presented by constructing the mixed formulation for a general elliptic problem. After a discussion of the object-oriented implementation of mixed FEM in `DIFFPACK`, we apply this software to some numerical experiments. We close the case study by solving the aluminium DC-casting model problem.

⋆ **Chapter 5:**

In this Chapter, we compare the conforming and mixed FEM by studying the similarities and differences in the formulations, the rates of convergence, the implementations and some physical aspects. Several numerical experiments are executed to study the rates of convergence before the summarizing remarks are given.

⋆ **Chapter 6:**

The difficulties arises in the aluminium DC-casting model problem are discussed in this Chapter. First, we set up a simple one-dimensional problem and study several relevant cases regarding the singularities in the solutions. Then, we introduce a new method for estimating the rates of convergence for the complicated problems whose analytical solutions are not available. Finally, we apply the method to our aluminium DC-casting problem.

Chapter 2

Model Description

Don't ask me...

2.1 Introduction

The problem defined here is taken from an aluminium DC-casting[1] process, cf. Figure 2.1. The solution domain for the mathematical model corresponds to the mushy-zone in the casting process. The mushy-zone is a region in space where the solid and liquid phases of aluminium co-exist. We quantify the amount of liquid aluminium at any point in the mushy-zone by the volume fraction of liquid g_l. The value of g_l varies from 0 to 1, by which $g_l = 0$ means solid, and $g_l = 1$ means liquid. Bigger g_l value indicates higher liquid percentage.

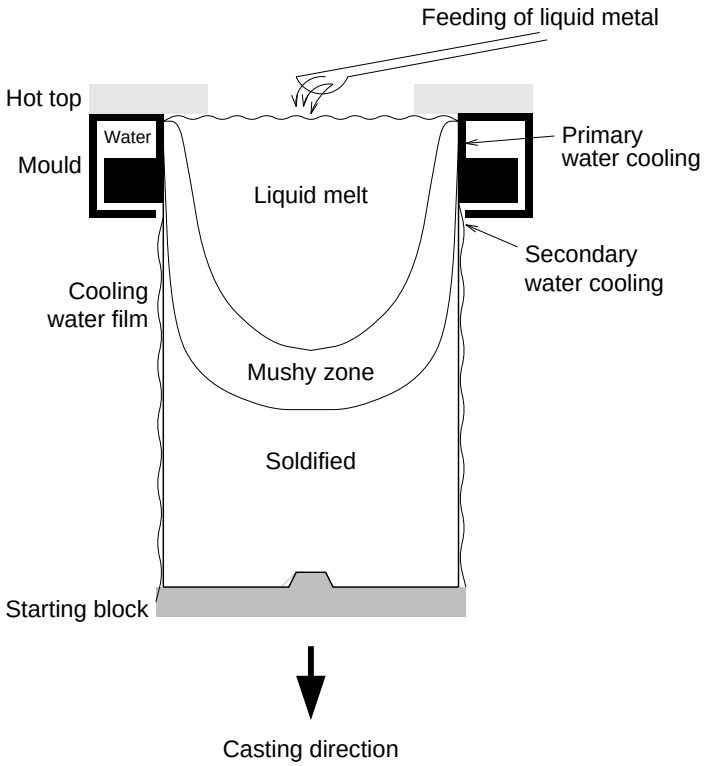


Figure 2.1: The aluminium DC-casting process.

It is regarded that the melt consists of two components, where pure aluminium is the major part (95 % of weight). The component which constitute the other 5 % of the melt, might, during the solidification, become unevenly distributed in space. To quantify the concentration of the "5 %"-constituent we introduce the species mass fraction c_k. A non-uniform spatial distribution of this quantity is called macro-segregation. In our model, metallostatic overpressure causes convection of species-rich melt towards the surface of the cast aluminium. This leads to a variant of macro-segregation named "surface segregation". Therefore, we study the movement in the mushy-zone so that better understanding of the melt on the surface can later be obtained. The convection of the liquid phase is governed by pressure gradients, gravity and interaction with the solid phase. The solid phase is considered as a fine network of solidified aluminium.

In Section 2.2, the governing equations are briefly described, while in Section 2.3, the

[1] The abbreviation "DC" stands for "direct chilling".

mathematical model problem which will be solved numerically is defined. During the derivation of the mathematical model, we assume that the process has reached a stationary state, i.e., all the parameters are independent of time. The two-phase volume-averaged conversation equations, which we use to deduce the mathematical model, are *not* derived here. The reader is referred to Ni and Beckermann [16].

2.2 The Conservation Equations

The governing equations of the model are the mass and linear momentum conservation equations. The following list shows all the parameters that appear in the equations.

Nomenclature

ρ_k	Density of phase k ($k = \{l, s\}$)	constant
g_k	Volume fraction of phase k	
$\vec{v}_k$	Velocity of phase k	
$\vec{v}_r$	Difference between average velocities in liquid and solid (equation (2.2))	
μ	Dynamic liquid viscosity	constant
K	Permeability (equation (2.9))	
p	Pressure	
p_l	Intrinsic liquid pressure	
$\vec{g}$	Acceleration due to gravity	constant
V_s	Magnitude of the casting speed	constant
g	Magnitude of the gravity constant	constant
u	Horizontal component of the relative velocity $\vec{v}_r$	
v	Vertical component of the relative velocity $\vec{v}_r$	
p_0	Metallostatic (over)pressure	constant
γ	Material constant for the permeability (equation (2.9))	constant
Γ_i	Solution domain boundary	
Y_i	x_2–value of boundary i for a given x_1-value	
L	Horizontal length of the solution domain	constant
H_c	Height of the solution domain at $x = 0$	constant
H_s	Highest y–value of the solution domain (at surface)	constant
L_2	Length of the Γ_2 boundary ($p = 0$)	constant
L_3	Length of the Γ_3 boundary ($u = 0$)	constant

2.2.1 Mass conservation

By adding the macroscopic mass conservation equations for the solid and liquid phases together, the following equation is obtained:

$$\frac{\partial}{\partial t}(\rho_l g_l + \rho_s g_s) + \nabla \cdot (\rho_l g_l \vec{v}_l + \rho_s g_s \vec{v}_s) = 0. \tag{2.1}$$

Here, ρ_k, g_k and $\vec{v}_k$ are the density, the volume fraction and the velocity of phase k within a small volume element (the averaging volume), respectively. Throughout this Chapter, subscript k means either l for liquid or s for solid.

By assuming that

- $\rho_l = \rho_s = \rho =$ constant,
- $\vec{v}_s =$ constant ,
- $g_l + g_s = 1$, (i.e. no pore formation)

and defining the relative velocity, $\vec{v}_r$, between the phases as

$$\vec{v}_r \equiv \vec{v}_l - \vec{v}_s, \tag{2.2}$$

equation (2.1) can be reduced to

$$\nabla \cdot (g_l \vec{v}_r) = 0, \tag{2.3}$$

where we have used the assumption that a stationary state is reached.

2.2.2 Momentum conservation

The conservation of momentum in the liquid phase is described by a Darcy equation, i.e.,

$$\frac{g_l \mu}{K}(\vec{v}_l - \vec{v}_s) + \nabla p_l - \rho_l \vec{g} = 0, \tag{2.4}$$

where μ, K and p are the dynamic viscosity, the permeability of the dendritic network, and the pressure, respectively. The permeability K, which is a function of g_l, i.e. $K = K(g_l)$, is a property associated with the material, it measures the state of being passed through by liquid or gas, etc. High value of K implies that liquid or gas can easily pass through the material, while low value of K indicates just the opposite. Not that the permeability is related to the volume fraction of liquid g_l, i.e., bigger g_l values give higher K values. In the solidified phase, the permeability is zero, while infinite permeability is expected in the liquid phase. These situations cause difficulties in solving the problem numerically. We will come to this point again later in the thesis.

By introducing the assumptions from Section 2.2.1, equation (2.4) can be rewritten as

$$\frac{g_l \mu}{K}\vec{v}_r + \nabla p_l - \rho \vec{g} = 0. \tag{2.5}$$

2.3 The Mathematical Model

We study a 2D–model of the mushy-zone in the DC-casting process. Since the projection of intersection is symmetrical in the stationary state (cf. Figure 2.1), we only need to study half part of the mushy-zone. The solution domain for the equations is indicated in Figure 2.2, which is the right-half part. Note also that the mathematical model described here is

only a small part of a big problem, i.e., we only study the pressure equation that is involved in the whole process.

As shown earlier in this Chapter, the volume fraction of liquid g_l appears in the conservation equations, and it will eventually be coupled in the derived pressure equation. In order to solve the pressure equation, we have to "choose" an approximation of g_l as the input parameter to our problem. The reader should be notified that in the whole model problem, g_l is also a main unknown and is governed by the energy conservation equation. We refer to Haug, Mo and Thevik [11] for more details of the complete model problem[2].

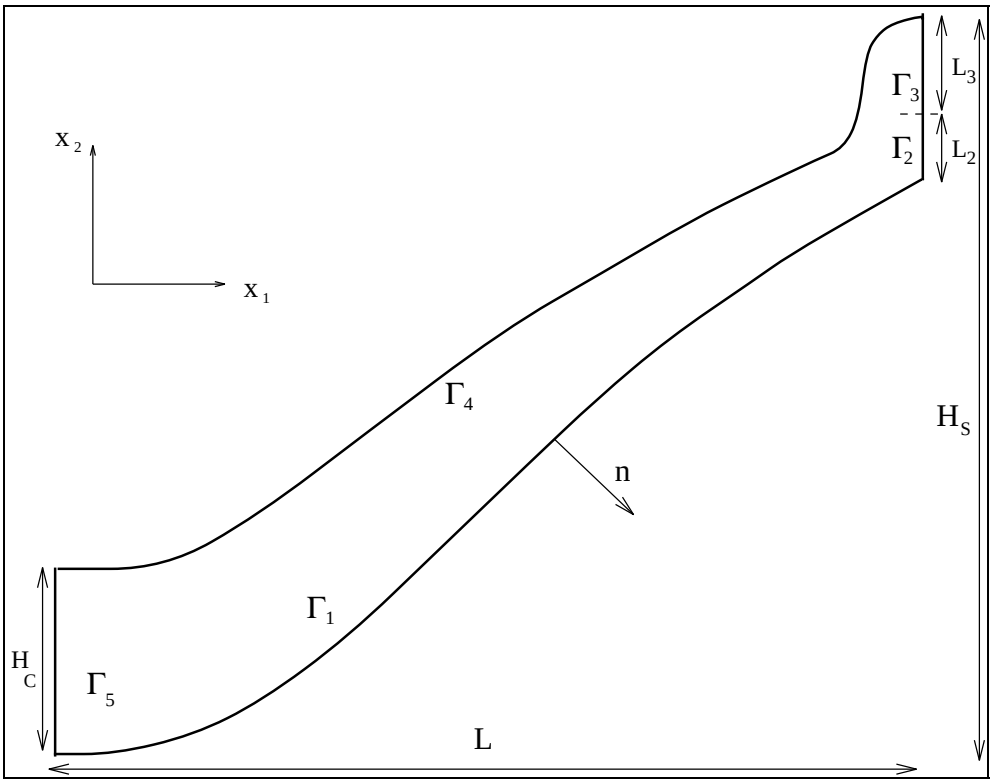


Figure 2.2: The solution domain Ω.

2.3.1 The system equations

The equations (2.3) and (2.5) from the previous section can be simplified by introducing

- $\vec{v}_s = -V_s \vec{\delta}_{i2}$ $(V_s > 0)$, where $\vec{\delta}_{i2} = (0, 1)$ is the unit vector pointing at x_2 direction,
- $\vec{g} = -g\vec{\delta}_{i2}$ $(g > 0)$,
- $\vec{v}_r = [u, v]$,
- $p_l = p$,

[2]The model discussed in the paper is a simplified one-dimensional problem.

and bear the following new forms

$$\frac{\partial}{\partial x_1}(g_l u) + \frac{\partial}{\partial x_2}(g_l v) = 0, \tag{2.6}$$

$$\frac{g_l \mu}{K} u + \frac{\partial p}{\partial x_1} = 0, \tag{2.7}$$

$$\frac{g_l \mu}{K} v + \frac{\partial p}{\partial x_2} + \rho g = 0. \tag{2.8}$$

The permeability, K, is assumed to obey the Kozeny-Carman relation

$$K = \gamma \frac{g_l^3}{(1 - g_l)^2}, \tag{2.9}$$

where γ is a constant, and the volume fraction of liquid g_l is given as input parameter to the problem. The plot of the function $K(g_l)$ with $\gamma = 1$ is included in Figure 2.3.

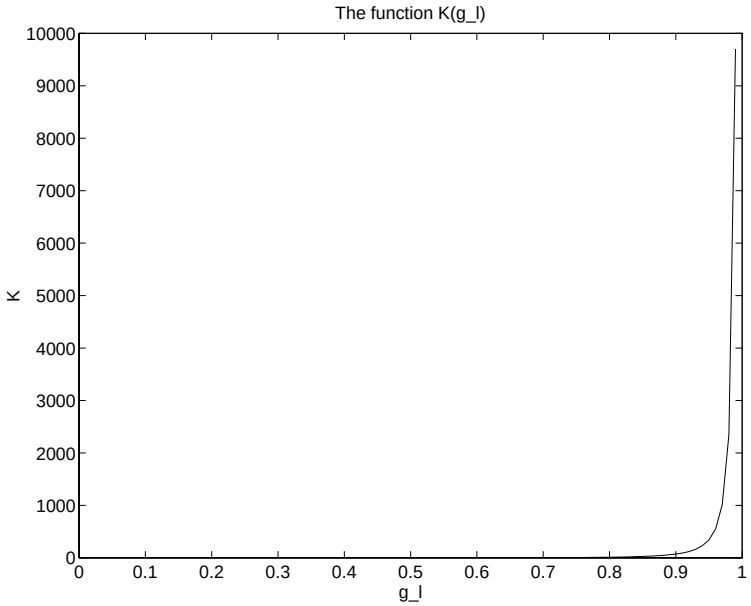


Figure 2.3: The function $K(g_l)$ with $\gamma = 1$.

2.3.2 The boundary conditions

The boundary of the solution domain is divided into five segments, $\Gamma_i, i = 1, \ldots, 5$ (cf. Figure 2.2). The conditions on each part of the boundary are given in Table 2.1, where $\mathbf{v} = (g_l u, g_l v)$ is the relative superficial velocity, and the vector $\mathbf{n}$ is the outwards directed normal vector at a boundary point. In addition, for a given x_1-coordinate, Y_i is the x_2-coordinate of the boundary segment Γ_i.

Here are the explanations of these boundary conditions:

- The boundary segment Γ_1 separates the mushy- and the solidified-zone. Since the solid aluminium is not permeable, we have $\mathbf{v} \cdot \mathbf{n} = 0$ here.

Segment	*Boundary Conditions*
Γ_1	$\mathbf{v} \cdot \mathbf{n} = 0$
Γ_2	$p = 0$
Γ_3	$\mathbf{v} \cdot \mathbf{n} = 0$
Γ_4	$p = p_0 + \rho g(H_s - Y_4)$
Γ_5	$\mathbf{v} \cdot \mathbf{n} = 0$

Table 2.1: Boundary conditions.

- The boundary segment Γ_2 is related to the air gap between the semi-solid shell and the mould, cf. Figures 2.1 and 2.4. If we neglect the atmosphere pressure, we have $p = 0$ here.

- The boundary segment Γ_3 is connected to the mould, so it is clear we have $\mathbf{v} \cdot \mathbf{n} = 0$ here.

- The boundary segment Γ_4 separates the liquid- and the mushy-zone. Since we neglect the atmosphere pressure and assume that the liquid density ρ is a constant, the pressure at a point on this segment is thus $\rho g H$, where g is the gravity acceleration constant and H is the height from this point to the surface of the liquid. Let p_0 denote the pressure at the top point of Γ_4, we have $p = p_0 + \rho g(H_s - Y_4)$ at this segment of the boundary.

- The boundary segment Γ_5 is the symmetry line. If we assume continuous velocity solution, we have $\mathbf{v} \cdot \mathbf{n} = 0$ here.

Remarks: Note that the boundary condition at Γ_2 is $p = 0$. By imposing this condition, we allow the liquid velocity to take a non-zero value at this boundary, i.e., liquid flows out of the solution domain at Γ_2 and causes surface segregation, cf. Figure 2.4. This flow phenomenon is referred to as *exudation*, cf. [4, page 252]. The surface layer being formed by the exuding interdendritic liquid is highly enriched in alloy element. Unfortunately, only the melt which includes 5% alloy element has the best physical qualities, so this surface layer must be removed before the aluminium is processed further. In industry this operation is a very expensive part of the total process. Therefore, the area close to Γ_2 is called the *critical region* of the problem.

2.3.3 Material and process parameters

There are three parameters which depend solely on the properties of the materials that are involved. The values of these parameters are given in Table 2.2. Geometrical data for the solution domain of an actual DC-casting process problem are summarized in Table 2.3, as well as other process specific information. For the choices of these parameters, we refer to [14], [15], [18], and [25].

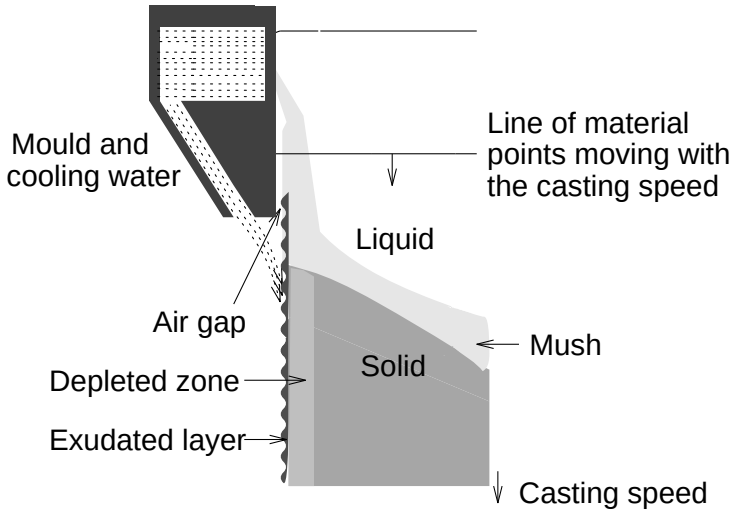


Figure 2.4: The region near the mould in the DC casting process.

Parameter	*Value*
ρ	$2385\ \mathrm{kg/m^3}$
μ	$1.2 \cdot 10^{-3}\ \mathrm{Ns/m^2}$
γ	$10^{-11}\ \mathrm{m^2}$

Table 2.2: The material properties.

Parameter	*Value*
p_0	$1900\ \mathrm{Pa}$
V_s	$7.5 \cdot 10^{-4}\ \mathrm{m/s}$
g	$9.8\ \mathrm{m/s^2}$
L	$0.24\ \mathrm{m}$
H_c	$0.1028\ \mathrm{m}$
H_s	$0.5\ \mathrm{m}$
L_2	$0.0469\ \mathrm{m}$
L_3	$0.0611\ \mathrm{m}$

Table 2.3: Process parameters.

2.3.4 Dimensionless equations

To facilitate further analysis and acquire numerical solution of the boundary value problem defined by equations (2.6)-(2.8) with the boundary conditions summarized in Table 2.1, it is convenient to pose the boundary value problem using dimensionless quantities. To this end, we introduce

$$\begin{aligned} x &= x'H_s,\ y = y'H_s, \\ u &= u'V_s,\ v = v'V_s, \\ p &= p'(p_0 + \rho g H_s), \\ K &= K'\gamma, \\ A &= \frac{\mu H_s V_s}{\gamma(p_0 + \rho g H_s)}, \\ B &= \frac{\rho g H_s}{p_0 + \rho g H_s}, \end{aligned}$$

where all quantities with a prime are dimensionless.

By using these definitions and equation (2.9), equations (2.6)-(2.8) can be rewritten in a dimensionless form, cf. Ni and Beckermann [16]. Hereafter, we drop the prime on the symbols, i.e. $x' \rightarrow x, u' \rightarrow u$ and so on. This leads to

$$\frac{\partial}{\partial x_1}(g_l u) + \frac{\partial}{\partial x_2}(g_l v) = 0, \tag{2.10}$$

$$A\frac{g_l u}{K} + \frac{\partial p}{\partial x_1} = 0, \tag{2.11}$$

$$A\frac{g_l v}{K} + \frac{\partial p}{\partial x_2} + B = 0. \tag{2.12}$$

Hence, the dimensionless permeability becomes $K(g_l) = \dfrac{g_l^3}{(1-g_l)^2}$. By choosing the parameters as in Tables 2.2 and 2.3, we have

$$A = 3.3121113 \quad \text{and} \quad B = 0.860155.$$

Furthermore, it is easy to express the boundary values in Table 2.1 in a dimensionless form, which is given in Table 2.4.

Segment	*Boundary conditions*
Γ_1	$\mathbf{v}\cdot\mathbf{n} = 0$
Γ_2	$p = 0$
Γ_3	$\mathbf{v}\cdot\mathbf{n} = 0$
Γ_4	$p = 1 - BY_4$
Γ_5	$\mathbf{v}\cdot\mathbf{n} = 0$

Table 2.4: Boundary conditions for the dimensionless equations.

2.3.5 Basis of the numerical solution

By combining equations (2.10)-(2.12), we get the following equation

$$\frac{\partial}{\partial x_1}\left(\frac{K}{A}\frac{\partial p}{\partial x_1}\right)+\frac{\partial}{\partial x_2}\left(\frac{K}{A}\left(\frac{\partial p}{\partial x_2}+B\right)\right)=0. \tag{2.13}$$

Introducing the linear operator for any scalar function in 2D

$$\nabla p=(\frac{\partial p}{\partial x_1},\frac{\partial p}{\partial x_2}),$$

and the linear operator for any vector-valued function in 2D

$$\nabla\cdot\mathbf{v}=\frac{\partial v_1}{\partial x_1}+\frac{\partial v_2}{\partial x_2},$$

where $\mathbf{v}=(v_1,v_2)$, we can write the equation (2.13) into the standard compact form as

$$-\nabla\cdot\left(\frac{K}{A}(\nabla p+\mathbf{E})\right)=0 \tag{2.14}$$

where $\mathbf{E}=(0,B)$ is a constant vector.

The equation (2.14) is a typical pressure equation. The boundary conditions for the equation follow from the dimensionless form in Table 2.4 with the relative superficial velocity being decided by the equation

$$\mathbf{v}=-\frac{K}{A}(\nabla p+\mathbf{E}).$$

We can summarize the boundary conditions as

$$\begin{aligned} p&=g_p \quad \text{on } \partial\Omega_a, \quad \text{where } \partial\Omega_a=\Gamma_2\cup\Gamma_4,\\ \mathbf{v}\cdot\mathbf{n}&=0 \quad \text{on } \partial\Omega_b, \quad \text{where } \partial\Omega_b=\Gamma_1\cup\Gamma_3\cup\Gamma_5. \end{aligned}$$

Here $\partial\Omega_a\cup\partial\Omega_b=\partial\Omega$ is the whole boundary with $\partial\Omega_a\cap\partial\Omega_b=\emptyset$.

Now it remains to find a "good" approximation of the volume fraction liquid g_l, which must be given as the input parameter. Physically, g_l varies from 0 on Γ_1 to 1 on Γ_4, so a linear function in x_2 for a given x_1 will be a natural choice. Hence, we choose a g_l which takes the values ε_1 at Γ_1 and $(1-\varepsilon_2)$ at Γ_4, i.e.,

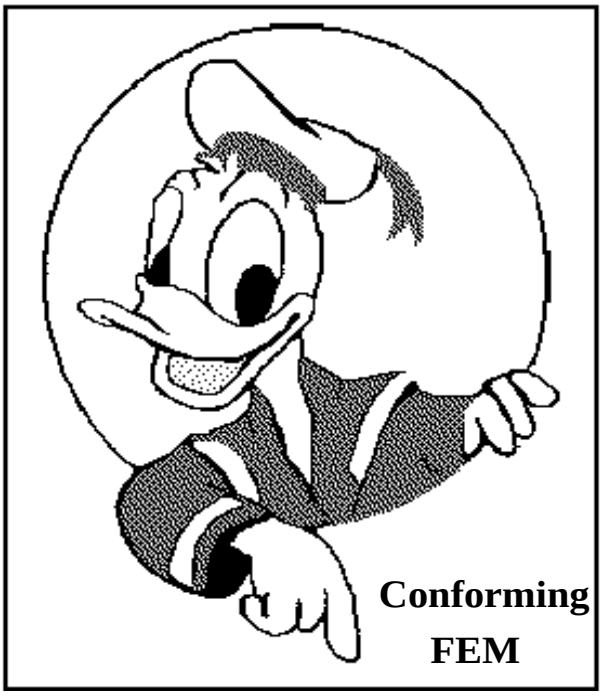
$$g_l(x_1,x_2)=1-\varepsilon_2-\frac{1-\varepsilon_1\varepsilon_2}{Y_4(x_1)-Y_1(x_1)}(Y_4(x_1)-x_2).$$

Here ε_1 and ε_2 are two small positive constants. The reason that we chose g_l in this way, as will be shown in Section 3.1.2, is to obtain a well-posed finite element formulation problem according to the Lax-Milgram Theorem. Finally, we may choose the following values for ε_1 and ε_2:

$$\varepsilon_1=\varepsilon_2=0.01.$$

Chapter 3

Conforming Finite Element Method



Using conforming finite element method to solve elliptic equations is widely discussed in many books and articles, cf [12] and [28]. In this Chapter, we study the numerical solutions of the general second-order elliptic boundary value problems with conforming FEM. The implementation together with the results of some case studies are also included.

3.1 The Finite Element Formulation of the Pressure Equation

In this section, we study the conforming finite element method which is often used in solving second-order elliptic problems. First, a weak formulation of a general elliptic boundary value problem is derived. Then we study whether the variational problem satisfies the conditions in the Lax-Milgram Theorem. The discretized problem is also derived, and we conclude this section with a brief description of the mapping theory used in finite element methods.

3.1.1 The weak formulation

Consider the second-order elliptic boundary value problem (1.1) presented in Chapter 1

$$-\nabla\cdot(\lambda(\nabla p+\mathbf{E}))=f \qquad \text{in } \Omega\in\mathbb{R}^2, \tag{3.1}$$

subject to the boundary conditions

$$p = g_p \qquad \text{on } \partial\Omega_1, \tag{3.2}$$

$$\mathbf{v}\cdot\mathbf{n} = g_v \qquad \text{on } \partial\Omega_2, \tag{3.3}$$

where $\mathbf{v}=-\lambda\left(\nabla p+\mathbf{E}\right)$.

In order to give the proper variational problem, we need to define some basic notations. We define the partial derivative of order α as

$$\mathrm{D}^{\alpha}q=\frac{\partial^{|\alpha|}q}{\partial x_1^{\alpha_1}\partial x_2^{\alpha_2}},$$

where $\alpha=(\alpha_1,\alpha_2)$, α_i is a non-negative integer and $|\alpha|=\alpha_1+\alpha_2$. As an example, a partial derivative of order 2 can then be written as $\mathrm{D}^{\alpha}q$ with $\alpha=(2,0)$, $\alpha=(1,1)$ or $\alpha=(0,2)$, for which $|\alpha|=2$.

Given an integer $m\geq 0$, the Sobolev spaces are then given by

$$H^m(\Omega)=\left\{q\in L^2(\Omega);\mathrm{D}^{\alpha}q\in L^2(\Omega),|\ \alpha\ |\leq m\right\},\ m=1,2,\cdots,$$

with associated norm and seminorm

$$\|q\|_{m,\Omega} = \left(\sum_{|\alpha|\leq m}\int_{\Omega}\left|\mathrm{D}^{\alpha}q\right|^2dx\right)^{\frac{1}{2}},$$

$$|q|_{m,\Omega} = \left(\sum_{|\alpha|=m}\int_{\Omega}\left|\mathrm{D}^{\alpha}q\right|^2dx\right)^{\frac{1}{2}}.$$

In order to formulate a variational problem, we also introduce the usual scalar product

$$< v, w >= \int_\Omega v(x)w(x)\, dx.$$

Let q be the test function that belongs to a Sobolev subspace which will be specified later. By multiplying the equation (3.1) with q and integrate over Ω, we get

$$< -\nabla \cdot (\lambda(\nabla p + \mathbf{E})), q >=< f, q > .$$

Using Green's formula, we have

$$\begin{aligned} < -\nabla \cdot (\lambda(\nabla p + \mathbf{E})), q > &= -\int_\Omega \nabla \cdot (\lambda(\nabla p + \mathbf{E}))q\, dx \\ &= \int_\Omega \nabla q \cdot (\lambda(\nabla p + \mathbf{E}))\, dx - \int_{\partial\Omega} q(\lambda(\nabla p + \mathbf{E})) \cdot \mathbf{n}\, ds \\ &= \int_\Omega \nabla q \cdot (\lambda(\nabla p + \mathbf{E}))\, dx + \int_{\partial\Omega} q\mathbf{v} \cdot \mathbf{n}\, ds. \end{aligned}$$

Since the boundary conditions are given as $p = g_p$ on $\partial\Omega_1$ and $\mathbf{v} \cdot \mathbf{n} = g_v$ on $\partial\Omega_2$, it is natural that we define these subspaces

$$\begin{aligned} H^1(\Omega) &= \left\{p : p \in L^2(\Omega);\, \nabla p \in \left(L^2(\Omega)\right)^2\right\}, \\ H^1_0(\Omega) &= \{p : p \in H^1(\Omega);\, p = 0 \text{ on } \partial\Omega_1\}, \end{aligned}$$

and the linear variety

$$H^1_g(\Omega) = \{p : p \in H^1(\Omega);\, p = g_p \text{ on } \partial\Omega_1\}.$$

Then the following variational problem can be given as

Problem (V): Find $p \in H^1_g(\Omega)$ such that

$$a(q, p) = L(q) \qquad \forall q \in H^1_0(\Omega), \tag{3.4}$$

where

$$a(q, p) = \int_\Omega \nabla q \cdot (\lambda \nabla p)\, dx$$

is a bilinear form on $H^1_0 \times H^1_0$, and

$$L(q) = \int_\Omega fq\, dx - \int_\Omega \nabla q \cdot (\lambda \mathbf{E})\, dx - \int_{\partial\Omega_2} g_v q\, ds$$

is a linear form on H^1_0.

Since the condition $p = g_p$ on $\partial\Omega_1$ is treated explicitly through the weak formulation, it is often referred to as the *essential* boundary condition. In contrast, the *natural* boundary condition is $\mathbf{v} \cdot \mathbf{n} = g_v$ on $\partial\Omega_2$ which is treated implicitly. The variational problem is also referred to as the weak formulation.

3.1.2 Stability test

Existence of a unique weak solution of the equation follows from the Lax-Milgram Theorem in Hilbert space theory. We shall use this Theorem to investigate the well-posedness of our variational problem. In particular, we will use the Theorem to derive the proper conditions on the function λ for a well-posed formulation problem.

In particular, the subspace $H^1(\Omega)$ is associated with the scalar product and the corresponding norm

$$\begin{aligned} < v, w >_{H^1(\Omega)} &= \int_\Omega (vw + \nabla v \cdot \nabla w)\, dx, \\ \|v\|_{H^1(\Omega)} &= < v, v >^{\frac{1}{2}}_{H^1(\Omega)} = \left(\int_\Omega (v^2 + \nabla v \cdot \nabla v)\, dx \right)^{\frac{1}{2}}, \end{aligned}$$

and the semi-norm is given as

$$|v|_{H^1(\Omega)} = \left(\int_\Omega (\nabla v \cdot \nabla v)\, dx \right)^{\frac{1}{2}}.$$

The Lax-Milgram Theorem reads: (cf. Johnson [12])

Theorem 3.1 (Lax-Milgram) *Let V be a Hilbert space with scalar product $(.,.)_V$ and the corresponding norm $\|.\|_V$. Suppose that $a : V \times V \to \mathbb{R}$ is a bilinear form on $V \times V$ and $L : V \to \mathbb{R}$ a linear form on V such that:*

(1) $a(.,.)$ is symmetric, i.e.,

$$a(v, w) = a(w, v) \quad \forall v, w \in V.$$

(2) $a(.,.)$ is continuous, i.e., there is a constant $\gamma > 0$ such that

$$|a(v, w)| \leq \gamma \|v\|_V \|w\|_V \quad \forall v, w \in V.$$

(3) $a(.,.)$ is V-elliptic, i.e., there is a constant $\alpha > 0$ such that

$$a(v, v) \geq \alpha \|v\|^2_V \quad \forall v \in V.$$

(4) $L(.)$ is continuous, i.e., there is a constant $\Lambda > 0$ such that

$$|L(v)| \leq \Lambda \|v\|_V \quad \forall v \in V.$$

Under these conditions, the variational problem: Find $p \in V$ such that

$$a(p, q) = L(q) \quad \forall q \in V$$

has a unique solution $p \in V$ and the following stability estimate holds

$$\|p\|_V \leq \frac{\Lambda}{\alpha}.$$

We now test whether these four conditions are satisfied in our problem.

- **Condition 1:**

  Since λ is a scalar-valued function here, we may write

  $$a(v,w) = \int_\Omega \lambda(\nabla v \cdot \nabla w)dx,$$

  and the symmetry is obvious.

- **Condition 2:**

  In order to increase the readability of the proof, we introduce some notations which only apply in this section:

  $$v_1 = \left\| \frac{\partial v}{\partial x_1} \right\|_{L^2(\Omega)} \quad , \quad v_2 = \left\| \frac{\partial v}{\partial x_2} \right\|_{L^2(\Omega)},$$
  $$w_1 = \left\| \frac{\partial w}{\partial x_1} \right\|_{L^2(\Omega)} \quad , \quad w_2 = \left\| \frac{\partial w}{\partial x_2} \right\|_{L^2(\Omega)}.$$

  Hence we have

  $$\begin{aligned} \|\nabla v\|_{(L^2(\Omega))^2} &= \left(\int_\Omega \nabla v \cdot \nabla v \, dx \right)^{\frac{1}{2}} \\ &= \left(\int_\Omega \left[\left(\frac{\partial v}{\partial x_1} \right)^2 + \left(\frac{\partial v}{\partial x_2} \right)^2 \right] dx \right)^{\frac{1}{2}} \\ &= \left(\left\| \frac{\partial v}{\partial x_1} \right\|^2_{L^2(\Omega)} + \left\| \frac{\partial v}{\partial x_2} \right\|^2_{L^2(\Omega)} \right)^{\frac{1}{2}} \\ &= (v_1^2 + v_2^2)^{\frac{1}{2}}, \end{aligned} \tag{3.5}$$

  and using Cauchy-Schwartz inequality, we have

  $$\begin{aligned} |a(v,w)| &= \left| \int_\Omega \lambda \nabla v \cdot \nabla w \, dx \right| \\ &\le \left| \int_\Omega \lambda \frac{\partial v}{\partial x_1} \frac{\partial w}{\partial x_1} \, dx \right| + \left| \int_\Omega \lambda \frac{\partial v}{\partial x_2} \frac{\partial w}{\partial x_2} \, dx \right| \\ &\le \|\lambda\|_{L^\infty(\Omega)} \left\| \frac{\partial v}{\partial x_1} \right\|_{L^2(\Omega)} \left\| \frac{\partial w}{\partial x_1} \right\|_{L^2(\Omega)} + \|\lambda\|_{L^\infty(\Omega)} \left\| \frac{\partial v}{\partial x_2} \right\|_{L^2(\Omega)} \left\| \frac{\partial w}{\partial x_2} \right\|_{L^2(\Omega)} \\ &= \|\lambda\|_{L^\infty(\Omega)} (v_1 w_1 + v_2 w_2). \end{aligned}$$

  Thus, in order to prove that

  $$|a(v,w)| \le \|\lambda\|_{L^\infty(\Omega)} (v_1 w_1 + v_2 w_2) \le \|\lambda\|_{L^\infty(\Omega)} \|\nabla v\|_{(L^2(\Omega))^2} \|\nabla w\|_{(L^2(\Omega))^2},$$

  it is sufficient to prove

  $$v_1 w_1 + v_2 w_2 \le \|\nabla v\|_{(L^2(\Omega))^2} \|\nabla w\|_{(L^2(\Omega))^2}, \tag{3.6}$$

since $\|\lambda\|_{L^\infty(\Omega)}$ is a positive constant. Using (3.5), we get

$$\begin{aligned}
&\|\nabla v\|_{(L^2(\Omega))^2}\|\nabla w\|_{(L^2(\Omega))^2} \\
&= (v_1^2+v_2^2)^{\frac{1}{2}}(w_1^2+w_2^2)^{\frac{1}{2}} \\
&= (v_1^2w_1^2+v_2^2w_1^2+v_1^2w_2^2+v_2^2w_2^2)^{\frac{1}{2}} \\
&= ([(v_1w_1)^2+2v_1v_2w_1w_2+(v_2w_2)^2]+[(v_2w_1)^2-2v_1v_2w_1w_2+(v_1w_2)^2])^{\frac{1}{2}} \\
&= \left([v_1w_1+v_2w_2]^2+[v_2w_1-v_1w_2]^2\right)^{\frac{1}{2}} \\
&\geq v_1w_1+v_2w_2,
\end{aligned}$$

so the inequality (3.6) holds. Then, since

$$\|\nabla v\|_{(L^2(\Omega))^2} = \left(\int_\Omega \nabla v\cdot\nabla v\,dx\right)^{\frac{1}{2}} \leq \left(\int_\Omega (v^2+\nabla v\cdot\nabla v)\,dx\right)^{\frac{1}{2}} = \|v\|_{H^1(\Omega)},$$

we can reach

$$|a(v,w)| \leq \gamma\|v\|_{H^1(\Omega)}\|w\|_{H^1(\Omega)},$$

where $\gamma = \|\lambda\|_{L^\infty(\Omega)}$.

- **Condition 3:**

We want to prove the inequality

$$\int_\Omega \nabla v\cdot(\lambda\nabla v)\,dx = a(v,v) \geq \alpha\|v\|^2_{H^1(\Omega)} = \alpha\left[\int_\Omega v^2\,dx+\int_\Omega \nabla v\cdot\nabla v\,dx\right], \tag{3.7}$$

it is sufficient to show that

$$\int_\Omega \nabla v\cdot(\lambda\nabla v)\,dx \geq a\int_\Omega \nabla v\cdot\nabla v\,dx,$$

and

$$\int_\Omega \nabla v\cdot(\lambda\nabla v)\,dx \geq b\int_\Omega v^2\,dx.$$

Consequently, (3.7) follows by putting $\alpha = \frac{1}{2}\min(a,b)$.

Let $\lambda_{\min} = \min_{x\in\Omega}\lambda(x) > 0$, then we have

$$\int_\Omega \nabla v\cdot(\lambda\nabla v)\,dx \geq \lambda_{\min}\int_\Omega \nabla v\cdot\nabla v\,dx, \tag{3.8}$$

and $a = \lambda_{\min}$. Using the Poincare's inequality

$$\int_\Omega v^2\,dx \leq C\int_\Omega \nabla v\cdot\nabla v\,dx,$$

for some independent constant C, we get

$$\int_\Omega \nabla v\cdot\lambda\nabla v\,dx \geq b\int_\Omega v^2\,dx, \tag{3.9}$$

where $b = a/C$. Then we have the inequality

$$a(v,v) \geq \alpha\|v\|^2_{H^1(\Omega)},$$

where $\alpha = \frac{1}{2}\min(a,b) = \frac{1}{2}\min(\lambda_{\min},\lambda_{\min}/C)$.

♦ **Condition 4:**

To prove the continuity of $L(.)$ is the same as to prove

$$\left|\int_\Omega f v\,dx\right| \leq a_1\|v\|_{H^1(\Omega)}, \tag{3.10}$$

$$\left|\int_\Omega \nabla v\cdot(\lambda\mathbf{E})\,dx\right| \leq a_2\|v\|_{H^1(\Omega)}, \tag{3.11}$$

$$\left|\int_{\partial\Omega_2} v g_v\,ds\right| \leq a_3\|v\|_{H^1(\Omega)}, \tag{3.12}$$

and it follows that $\Lambda = a_1 + a_2 + a_3$.

By remembering the Cauchy-Schwartz inequality

$$\left|\int_\Omega f v\,dx\right| \leq \|f\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)},$$

it is obvious that $\|v\|_{L^2(\Omega)} \leq \|v\|_{H^1(\Omega)}$, which means that we have (3.10) with $a_1 = \|f\|_{L^2(\Omega)}$. Furthermore, with $\mathbf{E} = (E_1, E_2)$, and using the Cauchy-Schwartz inequality again, we have

$$\begin{aligned}\left|\int_\Omega \nabla v\cdot(\lambda\mathbf{E})\,dx\right| &= \left|\int_\Omega \frac{\partial v}{\partial x_1}\lambda E_1\,dx + \int_\Omega \frac{\partial v}{\partial x_2}\lambda E_2\,dx\right| \\ &\leq \|\lambda\|_{L^\infty(\Omega)}\left[\|E_1\|_{L^2(\Omega)}\left\|\frac{\partial v}{\partial x_1}\right\|_{L^2(\Omega)} + \|E_2\|_{L^2(\Omega)}\left\|\frac{\partial v}{\partial x_2}\right\|_{L^2(\Omega)}\right].\end{aligned}$$

It is easy to prove the inequality $\left\|\frac{\partial v}{\partial x_i}\right\|_{L^2(\Omega)} \leq \|v\|_{H^1(\Omega)}$, $i = 1, 2$, and the inequality (3.11) follows with

$$a_2 = \|\lambda\|_{L^\infty(\Omega)}(\|E_1\|_{L^2(\Omega)} + \|E_2\|_{L^2(\Omega)}) = \|\lambda\|_{L^\infty(\Omega)}\|\mathbf{E}\|_{(L^2(\Omega))^2}.$$

At last, the final inequality (3.12) is obvious

$$\left|\int_{\partial\Omega_2} v g_v\,ds\right| \leq c\|v\|_{H^1(\Omega)}$$

for some constant c which is only dependent on g_v and $\partial\Omega_2$. So we have

$$|L(v)| \leq \Lambda\|v\|_{H^1(\Omega)},$$

where $\Lambda = a_1 + a_2 + a_3 = \|f\|_{L^2(\Omega)} + \|\lambda\|_{L^\infty(\Omega)}\|\mathbf{E}\|_{(L^2(\Omega))^2} + c$.

■

The uniqueness of the solution follows from the stability estimate. Suppose $p_1 \in H_g^1(\Omega)$ and $p_2 \in H_g^1(\Omega)$ are two solutions to the variational problem (V), then

$$\begin{aligned} a(p_1, q) &= L(q), \\ a(p_2, q) &= L(q), \end{aligned}$$

for all $q \in H_0^1(\Omega)$. By subtraction, we see that $p_1 - p_2 \in H_0^1(\Omega)$,

$$a(p_1 - p_2, q) = 0.$$

Applying the stability estimate to this situation with $L \equiv 0$, i.e., $\Lambda = 0$, we end with $\|p_1 - p_2\|_{H^1(\Omega)} = 0$, i.e., $p_1 = p_2$. Thus the uniqueness is verified.

Remarks: To this end, we can explain the choice of the volume fraction of liquid g_l in the aluminium DC-casting problem. Remember in the DC-casting problem, we have $\lambda = \frac{K}{A}$ where A is a constant and K is a scalar function of g_l that obeys the Kozeny-Carman relation

$$K(g_l) = \frac{g_l^3}{(1-g_l)^2}.$$

Here g_l is chosen to take the value ε_1 and $1-\varepsilon_2$ on Γ_1 and Γ_4 for two small positive values ε_1 and ε_2, and varies linearly in x_2 for a given x_1, cf Chapter 2. It is not difficult to see that with $g_l = 0$ on Γ_1, we have $K = 0$ on Γ_1, and consequently we loose the V-ellipticity. Meanwhile $g_l = 1$ on Γ_4 causes that K goes towards infinity on Γ_4 and we lose the continuity of $a(.,.)$. In order to avoid these difficulties and obtain a well-posed variational problem, we thus chose the g_l that varies from ε_1 to $1-\varepsilon_2$.

3.1.3 Discretization of the problem

For the discretization of the problem (V), we define the following finite-dimensional subspace and the linear variety as

$$V_h(\Omega) \subset H_0^1(\Omega), \qquad V_{h,g}(\Omega) \subset H_g^1(\Omega).$$

Hereby we define the discrete problem as

Problem (V_h): Find $p_h \in V_{h,g}(\Omega)$ such that

$$a(q_h, p_h) = L(q_h) \qquad \forall q_h \in V_h(\Omega). \tag{3.13}$$

The conditions in the Lax-Milgram Theorem will be satisfied in this discrete problem since $V_h(\Omega)$ is the finite-dimensional subspace of $H^1(\Omega)$, hence the discrete problem is also well-posed. Let $V_h(\Omega)$ be a finite-dimensional subspace of dimension M, and let φ_i, $i = 1, \cdots, M$, be the trial functions which constitute a basis for V_h, so that $\varphi_i \in V_h$ and any $q \in V_h$ has the unique representation

$$q = \sum_{i=1}^{M} \eta_i \varphi_i, \qquad \eta_i \in \mathbb{R}.$$

Then the discrete problem (3.13) is equivalent to

$$a(p_h, \varphi_j) = L(\varphi_j), \qquad j = 1, \cdots, M.$$

Note that $p = g_p$ on Γ_1 is the essential boundary condition which is explicitly imposed in the formulation. We can use the notation

$$p_h = \sum_{i=1}^{M} \xi_i \varphi_i, \qquad \xi_i \in \mathbb{R},$$

then we have

$$\sum_{i=1}^{M} a(\varphi_i, \varphi_j)\xi_i = L(\varphi_j), \qquad j = 1, \cdots, M.$$

The equivalent matrix form is $\mathbf{A}\boldsymbol{\xi} = \mathbf{b}$, where $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{M,M}$ with $a_{ij} = a(\varphi_i, \varphi_j)$, $\boldsymbol{\xi} = (\xi_i) \in \mathbb{R}^M$ and $\mathbf{b} = (b_i) \in \mathbb{R}^M$ with $b_i = L(\varphi_i)$. To prove that this linear system gives an unique solution, we notice

$$a(q, q) = a\left(\sum_{i=1}^{M} \eta_i\varphi_i, \sum_{j=1}^{M} \eta_j\varphi_j\right) = \sum_{i,j=1}^{M} \eta_i a(\varphi_i, \varphi_j)\eta_j = \boldsymbol{\eta}^T\mathbf{A}\boldsymbol{\eta},$$

and

$$\boldsymbol{\eta}^T\mathbf{A}\boldsymbol{\eta} = a(q, q) \geq \alpha\|q\|^2_{H^1(\Omega)} \geq 0,$$

with $\boldsymbol{\eta}^T\mathbf{A}\boldsymbol{\eta} = 0$ only when $q = 0$, i.e., only when $\boldsymbol{\eta} = 0$. In addition, it is obvious that $\mathbf{A}$ is symmetry since $a_{ij} = a_{ji}$. Therefore, we can conclude that the stiffness matrix $\mathbf{A}$ is symmetric and positive definite, and so the linear system has a unique solution.

Let $p \in H^1_g$ be the solution of the problem (V), and $p_h \in V_{h,g}$ be the solution of the problem (V_h). Following the standard discussion on error estimate of the finite element method for second-order elliptic equations, cf. Johnson [12], with the restriction on λ as $0 < \lambda_{\min} \leq \|\lambda\|_{L^\infty(\Omega)} < \infty$, we obtain the following standard error estimate for our problem

$$\begin{aligned} \|p - p_h\|_{L^2(\Omega)} &\leq Ch^2|p|_{H^2(\Omega)}, \\ |p - p_h|_{H^1(\Omega)} &\leq Ch|p|_{H^2(\Omega)}, \\ \|p - p_h\|_{H^1(\Omega)} &\leq Ch|p|_{H^2(\Omega)}, \end{aligned} \tag{3.14}$$

provided that p is "smooth" enough, i.e., $p \in H^2(\Omega)$.

3.1.4 The mapping theory

The trial functions in finite element methods are usually treated in context of the reference element $\widehat{\mathbf{K}}$. In order to compute the element matrix and the corresponding element righthand side vector for a given element, the necessary integrations are usually executed in the reference element and then mapped onto the actual element.

Let $F_{\mathbf{K}} : \widehat{x} \to F_{\mathbf{K}}(\widehat{x})$ be the unique invertible mapping such that

$$F_{\mathbf{K}}(\widehat{\mathbf{a}}_i) = \mathbf{a}_i,$$

where $\mathbf{a}_i$ denotes the coordinates of the ith vertex of $\mathbf{K}$. With any scalar function $\widehat{\phi}$ defined on $\widehat{\mathbf{K}}$ (respectively on $\partial\widehat{\mathbf{K}}$), we associate the function ϕ defined on $\mathbf{K}$ (respectively on $\partial\mathbf{K}$) by

$$\phi = \widehat{\phi} \circ F^{-1}_{\mathbf{K}}, \qquad \text{i.e., } \widehat{\phi} = \phi \circ F_{\mathbf{K}}.$$

When the underlying differential equation is a second-order elliptic equation, we have to compute the integrals

$$a^{\mathbf{K}}_{ij} = \int_{\mathbf{K}} \nabla\varphi_i \cdot (\lambda\nabla\varphi_j)\, dx, \ \ i, j = 1, \cdots, n, \tag{3.15}$$

where n is the number of shape functions in this element $\mathbf{K}$. By the chain rule

$$\frac{\partial \varphi_i}{\partial x_j} = \frac{\partial}{\partial x_j}\left(\widehat{\varphi}_i(F^{-1}(x))\right) = \frac{\partial \widehat{\varphi}_i}{\partial \widehat{x}_1}\frac{\partial \widehat{x}_1}{\partial x_j} + \frac{\partial \widehat{\varphi}_i}{\partial \widehat{x}_2}\frac{\partial \widehat{x}_2}{\partial x_j},$$

we have

$$\nabla \varphi_i = J^{-T}\nabla \widehat{\varphi}_i,$$

where J is the *Jacobian matrix* of the mapping F, and J^{-T} is the transposed Jacobian matrix of the mapping F^{-1}.

Now we can transform the integral in (3.15) to an integral over $\widehat{\mathbf{K}}$ using the mapping $F : \widehat{\mathbf{K}} \to \mathbf{K}$ as

$$a_{ij}^{\mathbf{K}} = \int_{\widehat{\mathbf{K}}} \left(J^{-T}\nabla \widehat{\varphi}_i\right) \cdot \left(\lambda J^{-T}\nabla \widehat{\varphi}_j\right) |\det J|\, d\widehat{x},$$

where $\det J$ is the determinant of J.

For more details of the conforming FEM, we refer to Johnson [12] and Zienkiewicz and Morgen [28].

3.2 The Implementation

In this section, we give the implementation of the aluminium DC-casting problem. Section 3.2.1 gives a short introduction to the finite element programming in `DIFFPACK` , and section 3.2.2 gives an outline of the class `Pressure` for our problem. The C++ code is not included in this thesis.

3.2.1 The finite element programming in `DIFFPACK`

Traditionally, the implementation of a finite element code is a big task. Such programs must deal with several topics including discretization, numerical integration, assembling of the stiffness matrix and the solution of linear systems. Basic finite element programming is already implemented in `DIFFPACK`, based on the object-oriented programming language C++. (For details of the C++ language, we refer to Stroustrup [26].) We obtain very short and clear source codes when we apply the finite element part of `DIFFPACK`. Usually, the problem-dependent part of the program will need only a few pages of codes, while the remaining parts are handled by the classes in the `DIFFPACK` libraries. Here is a short introduction to the finite element part of `DIFFPACK`. For further details we refer to the corresponding reports on `DIFFPACK`, especially [6], [7], [8] and [9].

- **Element types**

  The `ElmDef` hierarchy is designed to represent various types of reference elements that may be used by finite element methods. This hierarchy has an abstract base class `ElmDef` which contains data structures and functions for describing a finite element in its local coordinate system. The class is equipped with several member functions that take care of the basic operations regarding element geometry, trial functions and the mapping between a local and physical coordinates. Particular elements are

implemented as classes derived from `ElmDef`, i.e., new element types can be easily added to the hierarchy.

⁎ **Class GridFE**

The finite element method needs geometric information in terms of a grid. In addition to the coordinates of the nodal points, we may need the global node numbers for the nodes in each element, a "material" number associated with each element, the `ElmDef` type of each element and indicators for certain boundary conditions. Class `GridFE` has been designed to contain and handle this information. Usually, `GridFE` objects are declared by an empty constructor and filled by a finite element preprocessor.

⁎ **Class Preprocessors**

A preprocessor is used to generate a mesh grid and fill the `GridFE` object. `DIFFPACK` has a class hierarchy for diverse preprocessor methods, with class `Prepro` as the base class.

The available preprocessors are based on different strategies, such as the creation of single boxes, super elements and quadtrees. The organization of each preprocessor class follows common guidelines for separating the description of the geometry from the actual grid parameters. To exemplify, the super element processor in class `PreproSupElSet` consists essentially of two references, one to a `geometrySupElSet` object and the other to a `partitionSupElSet` object. The class `geometrySupElSet` contains the geometry information of all the super elements, including boundary indicators, and the class `partitionSupElSet` contains the partition of each super element into finite elements.

The technique used by `PreproSupElSet` is a conventional one: First, the domain is divided into a few large elements, so-called super elements. Afterwards, the preprocessor discretize each super element before combining these grid patches to a global finite element grid. The division of the geometry into a set of super elements is specified manually by the user, and is followed by parameters describing the partition of each super element. This information can be stored in files (for ex., `*.geom` and `*.part` files) with special formats. When these files are read by the preprocessor function "`scan`" in class `geometrySupElSet` and class `partitionSupElSet`, the grid can be generated automatically. Consult [9] for more details.

⁎ **Class FieldFE, class FieldsFE, class FieldSubDomain etc.**

The solution at the nodes can be stored as vector entries. The class `FieldFE`, which combines a `GridFE` with a vector, can be used to describe any scalar field. The class has several features that are convenient for a programmer working with finite element methods, e.g.,

- ◎ `FieldFE::operator=(FuncField&)` creates a finite element field that has values at all the nodal points defined by an explicit function.
- ◎ function `valuePt` returns the value at arbitrary points inside the field (by interpolating the finite element field).
- ◎ function `derivativePt` computes the gradient of the field at an arbitrary point.

Other classes, such as class `FieldsFE` for vector field and class `FieldSubDomain` for piecewise constant field etc., are also equipped with the corresponding functionalities.

* **Class FiniteElement**

  This class contains all the information of a finite element that is useful for a programmer. It generates a global finite element consisting of the definition of the element in its local coordinate system, the isoparametric mapping, the relation between local and global node numbers and the nodal coordinates of the element. Class `FiniteElement` also allows shape functions and their derivatives to be evaluated at a given point in a given element. Moreover, this class serves as an interface to numerical integration over a given element.

* **Numerical integration**

  For most non-trivial problems, it is difficult or even impossible to find the analytical expressions of the integrals involved in element matrices and vectors. This situation is even more complicated for distorted geometries, see for instance the variational formulation in Section 3.1. To overcome these difficulties, it is necessary to use numerical integration schemes that reduces the integrals to sums.

  The class `ElmItg` is developed for this purpose. It performs numerical integration over an element, or over a side. Frequently used types of numerical integration scheme is known as Gauss or Gauss-Legendre quadratures.

* **Linear system solver**

  Any finite element method ends up with solving a system of algebraic equations, either linear or non-linear. Class `LinEqAdm` is designed to be an interface to linear systems and the corresponding solvers. The main data structure is composed of a `LinEqAdm` class consisting of the coefficient matrix and two vectors, one for the load vector(right side of the equation), and one for the solution vector.

  Many kinds of matrix storage schemes and solution methods are allowed in `LinEqAdm`. The commonly used types for the coefficient matrix are `MatBand` and `MatSparse`. The solution methods are divided into two groups: direct methods such as Gaussian Elimination, and iterative methods such as Orthomin, conjugate gradients and many others.

* **Class FEM**

  When the finite element method is used to solve differential equations, there are always several operations that are independent of the given problem. Class `FEM` is developed to deal with these common tasks. Functions like `makeSystem` and `integrate` are quite general, except that they call some problem-dependent functions that should be defined by the user. However, the class is equipped with default versions of many of the functions that may be sufficient for simple problems.

  When implementing a specific application, the user defines his own problem-dependent class which is derived from `FEM`. This application class has to implement certain functions that are nonexistent in class `FEM`, and also redefine default functions that are unsuitable for the given problems.

3.2.2 Outline of class `Pressure`

The class `Pressure` is derived from class `FEM` to solve the pressure equation in the aluminium DC-casting process model. In the problem dependent functions like `fillEssBC` or `integrands` one has access to all the data in the problem.

The corresponding code is not included in this thesis, we only give an outline of the class `Pressure`. The solution of velocity is calculated by taking gradient of the pressure solution obtained by finite element method. Notice that for piecewise linear pressure field, its gradient is only continuous on each element, but discontinuous on the whole solution domain. To avoid such inconvenience, the velocity is calculated at the centroid of each element. The outline of the main functions of the class `Pressure` is listed here:

- Function `makegrid`, which uses the preprocessor to generate the mesh grid;
- Function `init_Gl`, which generates the initial g_l;
- Function `fillEssBC`, which defines the essential boundary conditions and which is inherited as a pure virtual function in class `FEM`;
- Function `makeSystem`, which creates a linear system of equations for the problem;
- Function `integrands`, which defines the integrand in the volume integral in the variational formulation;
- Function `cal_V`, which calculates the superficial relative velocity;
- Function `plot`, which outputs data for g_l (liquid volume fraction), p (pressure) and $\mathbf{v}$ (superficial relative velocity) into files, which can be visualized by `plotmtv`.
- Function `driver`, which contains the main algorithm for solving the finite element problem.

3.3 Numerical Experiments

In this section, we continue our study of the conforming FEM with some numerical experiments. In Section 3.3.1, we solve a pressure equation in which the numerical method gives the same solution as the analytical one. In Section 3.3.2, the pressure equation defined on an irregular geometry is solved and compared with its analytical solution so that the rates of convergence can be estimated. In Section 3.3.3, we solve a non-trivial problem defined on a highly irregular geometry — the aluminium DC-casting problem. Results in graphs are included at the end of the section.

3.3.1 Solving the pressure equation on a square domain

We want to solve the pressure equation

$$-\nabla \cdot (\lambda \nabla p) = 1 \qquad \text{in } \Omega,$$

where $\Omega = [0,1] \times [0,1]$. When the boundaries are shown in Figure 3.1, the boundary conditions are

$$\begin{aligned} p &= 1 && \text{on } \Gamma_1, \\ \frac{\partial p}{\partial \mathbf{n}} &= 0 && \text{on } \Gamma_2, \\ p &= 0 && \text{on } \Gamma_3, \\ \frac{\partial p}{\partial \mathbf{n}} &= 0 && \text{on } \Gamma_4. \end{aligned}$$

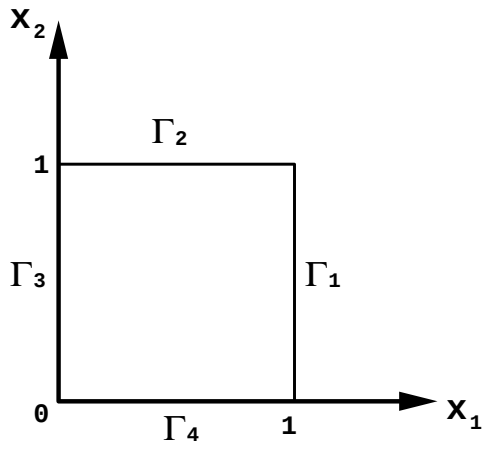


Figure 3.1: The square solution domain and the boundary indicators .

Observe that with $\mathbf{v} = -\lambda \nabla p$ one can easily prove that $\dfrac{\partial p}{\partial \mathbf{n}} = 0$ is equivalent to $\mathbf{v} \cdot \mathbf{n} = 0$. With the function $\lambda(x_1, x_2) = x_1 + 1$, we obtain the analytical solution for the problem simply as

$$p(x_1, x_2) = x_1.$$

The corresponding weak formulation is then given by (3.4) with $\lambda(x_1, x_2) = x_1 + 1$, $\mathbf{E} = (0,0)^T$, $f \equiv 1$, $\partial\Omega_1 = \Gamma_1 \cup \Gamma_3$ and $\partial\Omega_2 = \Gamma_2 \cup \Gamma_4$. When piecewise bilinear trial functions are used for p, the numerical method reproduces the analytical solution exactly: the maximum error for p solution is only 3.33067×10^{-16} when we divide the unit square with 4×4 partition.

3.3.2 Estimating the rate of convergence for the Pressure equation on an irregular geometry

Consider the Pressure equation

$$-\nabla \cdot (\lambda \nabla p) = f \qquad \text{in } \Omega,$$

where Ω is chosen be to the same as our aluminium DC-casting model problem, cf Figure 2.2. The boundary conditions are then given

$$\begin{aligned} p &= \tilde{p} && \text{on } \partial\Omega_A = \Gamma_1 \cup \Gamma_4, \\ \mathbf{v} \cdot \mathbf{n} = 0 \text{ or } \frac{\partial p}{\partial \mathbf{n}} &= 0 && \text{on } \partial\Omega_B = \Gamma_2 \cup \Gamma_3 \cup \Gamma_5. \end{aligned}$$

where $\mathbf{v} = -\lambda\nabla p$, and $\tilde{p}$ is a given function.

Clearly, this is a special case of the problem discussed in Section 3.1 with $\mathbf{E} = (0,0)^T$ and $g_v \equiv 0$. Here we study two cases:

- **Case 1:**

  With $\lambda \equiv 1$, $f \equiv 0$ and $\tilde{p}(x_1, x_2) = \cos(\omega x_1)\cosh(\omega x_2)$, the pressure equation reduces to the Laplace equation. And the analytical solution for p is

  $$p(x_1, x_2) = \cos(\omega x_1)\cosh(\omega x_2),$$

  where $\omega = \pi H_s/L$ and L and H_s are parameters for the problem given in Table 2.3.

  ■

- **Case 2:**

  With

  $$\begin{aligned} \lambda(x_1, x_2) &= \cosh(4x_1) + \cosh(2x_2) \\ \tilde{p}(x_1, x_2) &= \cos(\omega x_1)\cos(\omega x_2), \\ f(x_1, x_2) &= -\nabla \cdot (\lambda \nabla \tilde{p}), \end{aligned}$$

  we have the analytical solution of the problem

  $$p(x_1, x_2) = \cos(\omega x_1)\cos(\omega x_2),$$

  where ω is the same as in Case 1.

  ■

The rate of convergence of the method is related to different norms. We define the discrete norms for the scalar-valued function p in 2D as

$$\begin{aligned} \|p\|_{l^\infty(\Omega)} &= \max_{i=1}^{N} |p_i|, \\ \|p\|_{l^2(\Omega)} &= \left[\frac{1}{N}\sum_{i=1}^{N}(p_i)^2\right]^{\frac{1}{2}}. \end{aligned}$$

Here N is the number of nodes in the grid, and p_i denote the values evaluated at the node i. For the continuous norms, we study the usual L^2-norm and H^1-norm. Note that the discrete l^2-norm is actually an approximation to the continuous L^2-norm. The rate of convergence is calculated as follows:

We assume that the error e measured in some norm $\|.\|$ satisfies $\|e\| = ch^a$ for some constants c and a. Then for two different partitions, we have

$$e_1 = ch_1^a \quad \text{and} \quad e_2 = ch_2^a.$$

By dividing e_1 by e_2, we have

$$\frac{e_1}{e_2} = \left(\frac{h_1}{h_2}\right)^a.$$

Thus the rate of convergence is decided by

$$a = \frac{\ln(\frac{e_1}{e_2})}{\ln(\frac{h_1}{h_2})}.$$

Let p_h denote the numerical approximation of p. We use piecewise bilinear trial functions for p_h. The errors and the estimated rates of convergence of the numerical solutions measured in these norms are given in Tables 3.1 and 3.2. The results agree with the error estimate, cf (3.14).

The errors:				
partition(s)	$\|p-p_h\|_{l^2(\Omega)}$	$\|p-p_h\|_{l^\infty(\Omega)}$	$\|p-p_h\|_{L^2(\Omega)}$	$\|p-p_h\|_{H^1(\Omega)}$
1x1	9.37148e-01	3.61941e+00	4.54688e-01	2.08374e+01
2x2	2.03212e-01	9.23581e-01	1.16089e-01	1.01605e+01
4x4	4.74288e-02	2.09647e-01	2.93780e-02	5.06045e+00
8x8	1.17849e-02	5.29116e-02	7.37343e-03	2.52827e+00
16x16	3.14260e-03	1.41850e-02	1.83859e-03	1.26393e+00
32x32	1.01921e-03	4.55509e-03	4.53797e-04	6.31943e-01
The rates of convergence:				
1x1 $\sim$ 2x2	2.2059446	1.9704441	1.9696457	1.0362039
2x2 $\sim$ 4x4	2.0991503	2.1392763	1.9824232	1.0056338
4x4 $\sim$ 8x8	2.0088239	1.9863062	1.9943284	1.0011151
8x8 $\sim$ 16x16	1.9069090	1.8992179	2.0037361	1.0002340
16x16 $\sim$ 32x32	1.6245073	1.6388147	2.0184808	1.0000502

Table 3.1: The errors and the rates of convergence of p_h for case 1.

The errors:				
partition(s)	$\|p-p_h\|_{l^2(\Omega)}$	$\|p-p_h\|_{l^\infty(\Omega)}$	$\|p-p_h\|_{L^2(\Omega)}$	$\|p-p_h\|_{H^1(\Omega)}$
1x1	2.28311e-03	1.00476e-02	3.10035e-02	6.41714e-01
2x2	1.45513e-03	8.00804e-03	9.79429e-03	3.57426e-01
4x4	3.88188e-04	2.01233e-03	2.63298e-03	1.84675e-01
8x8	9.77409e-05	4.85885e-04	6.71033e-04	9.30728e-02
16x16	2.44625e-05	1.22885e-04	1.68585e-04	4.66289e-02
The rates of convergence:				
1x1 $\sim$ 2x2	0.6498523	0.3273298	1.6624183	0.84428588
2x2 $\sim$ 4x4	1.9063206	1.9925823	1.8952444	0.95365601
4x4 $\sim$ 8x8	1.9897212	2.0501801	1.9721971	0.98855706
8x8 $\sim$ 16x16	1.9983905	1.9833061	1.9929075	0.99713521

Table 3.2: The errors and the rates of convergence of p_h for case 2.

3.3.3 The aluminium DC-casting problem - the pressure equation posed on an irregular geometry

In this section, we present the numerical solutions for the aluminium DC-casting problem graphically. The outline of the implementation of class `Pressure` for the DC-casting model

is described in subsection 3.2.2. We use class `PreproSupElSet` in the preprocessor hierarchy to generate the `GridFE` object with the solution domain being divided into 9 super elements. We only present the solutions using quadrilateral elements since triangular elements give the same results. The linear system is solved by both Gauss Elimination and the conjugate gradient method. The two solvers give almost the same results.

Figure 3.2 shows the finite element grid for our problem with 5×5 partition on each super element. Note that we have finer grid in the critical region. According to the discussion in Johnson [12, p. 90, eq. (4.16)], we have the following error estimate

$$|p - p_h|_{H^1(\Omega)} \leq C \left[\sum_K \left(h_{\mathbf{K}} |p|_{H^2(\mathbf{K})} \right)^2 \right]^{1/2}.$$

It is clear that we would like to balance the size of $h_{\mathbf{K}}$ with that of $|p|_{H^2(\mathbf{K})}$, and in particular we should choose small $h_{\mathbf{K}}$ where $|p|_{H^2(\mathbf{K})}$ is large, i.e., the critical region in our problem.

The computing results for the variables involved in the problem equations are shown in Figures 3.3-3.7. These results are obtained from the grid which takes a 5×5 partition on each of the 9 super elements, and the trial functions for pressure are piecewise bilinear. Figures of the critical region for pressure and velocity solutions are also given. In order to present the velocity solution more clearly, we include a plot of streamlines[1] in the critical region, cf. Figure 3.8.

Here we give some general remarks about these results:

- The volume fraction of the liquid aluminium g_l is the input argument of our problem, which is given to be 0.99 on Γ_4 and 0.01 on Γ_1, as shown in Figure 3.3.
- The solution for the pressure p shows that in the lower part of the domain, p varies slowly and smoothly, with almost horizontal contour lines. But in the critical region, p has rapid variation. Especially at the part which is very close to Γ_2 where the aluminium is supposed to run out, we find almost vertical contour lines.
- The relation between the velocity $\mathbf{v}$ and the pressure p is given by

  $$\mathbf{v} = -\frac{K}{A}(\nabla p + \mathbf{E}).$$

  As we see from the solution for the pressure, p varies vertically and smoothly in the lower part of the domain. Since the effect of ∇p is totally cancelled out by the constant vector $\mathbf{E}$ in this area, the corresponding solution $\mathbf{v}$ is zero. In contrast, p varies quickly in the critical region, which means the corresponding velocity is no longer zero. A closer look at the part which is very close to the boundary Γ_2 reveals that the direction of the velocity is almost horizontal, which is indicated by the horizontal-varying pressure solution. The plot of the streamlines in the critical region shows the same result more clearly.

[1] The plot of streamlines is produced with the help of the corresponding `DIFFPACK` software implemented by Xing Cai.

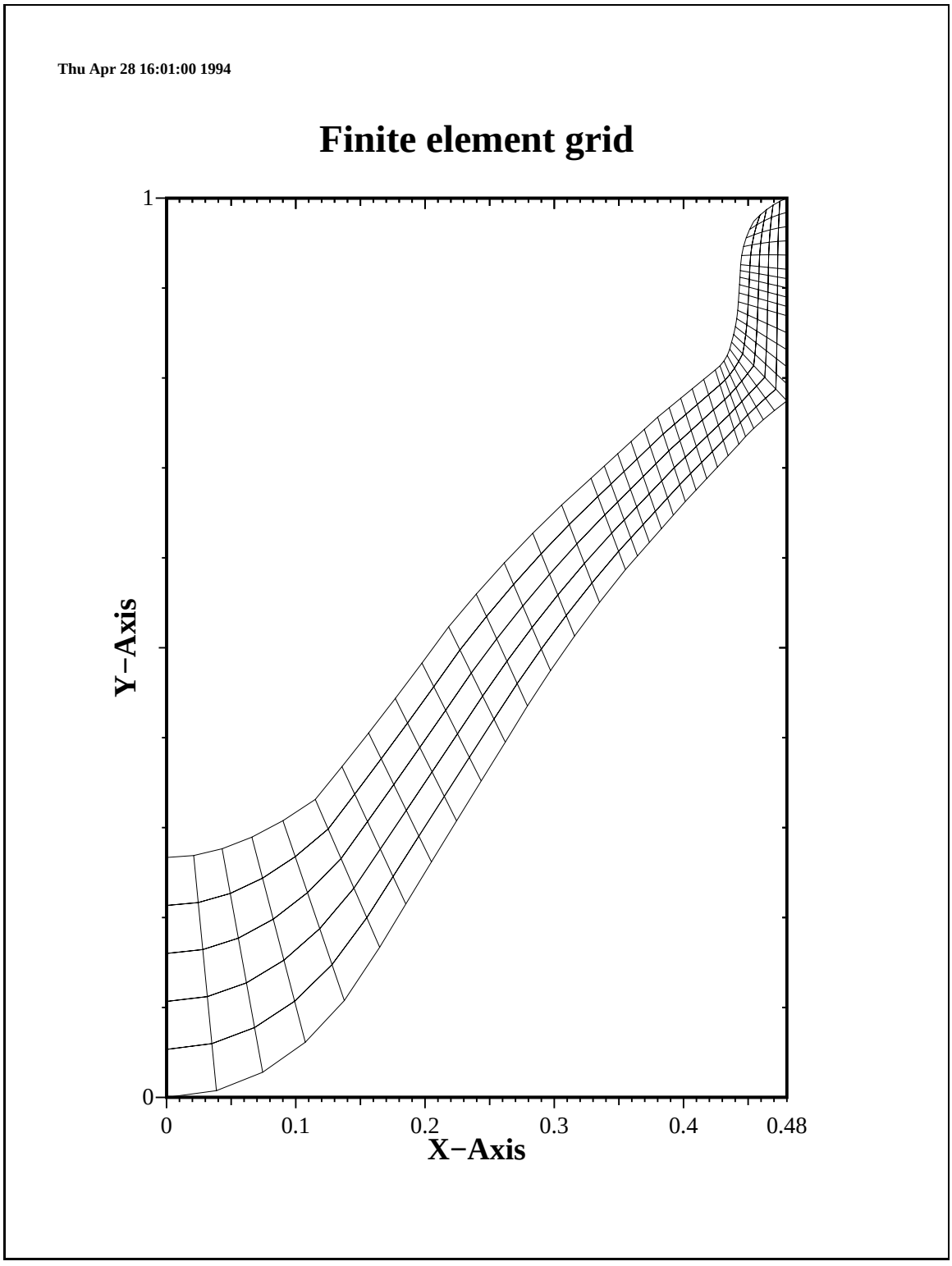


Figure 3.2: Finite element grid of rectangular elements.

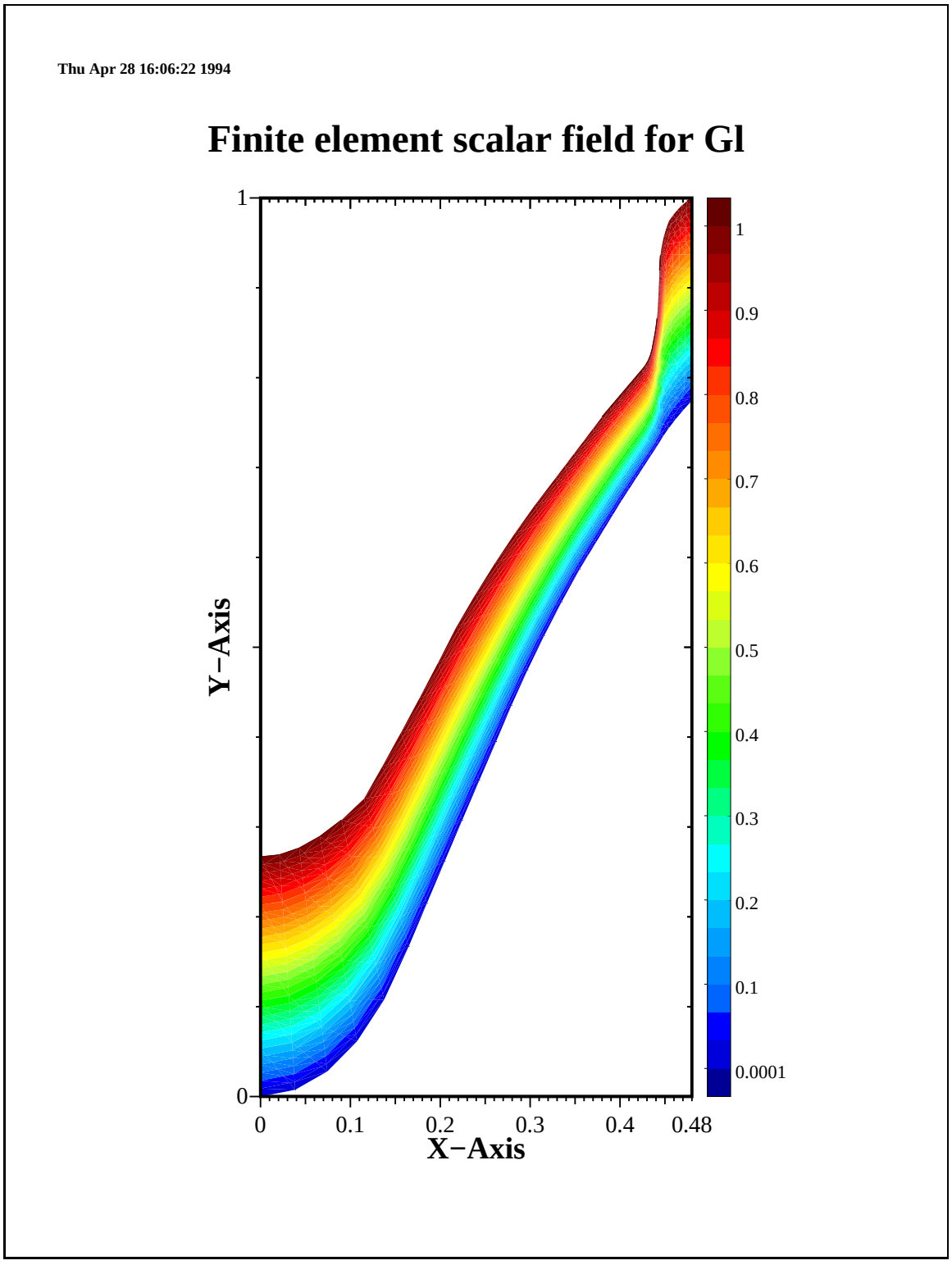


Figure 3.3: Plot of the volume fraction liquid aluminium G_l.

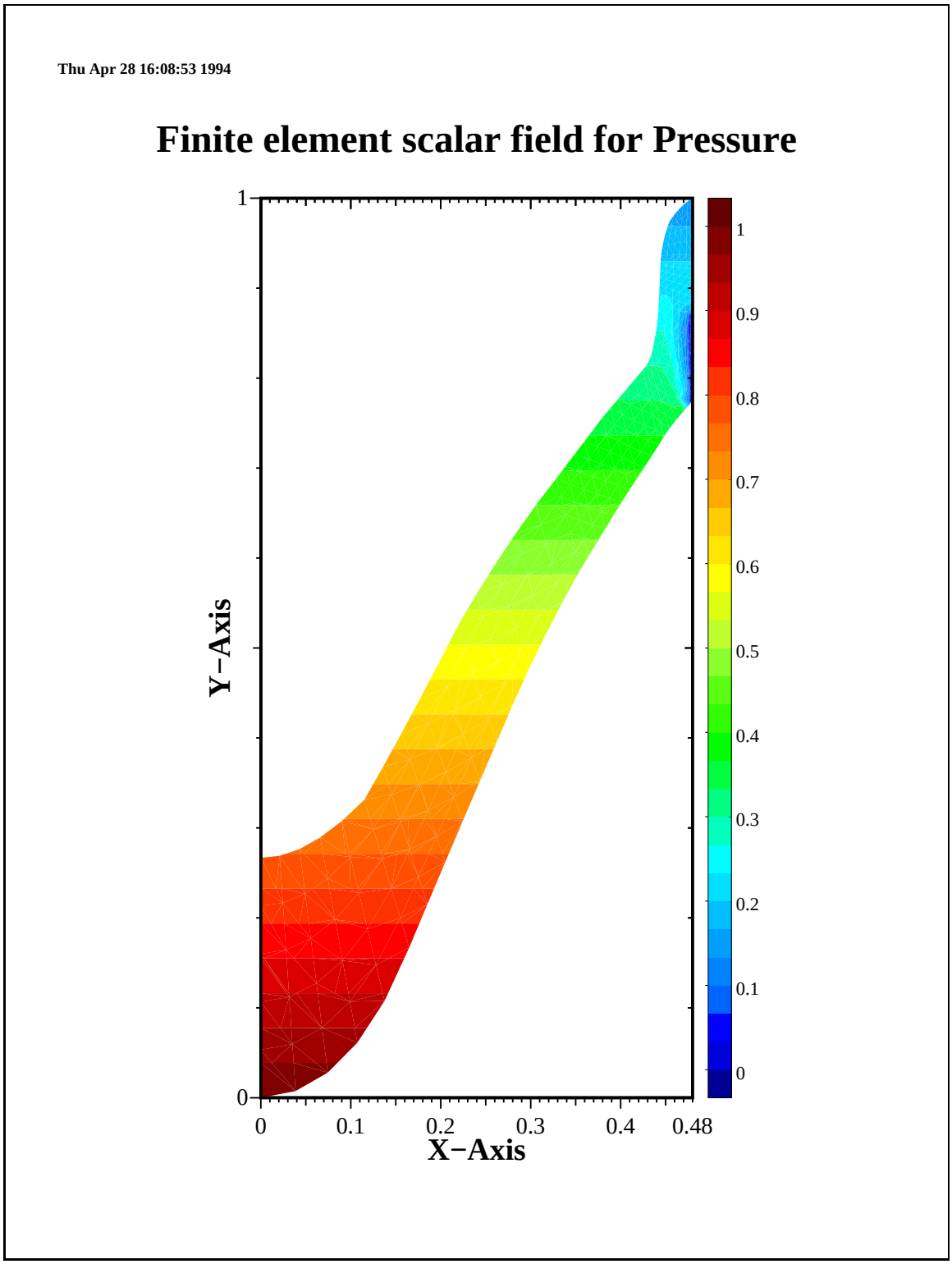


Figure 3.4: Plot of the conforming FEM solution of the pressure P.

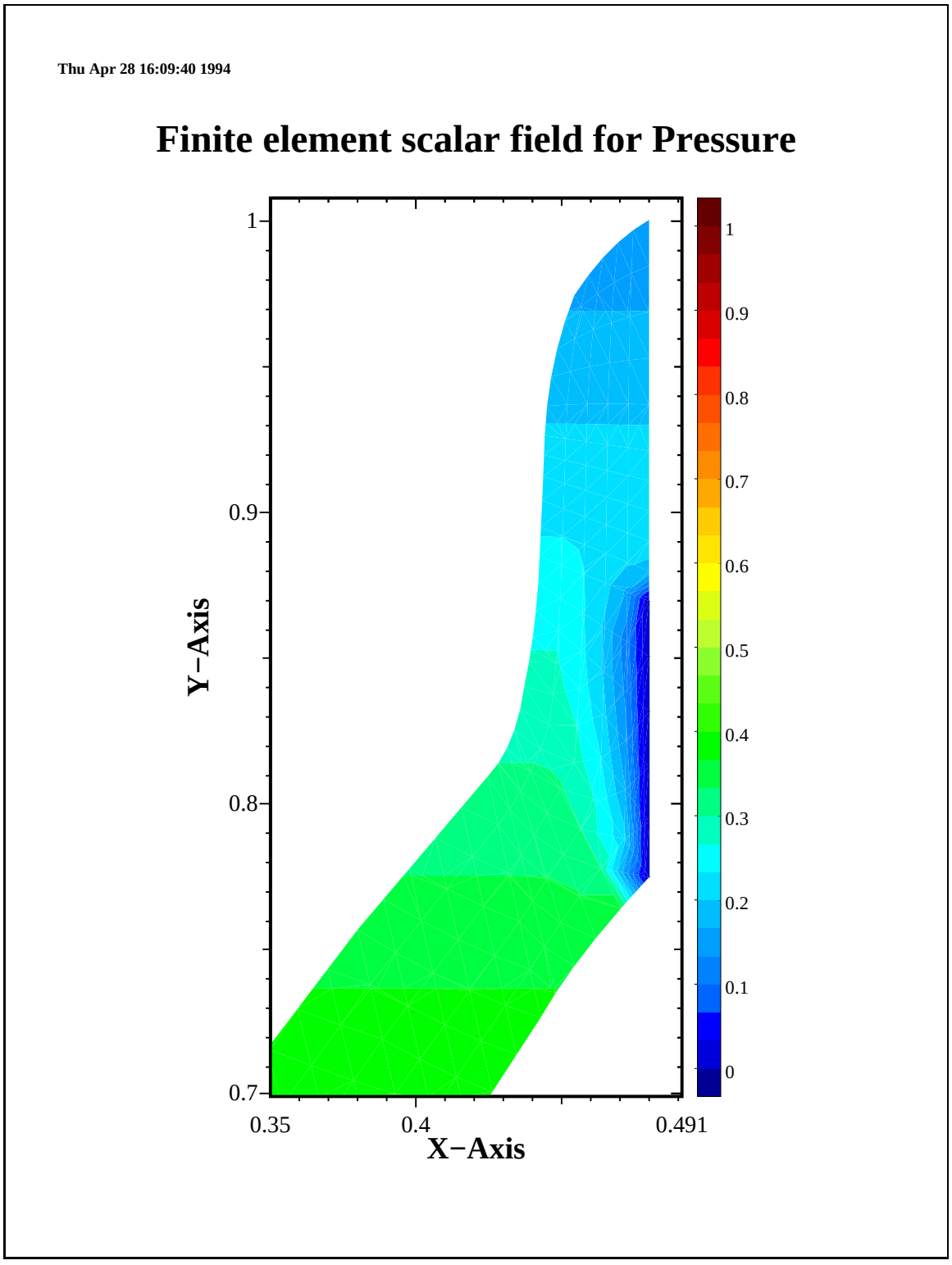


Figure 3.5: Plot of the conforming FEM solution of P in the critical region.

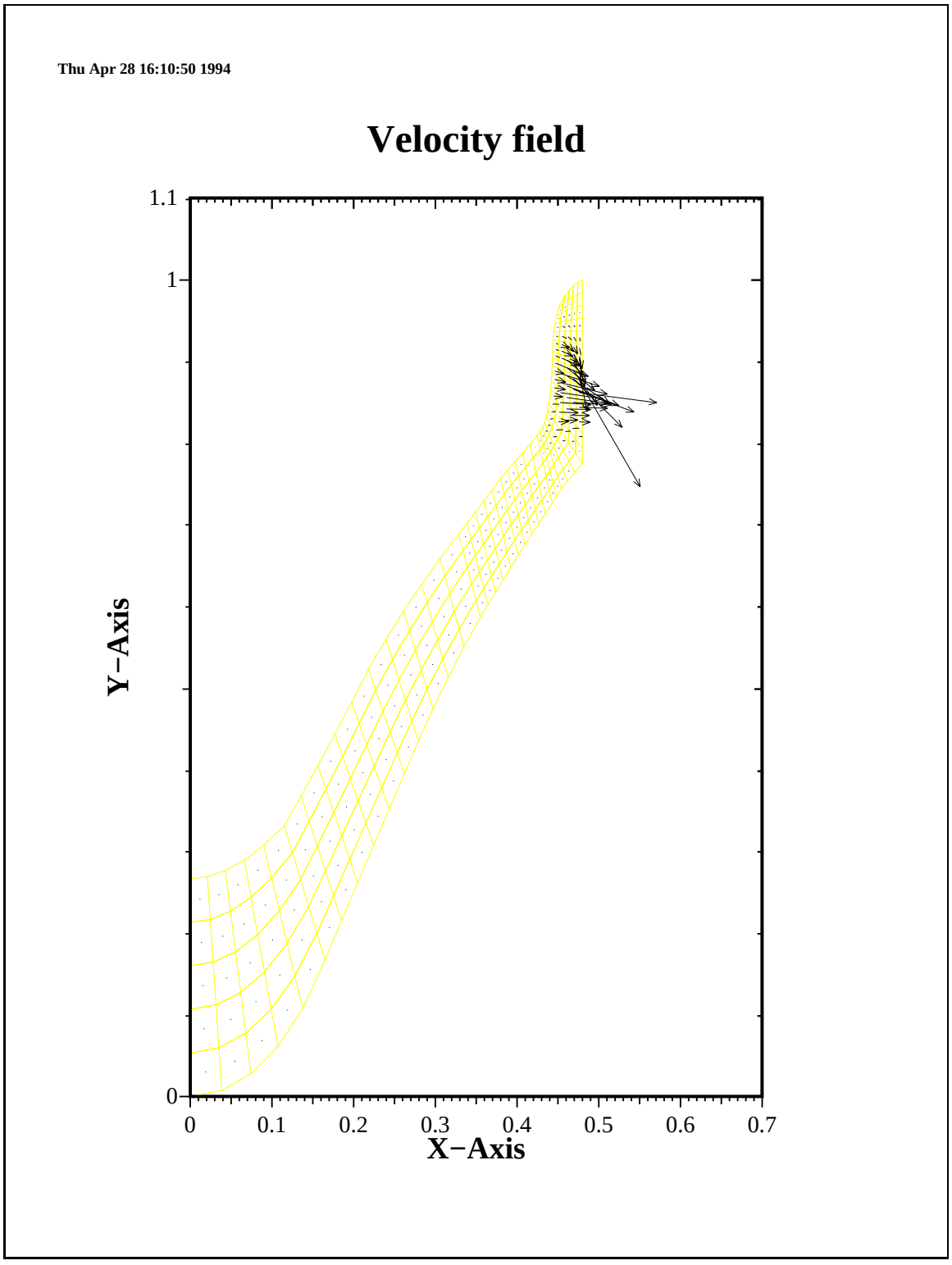


Figure 3.6: Plot of the conforming FEM solution of $\mathbf{v}$.

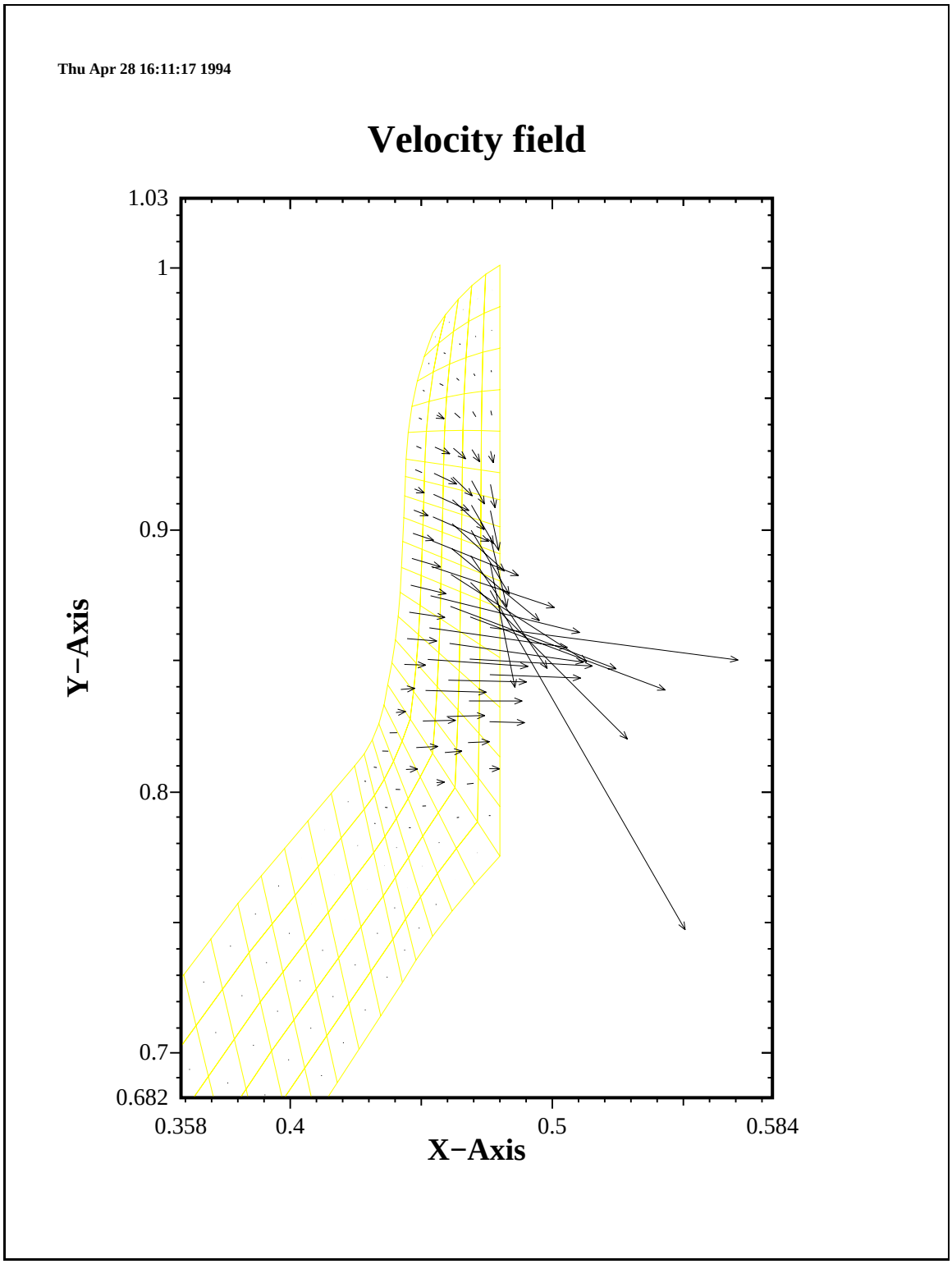


Figure 3.7: Plot of the conforming FEM solution of $\mathbf{v}$ in the critical region.

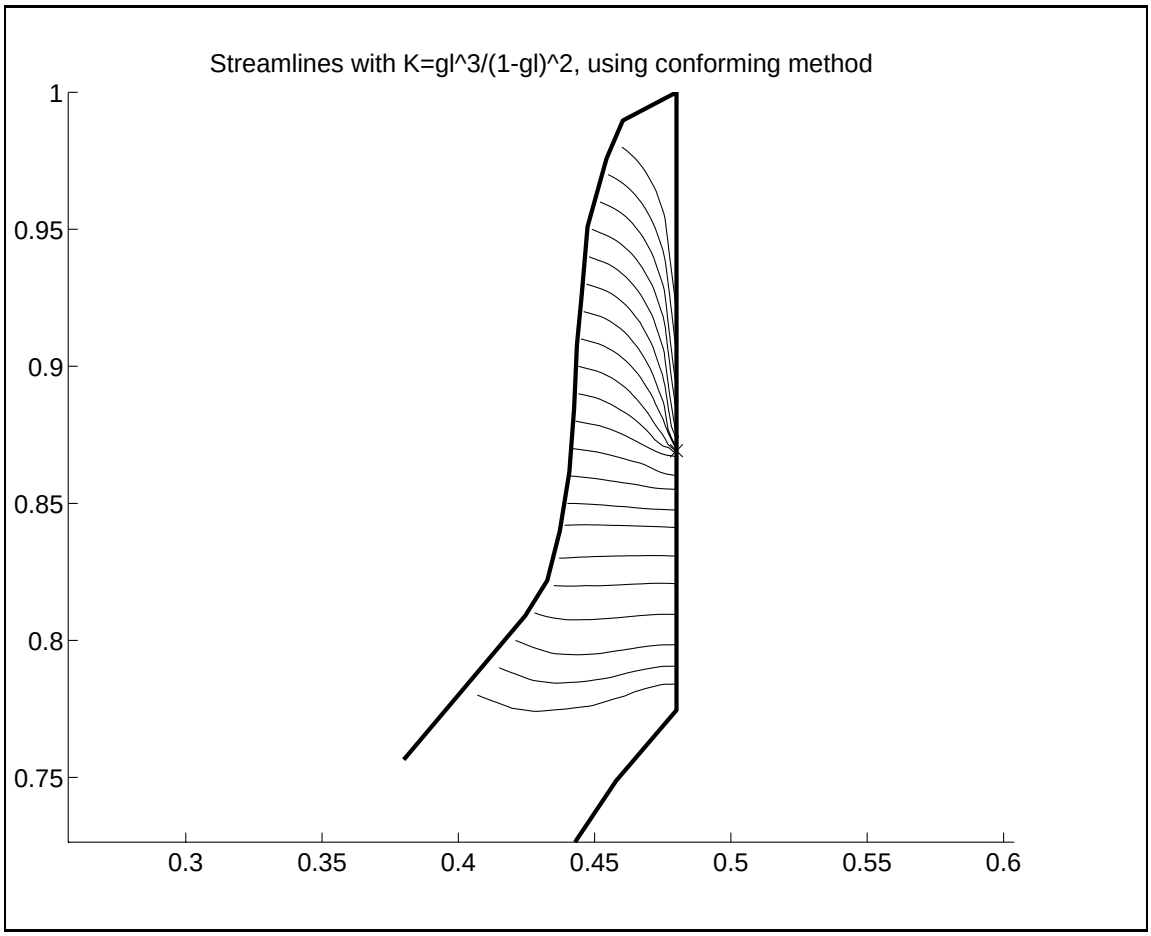


Figure 3.8: Plot of streamlines in the critical region, conforming FEM.

Chapter 4

Mixed Finite Element Method

4.1 Introduction

In this Chapter we will discuss the theoretical and practical aspects of the mixed finite element method applied to second order elliptic equations, such as the Poisson equation and the pressure equation. The mixed FEM is described in section 4.2, and its present implementation in `DIFFPACK` will be explained in section 4.3. In section 4.4, we give some examples together with the results of several numerical experiments.

Consider the same elliptic model problem (1.1) presented in Chapter 1

$$-\nabla\cdot(\lambda(\nabla p+\mathbf{E}))=f \qquad \text{in } \Omega\in\mathbb{R}^2, \tag{4.1}$$

subject to the boundary conditions

$$p = g_p \qquad \text{on } \partial\Omega_1, \tag{4.2}$$
$$\mathbf{v}\cdot\mathbf{n} = g_v \qquad \text{on } \partial\Omega_2, \tag{4.3}$$

where $\mathbf{v}=-\lambda\left(\nabla p+\mathbf{E}\right)$. Typically, the problem (4.1)-(4.3) can be solved by a *conforming* finite element method, as we have discussed in Chapter 3.

The variable $\mathbf{v}$ derived above may also have a physical meaning, and it is vital in many applications. In the context of the pressure equation, it represents the corresponding velocity field. Using the conforming FEM, one can obtain the velocity by calculating the gradient of the pressure field produced by the finite element solution. Inevitably, the rate of convergence for the velocity solution will be of one lower order than that of the pressure solution. In order to get the same order of accuracy in both solutions, one can view both the pressure and the velocity as the primary unknowns, and thus solve the two entities simultaneously. This method is called the *mixed* finite element method.

Mixed FEM is widely used for the solution of second-order elliptic equations, especially for problems where $\mathbf{v}$ has a physical interpretation. To this end, we rewrite the pressure equation (4.1) into two coupled first order partial differential equations as

$$\mathbf{v} = -\lambda\left(\nabla p+\mathbf{E}\right), \tag{4.4}$$
$$\nabla\cdot\mathbf{v} = f \qquad \text{in } \Omega. \tag{4.5}$$

These equations are still subject to the boundary conditions (4.2)-(4.3). However, as will be seen in Section 4.2, the condition (4.2) is now a *natural* boundary, while (4.3) becomes an *essential* boundary condition, which is the opposite case of the conforming FEM.

4.2 The Mixed Finite Element Formulation of the Pressure Equation

In this section, we present the mixed finite element method applied to the model problem. First, we give the weak formulation of the elliptic problem. After that, we discuss the use of mixed triangular elements and mixed quadrilateral elements. The special mapping theory used in the mixed FEM is also briefly explained. Finally, we end the section by describing

the choice of the trial functions and the setting up of the corresponding linear system of algebraic equations. For more detailed discussion of the mixed FEM, we refer to the recent survey by Roberts and Thomas [20] and the references therein.

4.2.1 The weak formulation

First, we need to define some basic notations. Given an integer $m \geq 0$, the usual Sobolev spaces $H^m(\Omega)$ are defined in Section 3.1.1. In order to derive the appropriate variational form of the current problem, we also introduce the space

$$H(\mathrm{div};\Omega) = \{\mathbf{v} \in (L^2(\Omega))^n; \nabla \cdot \mathbf{v} \in L^2(\Omega)\}$$

together with its norm

$$\| \mathbf{v} \|_{H(\mathrm{div};\Omega)} = \left(\| \mathbf{v} \|^2_{0,\Omega} + \| \nabla \cdot \mathbf{v} \|^2_{0,\Omega}\right)^{\frac{1}{2}}.$$

Defining $\mathbf{u} = (u_1, u_2)$ and q as functions in appropriate subspaces of $H(\mathrm{div};\Omega)$ and $L^2(\Omega)$, the weak formulation of (4.4) and (4.5) is derived as the following:

First we take the vector inner product with $\mathbf{u}$ on both sides of equation (4.4) and multiply equation (4.5) with q, thereafter we integrate both equations over Ω, which leads to

$$\begin{aligned} \int_\Omega \mathbf{u} \cdot \frac{1}{\lambda}\mathbf{v}\, dx + \int_\Omega \mathbf{u} \cdot \nabla p\, dx &= -\int_\Omega \mathbf{u} \cdot \mathbf{E}\, dx, \\ -\int_\Omega q \nabla \cdot \mathbf{v}\, dx &= \int_\Omega f q\, dx. \end{aligned}$$

The Green's formula

$$\int_\Omega (\nabla \cdot \mathbf{u}) q\, dx + \int_\Omega \mathbf{u} \cdot \nabla q\, dx = \int_\Gamma q(\mathbf{u} \cdot \mathbf{n})\, ds,$$

which is valid for all $q \in H^1(\Omega)$, can be applied to the term $\int_\Omega \mathbf{u} \cdot \nabla p\, dx$, i.e.,

$$\int_\Omega \mathbf{u} \cdot \nabla p\, dx = \int_{\partial\Omega} p(\mathbf{u} \cdot \mathbf{n})\, ds - \int_\Omega p(\nabla \cdot \mathbf{u})\, dx.$$

To enforce the essential boundary conditions $\mathbf{v} \cdot \mathbf{n} = 0$ on $\partial\Omega_2$, it is convenient to define the subspace

$$H_0(\mathrm{div};\Omega) = \{\mathbf{u} \in H(\mathrm{div};\Omega);\ \mathbf{u} \cdot \mathbf{n} = 0 \ \text{ on } \partial\Omega_2\}.$$

Now we choose $\mathbf{u} \in H_0(\mathrm{div};\Omega)$. Consequently, the boundary integral is reduced to

$$\int_\Omega \nabla p \cdot \mathbf{u}\, dx = \int_{\partial\Omega_1} g_p(\mathbf{u} \cdot \mathbf{n})\, ds - \int_\Omega p(\nabla \cdot \mathbf{u})\, dx.$$

Moreover, let the variety be defined as

$$\tilde{H}(\mathrm{div};\Omega) = \{\mathbf{u} \in H(\mathrm{div};\Omega);\ \mathbf{u} \cdot \mathbf{n} = g_v \ \text{ on } \partial\Omega_2\},$$

the weak formulation of the model problem (4.4)-(4.5) can then be stated as

Find a pair of functions $(p, \mathbf{v}) \in L^2(\Omega) \times \tilde{H}(\mathrm{div};\Omega)$ such that:

$$\begin{array}{lll} a(\mathbf{u},\mathbf{v}) + b(p,\mathbf{u}) & = \alpha(\mathbf{u}) & \forall \mathbf{u} \in H_0(\mathrm{div};\Omega), \\ b(q,\mathbf{v}) & = \beta(q) & \forall q \in L^2(\Omega). \end{array} \tag{4.6}$$

Here the bilinear forms $a : H(\mathrm{div};\Omega) \times H(\mathrm{div};\Omega) \mapsto \mathbb{R}$ and $b : L^2(\Omega) \times H_0(\mathrm{div};\Omega) \mapsto \mathbb{R}$ are given as

$$\begin{aligned} a(\mathbf{u},\mathbf{v}) &= \int_\Omega \mathbf{u} \cdot \left(\frac{1}{\lambda}\mathbf{v}\right) dx, \\ b(p,\mathbf{u}) &= -\int_\Omega p \nabla \cdot \mathbf{u}\, dx, \end{aligned}$$

while the linear functionals $\alpha : H(\mathrm{div};\Omega) \mapsto \mathbb{R}$ and $\beta : L^2(\Omega) \mapsto \mathbb{R}$ are defined by

$$\begin{aligned} \alpha(\mathbf{u}) &= -\int_\Omega \mathbf{u} \cdot \mathbf{E}\, dx - \int_{\partial\Omega_1} g_p (\mathbf{u} \cdot \mathbf{n})\, ds, \\ \beta(q) &= -\int_\Omega f q\, dx. \end{aligned}$$

Discretization of the problem based on the mixed variational formulation (4.6) can be stated in terms of the finite-dimensional subspaces Q_h and U_h, and the variety $\tilde{U}_h$, such that

$$Q_h \subset L^2(\Omega), \qquad U_h \subset H_0(\mathrm{div};\Omega) \qquad \text{and} \qquad \tilde{U}_h \subset \tilde{H}(\mathrm{div};\Omega).$$

The corresponding finite-dimensional problem is then

Find a pair of functions $(p_h, \mathbf{v}_h) \in Q_h \times \tilde{U}_h$ such that

$$\begin{array}{lll} a(\mathbf{u}_h,\mathbf{v}_h) + b(p_h,\mathbf{u}_h) & = \alpha(\mathbf{u}_h) & \forall \mathbf{u}_h \in U_h, \\ b(q_h,\mathbf{v}_h) & = \beta(q_h) & \forall q_h \in Q_h. \end{array} \tag{4.7}$$

It remains to construct the finite-dimensional subspaces Q_h and U_h such that they satisfy "good" approximation properties and certain compatibility conditions. More precisely, given $\mathbf{u}_h \in U_h$, we want the relation

$$\int_\Omega q_h \nabla \cdot \mathbf{u}_h\, dx = 0 \quad \forall q_h \in Q_h \quad \Rightarrow \quad \nabla \cdot \mathbf{u}_h = 0 \tag{4.8}$$

In addition, Q_h and U_h must satisfy the inf-sup condition, i.e., there must exist a constant c independent of the mesh size h such that

$$\inf_{q_h \in Q_h} \sup_{u_h \in U_h} \frac{b(q_h,\mathbf{u}_h)}{\|q_h\|_{0,\Omega} \|\mathbf{u}_h\|_{H(\mathrm{div};\Omega)}} \geq c > 0. \tag{4.9}$$

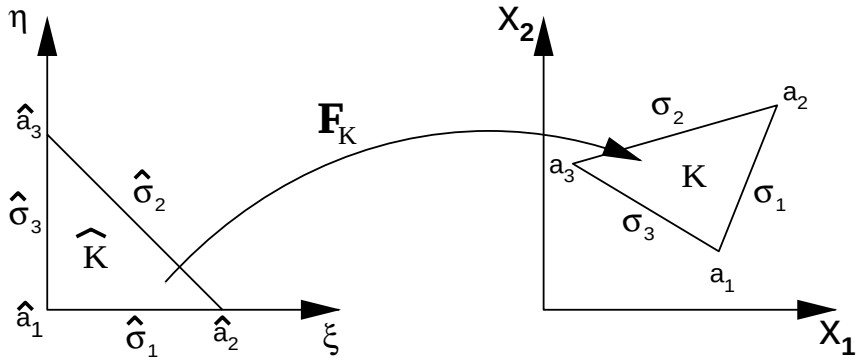


Figure 4.1: The mapping F_K from $\widehat{\mathbf{K}}$ to $\mathbf{K}$.

This means that for a given "pressure space" Q_h, the "velocity space" U_h has to be sufficiently large. In addition, Q_h and U_h must satisfy the condition $\sup_{q_h \in Q_h} b(q_h, \mathbf{u}_h) > 0$ for all $\mathbf{u}_h \in U_h$ such that $\nabla \cdot \mathbf{u}_h \neq 0$. For more details, we refer to Raviart and Thomas [19] and Brezzi [1].

The use of both triangular elements and quadrilateral elements which leads to different choices of Q_h and U_h, will be discussed next.

4.2.2 Mixed triangular elements

Following the discussion by Raviart and Thomas [19], we look closer at a particular choice of subspaces Q_h and U_h based on triangular elements. Let $\mathbf{K}$ be a triangle, we define a space $U_{\mathbf{K}}$ of vector-valued functions $\mathbf{u} \in H(\mathrm{div}; \mathbf{K})$ on $\mathbf{K}$ such that:

(i) $\nabla \cdot \mathbf{u}$ is a polynomial of degree $\leq k$,

(ii) the restriction of $\mathbf{u} \cdot \mathbf{n}_{\mathbf{K}}$ to any side σ_i, $i = 1, 2, 3$, of $\mathbf{K}$ is a polynomial of degree $\leq k$, where k is a non-negative integer.

Since any triangle can be mapped onto the unit right triangle $\widehat{\mathbf{K}}$ in the (ξ, η)-plane whose vertices are $\hat{\mathbf{a}}_1 = (1, 0)$, $\hat{\mathbf{a}}_2 = (0, 1)$ and $\hat{\mathbf{a}}_3 = (0, 0)$ (cf. Figure 4.1), we need only to study the space $\widehat{U}$ associated with $\widehat{\mathbf{K}}$. This space $\widehat{U}$ is required to satisfy the following properties:

$$\begin{aligned} (P_k)^2 &\subset \widehat{U}, \\ \dim(\widehat{U}) &= (k+1)(k+3), \\ \nabla \cdot \mathbf{u} = \frac{\partial u_1}{\partial \xi} + \frac{\partial u_2}{\partial \eta} &\in P_k \qquad \text{for all } \mathbf{u} \in \widehat{U}, \\ \mathbf{u} \cdot \mathbf{n} &\in \widehat{S}_k \qquad \text{for all } \mathbf{u} \in \widehat{U}, \\ \text{and} \quad \widehat{U}_0 = \Big\{ \mathbf{u} \in \widehat{U}; \nabla \cdot \mathbf{u} = 0 \Big\} &\in (P_k)^2. \end{aligned} \tag{4.10}$$

Here P_k denotes the space of all polynomials of degree $\leq k$ in the two variables ξ and η, and $\widehat{S}_k$ denotes the space of all functions defined over the boundary $\partial\widehat{\mathbf{K}}$ whose restrictions

to any side $\hat{\sigma}_i$ of $\widehat{\mathbf{K}}$ are polynomials of degree $\leq k$. These properties will also hold if the unit right triangle is mapped back to the general triangle $\mathbf{K}$. We will only use these results, and refer to Raviart and Thomas [19] for a proof.

Assume that $\mathcal{K}_h$ is a triangulation of $\overline{\Omega} = \Omega \cup \partial\Omega$ made up of triangles $\mathbf{K}$ whose diameters are $\leq h$, then we introduce the space

$$U_h = \{\mathbf{u}_h \in H(\mathrm{div};\Omega);\ \mathbf{u}_h\ |_{\mathbf{K}} \in U_{\mathbf{K}} \text{ for all } \mathbf{K} \in \mathcal{K}_h, \},$$

where $U_{\mathbf{K}}$ is defined as

$$U_{\mathbf{K}} = \left\{\mathbf{u} \in H(\mathrm{div};\mathbf{K});\ \mathbf{u} \in \widehat{U}\right\}.$$

The degrees of freedom for a function $\mathbf{u}_h \in U_h$ are easily determined: for any $\mathbf{u}_h \in U_h$ and $\mathbf{K} \in \mathcal{K}_h$, we have $(\nabla \cdot \mathbf{u}_h)|_{\mathbf{K}} \in P_k$. Therefore, a natural choice for the space Q_h is

$$Q_h = \{q_h \in L^2(\Omega);\ q_h|_{\mathbf{K}} \in P_k \text{ for all } \mathbf{K} \in \mathcal{K}_h, \},$$

such that condition (4.8) will be satisfied automatically. Note that the function $q_h \in Q_h$ does not need to satisfy any continuity constraint at the inter-element boundaries.

One particular way to choose the space $\widehat{U}$ is to use a polynomial basis. For example, let $k \geq 0$ be an even integer, and define $\widehat{U}$ to be the space of all functions $\mathbf{u} = (u_1, u_2)$ of the form

$$\begin{cases} u_1 = pol_k(\xi,\eta) + \alpha_0\xi^{k+1} + \alpha_1\xi^k\eta + \cdots + \alpha_{\frac{k}{2}}\xi^{\frac{k}{2}+1}\eta^{\frac{k}{2}} \\ u_2 = pol_k(\xi,\eta) + \beta_0\eta^{k+1} + \beta_1\eta^k\xi + \cdots + \beta_{\frac{k}{2}}\eta^{\frac{k}{2}+1}\xi^{\frac{k}{2}} \end{cases},$$

with

$$\sum_{i=0}^{\frac{k}{2}}(-1)^i(\alpha_i - \beta_i) = 0.$$

Here $pol_k(\xi,\eta)$ denotes any polynomial of degree k in the two variables ξ and η. It is not difficult to prove that this choice of $\widehat{U}$ satisfies the required properties in (4.10).

In the special case when $k = 0$, the individual components of a function $\mathbf{u} \in \widehat{U}$ are of the form

$$\begin{cases} u_1 = a_0 + a_1\xi \\ u_2 = b_0 + b_1\eta \end{cases},$$

with $a_1 = b_1$. And a function $q \in \widehat{Q}$ can simply be any constant.

For odd integers k, the case is different. If $k \geq 1$ is an odd integer, we can define $\widehat{U}$ to be the space of all functions $\mathbf{u} = (u_1, u_2)$ of the form

$$\begin{cases} u_1 = pol_k(\xi,\eta) + \alpha_0\xi^{k+1} + \alpha_1\xi^k\eta + \cdots + \alpha_{\frac{k+1}{2}}\xi^{\frac{k+1}{2}}\eta^{\frac{k+1}{2}} \\ u_2 = pol_k(\xi,\eta) + \beta_0\eta^{k+1} + \beta_1\eta^k\xi + \cdots + \beta_{\frac{k+1}{2}}\eta^{\frac{k+1}{2}}\xi^{\frac{k+1}{2}} \end{cases},$$

satisfying the constraint

$$\sum_{i=0}^{\frac{k+1}{2}}(-1)^i\alpha_i = \sum_{i=0}^{\frac{k+1}{2}}(-1)^i\beta_i = 0.$$

Here one can easily check that conditions in (4.10) hold.

Taking $k = 1$ as an example, the components of a function $\mathbf{u} \in \widehat{U}$ look like

$$\begin{cases} u_1 = a_0 + a_1\xi + a_2\eta + a_3\xi(\xi + \eta) \\ u_2 = b_0 + b_1\xi + b_2\eta + b_3\eta(\xi + \eta) \end{cases}.$$

With this choice of $\mathbf{u}$, the inf-sup condition (4.9) will be satisfied if $q \in \widehat{Q}$ is a polynomial of degree 1. For higher k, it is not difficult to express a function $\mathbf{u} \in \widehat{U}$ or $q \in \widehat{Q}$. However, such values of k lead to complicated expressions and will not be discussed further in this thesis.

4.2.3 Mixed quadrilateral elements

Using the results in the previous section on triangular elements, it is now easy to get the results for quadrilateral elements. We begin with introducing the space $\widehat{U}$ associated with the unit square $\widehat{\mathbf{K}} = [-1, 1]^2$ in the (ξ, η)-plane. Any parallelogram in the (x_1, x_2)-plane can be mapped onto the unit square, and this mapping is unique and invertible. Given two integers $k, l \geq 0$, let us denote by $P_{k,l}$ the space of all polynomials in the two variables ξ and η of the form

$$P(\xi, \eta) = \sum_{i=0}^{k} \sum_{j=0}^{l} c_{ij} \xi^i \eta^j, \qquad c_{ij} \in \mathbb{R}.$$

Now we define the space $\widehat{U}$ as

$$\widehat{U} = \{\mathbf{u} = (u_1, u_2);\, u_1 \in P_{k+1,k}, u_2 \in P_{k,k+1}\}. \tag{4.11}$$

Note that for $\mathbf{u} \in \widehat{U}$, we have

(i) $\nabla \cdot \mathbf{u} = \dfrac{\partial u_1}{\partial \xi} + \dfrac{\partial u_2}{\partial \eta} \in P_{k,k}$,

(ii) the restriction of $\mathbf{u} \cdot \mathbf{n}$ to any side $\widehat{\sigma}_i$, $i = 1, 2, 3, 4$, of $\widehat{\mathbf{K}}$ is a polynomial of degree $\leq k$.

It is also easy to check that conditions in (4.10) hold. Corresponding to (4.11), the space $\widehat{Q}$ can be defined as

$$\widehat{Q} = \{q \in L^2(\Omega);\, q \in P_{k,k}\}.$$

For the case $k = 0$, the individual components of a function $\mathbf{u} \in \widehat{U}$ look like

$$\begin{cases} u_1 = a_0 + a_1\xi \\ u_2 = b_0 + b_1\eta \end{cases}.$$

With this choice of $\mathbf{u}$, a function $q \in \widehat{Q}$ can just be constant. It is then easy to verify that the inf-sup condition (4.9) will be satisfied. We will not discuss higher-order elements in this thesis.

4.2.4 The error estimate

The analysis of Brezzi [1], Raviart and Thomas [19] and Falk and Osborn [10] yield the following error estimate for the mixed FEM:

$$\begin{aligned} \|p - p_h\|_{L^2(\Omega)} &\leq C_1 h^{k+1}, \\ \|\mathbf{v} - \mathbf{v}_h\|_{(L^2(\Omega))^2} &\leq C_2 h^{k+1}, \\ \|\mathbf{v} - \mathbf{v}_h\|_{H(\mathrm{div};\Omega)} &\leq C_3 h^{k+1}, \end{aligned} \tag{4.12}$$

where k is the order of polynomials defined in subsections 4.2.2 and 4.2.3, and C_1, C_2 and C_3 are constants depending upon the smoothness of p and $\mathbf{v}$.

In particular, when $k = 0$, we have

$$\begin{aligned} \|p - p_h\|_{L^2(\Omega)} &\leq C_1 h, \\ \|\mathbf{v} - \mathbf{v}_h\|_{(L^2(\Omega))^2} &\leq C_2 h, \\ \|\mathbf{v} - \mathbf{v}_h\|_{H(\mathrm{div};\Omega)} &\leq C_3 h, \end{aligned} \tag{4.13}$$

where C_1, C_2 and C_3 are also constants depending upon the smoothness of p and $\mathbf{v}$.

4.2.5 Mapping for mixed finite elements

Same as in the conforming FEM, the trial functions in the mixed FEM are also treated in the context of the reference element $\widehat{\mathbf{K}}$. In order to compute the element matrix and the corresponding element right-hand side vector for a given element, the necessary integrations are usually mapped to the reference element. However, the mapping theory for the mixed FEM is different from the one used for the conforming FEM.

Mapping of mixed triangular elements. Let $F_{\mathbf{K}} : \widehat{x} \rightarrow F_{\mathbf{K}}(\widehat{x})$ be the unique invertible mapping such that

$$F_{\mathbf{K}}(\widehat{\mathbf{a}}_i) = \mathbf{a}_i, \qquad i = 1, 2, 3,$$

where $\mathbf{a}_i$ denotes the coordinates of the ith vertex of $\mathbf{K}$ (cf. Figure 4.1). With any scalar function $\widehat{\phi}$ defined on $\widehat{\mathbf{K}}$ (respectively on $\partial\widehat{\mathbf{K}}$), we associate the function ϕ defined on $\mathbf{K}$ (respectively on $\partial\mathbf{K}$) by

$$\begin{aligned} \phi &= \widehat{\phi} \circ F_{\mathbf{K}}^{-1}, \\ \text{i.e.,} \qquad \widehat{\phi} &= \phi \circ F_{\mathbf{K}}. \end{aligned}$$

On the other hand, with any vector-valued function $\widehat{\boldsymbol{\psi}} = (\widehat{\psi}_1, \widehat{\psi}_2)$ defined on $\widehat{\mathbf{K}}$, we associate the function $\boldsymbol{\psi}$ defined on $\mathbf{K}$ by

$$\boldsymbol{\psi} = \frac{1}{|J_{\mathbf{K}}|} J_{\mathbf{K}} \widehat{\psi} \circ F_{\mathbf{K}}^{-1}, \tag{4.14}$$

$$\text{i.e.,} \qquad \widehat{\boldsymbol{\psi}} = |J_{\mathbf{K}}| J_{\mathbf{K}}^{-1} \psi \circ F_{\mathbf{K}}. \tag{4.15}$$

Here $J_{\mathbf{K}}$ is the Jacobian matrix of the mapping, and $|J_{\mathbf{K}}| = \det(J_{\mathbf{K}})$. The relation between $\widehat{\phi}$ and ϕ, as well as that between $\widehat{\boldsymbol{\psi}}$ and $\boldsymbol{\psi}$, is one-to-one.

The transformations (4.14) and (4.15) are based on the following standard result:

For any function $\widehat{\boldsymbol{\psi}} \in (H^1(\widehat{\mathbf{K}}))^2$, we have:

$$\begin{aligned} \int_{\widehat{\mathbf{K}}} \widehat{\phi}(\nabla \cdot \widehat{\boldsymbol{\psi}})\, d\widehat{x} &= \int_{\mathbf{K}} \phi(\nabla \cdot \boldsymbol{\psi})\, dx \qquad \forall \widehat{\phi} \in L^2(\widehat{\mathbf{K}}), \\ \int_{\partial\widehat{\mathbf{K}}} \widehat{\phi}(\widehat{\boldsymbol{\psi}} \cdot \widehat{\mathbf{n}})\, d\widehat{s} &= \int_{\partial\mathbf{K}} \phi(\boldsymbol{\psi} \cdot \mathbf{n}_{\mathbf{K}})\, ds \qquad \forall \widehat{\phi} \in L^2(\partial\widehat{\mathbf{K}}). \end{aligned}$$

For more details, we refer to Raviart and Thomas [19].

Mapping of mixed quadrilateral elements. The special mapping theory for mixed triangular elements is applied here in the same way. Let $F_{\mathbf{K}} : \widehat{\mathbf{K}} \to \mathbf{K}$ denote the affine invertible mapping $F_{\mathbf{K}}(\widehat{x}) = B_{\mathbf{K}}\widehat{x} + b_{\mathbf{K}}$. Assume that $\mathcal{K}_h$ is a triangulation[1] of $\overline{\Omega} = \Omega \cup \partial\Omega$ consisting of parallelograms $\mathbf{K}$ whose diameters are $\leq h$, we define

$$U_h = \left\{ \mathbf{u}_h \in H(\mathrm{div};\Omega);\ \mathbf{u}_h|_{\mathbf{K}} \in \widehat{U} \text{ for all } \mathbf{K} \in \mathcal{K}_h \right\}.$$

Since for any $\mathbf{u}_h \in U_h$ and any $\mathbf{K} \in \mathcal{K}_h$ we have

$$(\nabla \cdot \mathbf{u}_h)|_{\mathbf{K}} \circ F_{\mathbf{K}} \in P_{k,k},$$

then we may use

$$Q_h = \{ q_h \in L^2(\Omega); q_h|_{\mathbf{K}} \circ F_{\mathbf{K}} \in P_{k,k} \text{ for all } \mathbf{K} \in \mathcal{K}_h \}.$$

4.2.6 Deriving the linear system of equations

Given suitable basis functions for the involved subspaces, the variational problem (4.7) may be reformulated in terms of a system of linear algebraic equations. Let $\{\phi_i\}_{i=1}^{n_p}$ and $\{\boldsymbol{\psi}_i\}_{i=1}^{n_v} = \{(\psi_{1,i}, \psi_{2,i})\}_{i=1}^{n_v}$ denote the bases for the subspaces Q_h (pressure) and U_h (velocity), respectively. We may thus express the functions $p_h \in Q_h$ and $\mathbf{v}_h \in U_h$ as linear combinations of these basis functions, i.e.,

$$p_h = \sum_{i=1}^{n_p} p_i \phi_i,$$

and

$$\mathbf{v}_h = \sum_{i=1}^{n_v} v_i \boldsymbol{\psi}_i = \left(\sum_{i=1}^{n_v} v_i \psi_{1,i}, \sum_{i=1}^{n_v} v_i \psi_{2,i} \right),$$

[1] Triangulation is the conventional name of the method for dividing the solution domain into a mesh grid of finite elements, with either triangular or quadrilateral.

for $n_p = \dim(Q_h)$ and $n_v = \dim(U_h)$. Using the traditional Galerkin approach, i.e., choosing $q_h = \phi_j$ and $\mathbf{u}_h = \boldsymbol{\psi}_j$, we rewrite problem (4.7) as

$$\begin{aligned} & \int_\Omega \boldsymbol{\psi}_j \cdot \frac{1}{\lambda} \left(\sum_{i=1}^{n_v} v_i \boldsymbol{\psi}_i \right) dx - \int_\Omega \left(\sum_{i=1}^{n_p} p_i \phi_i \right) \nabla \cdot \boldsymbol{\psi}_j \, dx \\ = \quad & - \int_\Omega \boldsymbol{\psi}_j \cdot \mathbf{E} \, dx - \int_{\partial\Omega_2} g_p (\boldsymbol{\psi}_j \cdot \mathbf{n}) \, ds, \qquad j = 1, \cdots, n_v, \end{aligned}$$

and

$$- \int_\Omega \phi_j \nabla \cdot \left(\sum_{i=1}^{n_v} v_i \boldsymbol{\psi}_i \right) dx = - \int_\Omega f \phi_j \, dx, \qquad j = 1, \cdots, n_p.$$

Clearly, we have a linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$ with $n_v + n_p$ unknowns. The nonzero contributions to $\mathbf{A}\mathbf{x} = \mathbf{b}$ from each element depend on the element type and the polynomial order k. In this paper, we concentrate on mixed quadrilateral elements with $k = 0$. When using block-centered finite elements, cf. Figure 4.3, there are 5 nodal variables associated with a general element Ω^e, namely 4 velocity nodes and 1 pressure node. If we locally denote the trial functions associated with the nodes v_1, v_2, v_3, v_4 and p_5 by $\boldsymbol{\psi}_1$, $\boldsymbol{\psi}_2$, $\boldsymbol{\psi}_3$, $\boldsymbol{\psi}_4$ and ϕ_5, then the nonzero contributions from this element e to the system are contained in $\mathbf{A}^e \mathbf{x}^e = \mathbf{b}^e$, where

$$\mathbf{A}^e = \int_{\Omega^e} \begin{bmatrix} \frac{1}{\lambda}\boldsymbol{\psi}_1 \cdot \boldsymbol{\psi}_1 & \frac{1}{\lambda}\boldsymbol{\psi}_1 \cdot \boldsymbol{\psi}_2 & \frac{1}{\lambda}\boldsymbol{\psi}_1 \cdot \boldsymbol{\psi}_3 & \frac{1}{\lambda}\boldsymbol{\psi}_1 \cdot \boldsymbol{\psi}_4 & -\phi_5 \nabla \cdot \boldsymbol{\psi}_1 \\ \frac{1}{\lambda}\boldsymbol{\psi}_2 \cdot \boldsymbol{\psi}_1 & \frac{1}{\lambda}\boldsymbol{\psi}_2 \cdot \boldsymbol{\psi}_2 & \frac{1}{\lambda}\boldsymbol{\psi}_2 \cdot \boldsymbol{\psi}_3 & \frac{1}{\lambda}\boldsymbol{\psi}_2 \cdot \boldsymbol{\psi}_4 & -\phi_5 \nabla \cdot \boldsymbol{\psi}_2 \\ \frac{1}{\lambda}\boldsymbol{\psi}_3 \cdot \boldsymbol{\psi}_1 & \frac{1}{\lambda}\boldsymbol{\psi}_3 \cdot \boldsymbol{\psi}_2 & \frac{1}{\lambda}\boldsymbol{\psi}_3 \cdot \boldsymbol{\psi}_3 & \frac{1}{\lambda}\boldsymbol{\psi}_3 \cdot \boldsymbol{\psi}_4 & -\phi_5 \nabla \cdot \boldsymbol{\psi}_3 \\ \frac{1}{\lambda}\boldsymbol{\psi}_4 \cdot \boldsymbol{\psi}_1 & \frac{1}{\lambda}\boldsymbol{\psi}_4 \cdot \boldsymbol{\psi}_2 & \frac{1}{\lambda}\boldsymbol{\psi}_4 \cdot \boldsymbol{\psi}_3 & \frac{1}{\lambda}\boldsymbol{\psi}_4 \cdot \boldsymbol{\psi}_4 & -\phi_5 \nabla \cdot \boldsymbol{\psi}_4 \\ -\phi_5 \nabla \cdot \boldsymbol{\psi}_1 & -\phi_5 \nabla \cdot \boldsymbol{\psi}_2 & -\phi_5 \nabla \cdot \boldsymbol{\psi}_3 & -\phi_5 \nabla \cdot \boldsymbol{\psi}_4 & 0 \end{bmatrix} dx, \qquad (4.16)$$

with

$$\mathbf{x}^e = (v_1, v_2, v_3, v_4, p_5)^T,$$

and

$$\mathbf{b}^e = \begin{bmatrix} - \int_{\Omega^e} \boldsymbol{\psi}_1 \cdot \mathbf{E} \, dx - \int_{\partial\Omega_2^e} g_p (\boldsymbol{\psi}_1 \cdot \mathbf{n}) \, ds \\ - \int_{\Omega^e} \boldsymbol{\psi}_2 \cdot \mathbf{E} \, dx - \int_{\partial\Omega_2^e} g_p (\boldsymbol{\psi}_2 \cdot \mathbf{n}) \, ds \\ - \int_{\Omega^e} \boldsymbol{\psi}_3 \cdot \mathbf{E} \, dx - \int_{\partial\Omega_2^e} g_p (\boldsymbol{\psi}_3 \cdot \mathbf{n}) \, ds \\ - \int_{\Omega^e} \boldsymbol{\psi}_4 \cdot \mathbf{E} \, dx - \int_{\partial\Omega_2^e} g_p (\boldsymbol{\psi}_4 \cdot \mathbf{n}) \, ds \\ - \int_{\Omega^e} f \phi_5 \, dx \end{bmatrix}.$$

In particular, for the aluminium DC-casting problem discussed in Chapter 2, we have

- $\lambda = K/A$,

- $f \equiv 0$,
- $g_v \equiv 0$,
- $g_p \neq 0$ on Γ_4,
- $\partial\Omega_1 = \Gamma_2 \cup \Gamma_4$,
- and $\partial\Omega_2 = \Gamma_1 \cup \Gamma_3 \cup \Gamma_5$.

Consequently, we obtain

$$\begin{aligned} a(\mathbf{u},\mathbf{v}) &= \int_\Omega \mathbf{u}\cdot\left(\frac{A}{K}\mathbf{v}\right) dx, \\ b(p,\mathbf{u}) &= -\int_\Omega p(\nabla\cdot\mathbf{u})\, dx \end{aligned}$$

and

$$\begin{aligned} \alpha(\mathbf{u}) &= -\int_\Omega \mathbf{u}\cdot\mathbf{E}\, dx - \int_{\Gamma_4}(1-BY_4)(\mathbf{u}\cdot\mathbf{n})\, ds, \\ \beta(p) &= 0. \end{aligned}$$

We observe that the element matrix $\mathbf{A}^e$ is symmetric.

Let $\mathbf{v}$ denote the solution vector for velocity, and $\mathbf{p}$ be the solution vector for pressure. The resulting linear system may then be written as

$$\begin{pmatrix} \mathbf{C} & \mathbf{B} \\ \mathbf{B}^T & 0 \end{pmatrix} \begin{pmatrix} \mathbf{v} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{R}_v \\ \mathbf{R}_p \end{pmatrix}. \tag{4.17}$$

The submatrix $\mathbf{C}$ will typically be symmetric and banded (tridiagonal or even diagonal in some special cases). However, the complete coefficient matrix is symmetric indefinite with a sparse structure. Such systems can be solved iteratively using indefinite generalizations of the conjugate gradient method, e.g., the minimum residual method proposed by Paige and Saunders [17]. When using an iterative solver, special caution should be taken to find a suitable preconditioner, see Rusten and Winther [22] and [23].

When solving the system (4.17) by a direct method, we should renumber the nodes such that the bandwidth of the matrix is minimized. In this case it will be better to merge the sequences of v-nodes and p-nodes.

4.3 The Implementation of the Mixed Finite Element Method in DIFFPACK

The present implementation of mixed FEM in DIFFPACK is a preliminary version that may be replaced by a more general code in the future. The class hierarchy for the mixed FEM is based on the one for the conforming FEM. The organization is also quite similar to that of the conforming FEM. Actually, it takes advantage of some of the tools for the conforming

FEM programming, and enforces some necessary developments. In order to understand the details discussed in this Section, one must be familiar with the concepts of the reports [5], [6], [7], [8] and [9]. A short introduction to the established procedures for finite element programming in `DIFFPACK` was given earlier in Section 3.2. Here we only describe briefly the implementation of the mixed FEM, with the discussion of the differences between the implementations.

4.3.1 Block-centered mixed finite elements and their trial functions

Consider the block-centered mixed finite elements, i.e., $k = 0$. The nodes defining pressure and velocity values for triangular elements and quadrilateral elements are indicated in Figure 4.2 and Figure 4.3, respectively, where the symbol "•" represents a pressure node, and the symbol "**X**" represents a velocity node. Consequently, the pressure is calculated at the centroid of the element, while the velocity nodes reside in the middle point of each element side.

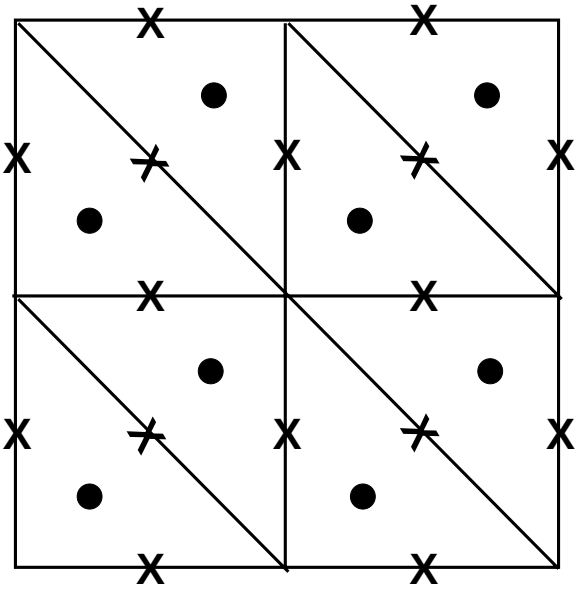


Figure 4.2: Triangular block-centered mixed finite elements in 2D.

When implementing mixed FEM in `DIFFPACK`, we have defined four new element types in the `ElmDef` hierarchy, two for triangular elements and two for quadrilateral elements. To be precise, `ElmT1n2D` and `ElmB1n2D` are pressure-node elements, while `ElmT3nS2D` and `ElmB4nS2D` are velocity-node elements. The numbering of the local nodes of the new element types is indicated in Figures 4.4-4.5.

The conforming FEM chooses the trial functions with local support in order to obtain a banded or sparse matrix structure. Usually, such a trial function equals 1 at its own node and vanishes at all other nodes. In this case, the degree of freedom associated with each node is simply the solution value in that particular point. We will now discuss possible trial functions for the mixed FEM, both for p-nodes (the nodes where the pressure is calculated) and v-nodes (the nodes where the velocity is calculated) associated with the

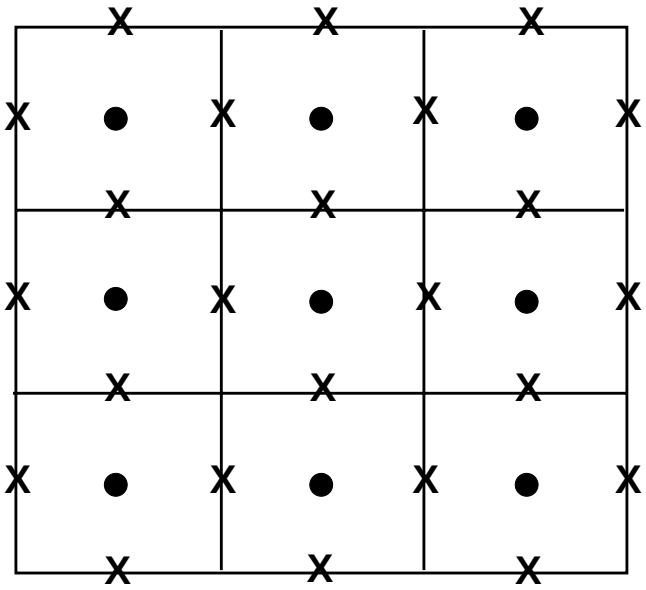


Figure 4.3: Quadrilateral block-centered mixed finite elements in 2D.

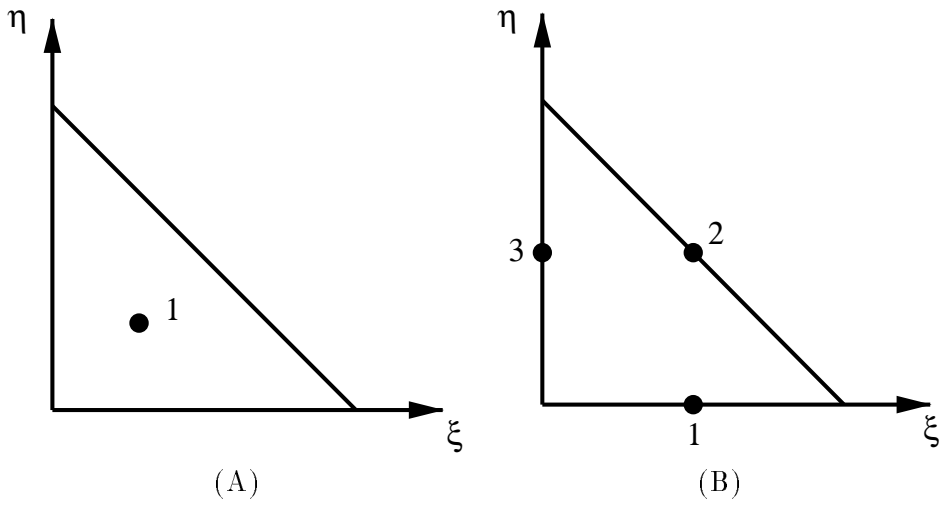


Figure 4.4: The local numbering of nodes for mixed triangular elements. (A) element type `ElmT1n2D`, (B) element type `ElmT3nS2D`.

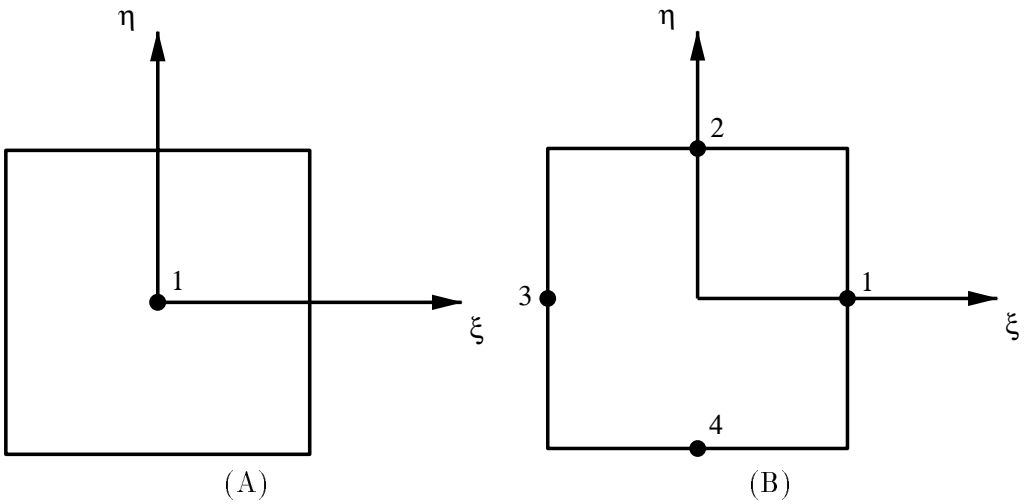


Figure 4.5: The local numbering of nodes for mixed quadrilateral elements. (A) element type `ElmB1n2D`, (B) element type `ElmB4nS2D`.

reference element $\widehat{\mathbf{K}}$.

The choice of the trial functions for p-nodes is simple. Since the pressure is required to be constant over any given element, one can simply choose the trial function to be 1 over the element and zero elsewhere.

The choice of the trial functions for v-nodes is a little more complicated. If the degree of freedom for v-nodes is chosen to be the value of $\mathbf{v} \cdot \mathbf{n}$ at the middle point of each side of the element, we want that

The trial function $\boldsymbol{\psi}_i$ for the node i in element e must satisfy:

- $\boldsymbol{\psi}_i \cdot \mathbf{n} = 1$ on the node i;
- $\boldsymbol{\psi}_i \cdot \mathbf{n} = 0$ on all other nodes.

The following vector-valued trial functions are used for the triangular reference element `ElmT3nS2D` (cf. Figure 4.4(B)):

$$\begin{aligned} \boldsymbol{\psi}_1 &= (\xi, \eta - 1), \\ \boldsymbol{\psi}_2 &= \left(\sqrt{2}\xi, \sqrt{2}\eta\right), \\ \boldsymbol{\psi}_3 &= (\xi - 1, \eta). \end{aligned}$$

It is easy to prove that these trial functions are linearly independent such that the value of $\mathbf{v}$ at a point (ξ, η) in the element can be evaluated by interpolation, i.e.,

$$v(\xi, \eta) = k_1\boldsymbol{\psi}_1 + k_2\boldsymbol{\psi}_2 + k_3\boldsymbol{\psi}_3$$

$$\begin{aligned} &= k_1(\xi, \eta - 1) + k_2\left(\sqrt{2}\xi, \sqrt{2}\eta\right) + k_3(\xi - 1, \eta) \\ &= \left((k_1 + \sqrt{2}k_2 + k_3)\xi - k_3, (k_1 + \sqrt{2}k_2 + k_3)\eta - k_1\right). \end{aligned}$$

For the quadrilateral elements, the trial functions associated with the reference element `ElmB4nS2D` (cf. Figure 4.5(B)) are:

$$\begin{aligned} \boldsymbol{\psi}_1 &= \left(\frac{1}{2}(\xi + 1), 0\right), \\ \boldsymbol{\psi}_2 &= \left(0, \frac{1}{2}(\eta + 1)\right), \\ \boldsymbol{\psi}_3 &= \left(\frac{1}{2}(\xi - 1), 0\right), \\ \boldsymbol{\psi}_4 &= \left(0, \frac{1}{2}(\eta - 1)\right). \end{aligned}$$

Obviously, these trial functions are also linearly independent. The value of $\mathbf{v}$ at a point (ξ, η) in the element can then be evaluated by

$$\begin{aligned} v(\xi, \eta) &= k_1\boldsymbol{\psi}_1 + k_2\boldsymbol{\psi}_2 + k_3\boldsymbol{\psi}_3 + k_4\boldsymbol{\psi}_4 \\ &= \left(\frac{1}{2}(\xi + 1)k_1 + \frac{1}{2}(\xi - 1)k_3, \frac{1}{2}(\eta + 1)k_2 + \frac{1}{2}(\eta - 1)k_4\right). \end{aligned}$$

Note that the ξ-component of $\mathbf{v}$ is decided only by $\boldsymbol{\psi}_1$ and $\boldsymbol{\psi}_3$, while the η-component is decided by $\boldsymbol{\psi}_2$ and $\boldsymbol{\psi}_4$.

4.3.2 The implementation of the mixed FEM

In addition to the extensions of the `ElmDef` hierarchy, some other classes have also been developed or generated. Since the C++ code for the implementation of the mixed FEM is too large to be included in the thesis, we simply introduce the basic ideas.

◀ **Class `ShapeFuncGrid`:**

Obviously, the implementations of the mixed FEM and the conforming FEM need different representations of spatial information. The conforming FEM relies on isoparametric grids, where the element vertices represent nodes with corresponding shape functions. However, for the mixed FEM, the vertices are only used for geometrical purposes. In order to represent the shape function nodes, we need the class `ShapeFuncGrid`. Roughly speaking, this class can be thought of as an overlay to its accompanying `GridFE`, thus providing the extra information that is needed by mixed FEM.

For our model problem, it is natural to allocate two `ShapeFuncGrid` objects, one for the pressure nodes and one for the velocity nodes. Since the basic geometry of these `ShapeFuncGrid`s remain unchanged, they both hold a reference to a common instance of `GridFE`.

In order to provide a general interface to grid information, a `ShapeFuncGrid` remembers whether it is isoparametric or not. If used by a conforming FEM, i.e., in isoparametric mode, the `ShapeFuncGrid` member functions are simply calls to the corresponding functions in class `GridFE`. Otherwise, it provides redefined or additional functionality not offered by `GridFE`.

◀ **Class `FiniteElement`:**

The class `FiniteElement` has been extended so that it can also been used for mixed finite element programming. A boolean variable is added to keep trace of the element type, i.e., whether it is isoparametric (conforming case) or non-isoparametric (mixed case). We have also added some member functions to the class in order to evaluate the trial functions for both pressure nodes and velocity nodes. The user should be very careful regarding the type match when calling these functions. Warning messages will be given if calls are issued to the wrong functions.

◀ **Class `MXDegFreeFE` :**

Class `MXDegFreeFE` is designed to represent the degrees of freedom in a mixed finite element field associated with a finite element grid (eventually, a `ShapeFuncGrid`). This class is conceptually similar to the class `DegFreeFE`. The present version merges element-wise the sequences of velocity nodes and pressure nodes, resulting in a banded matrix structure. The actual bandwidth will depend heavily on the element numbering in the original `GridFE` object.

◀ **Class `MXElmMatVec`:**

Class `MXElmMatVec` is almost a copy of the class `ElmMatVec`. The only difference is that it uses a class `MXDegFreeFE` object to represent the data for mapping between local and global degrees of freedom and the boundary conditions.

◀ **Class `MXFEM`:**

Similar to class `FEM`, a class `MXFEM` is created to handle the basic programming of mixed FEM. These two classes are almost identical, except for some small differences. The most important difference concerns the computation of element matrices and vectors. Typically, the contribution to the linear system from element number e will be

$$\begin{bmatrix} \mathbf{C}^e & \mathbf{B}^e \\ (\mathbf{B}^e)^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{v}^e \\ p^e \end{bmatrix} = \begin{bmatrix} \mathbf{R}_v^e \\ R_p^e \end{bmatrix}$$

in the mixed case. When using mixed quadrilateral elements, we have $\mathbf{C}^e \in \mathbb{R}^{4\times 4}$, $\mathbf{B}^e \in \mathbb{R}^{4\times 1}$, $\mathbf{v}^e \in \mathbb{R}^4$, $\mathbf{R}_v^e \in \mathbb{R}^4$, while p^e and R_p^e are scalars. Usually, $\mathbf{R}_v^e$ involves side integration over the sides where natural boundary conditions are given, while both $\mathbf{R}_v^e$ and R_P^e may involve integration over the element. That is, we need four integrands to calculate the element matrices for the mixed finite element. Compared to the single function `integrands` in class `FEM`, the `MXFEM` class assumes the availability of four user-defined functions.

4.4 Numerical Experiments

In this section, we conclude our study of mixed FEM by looking at three numerical experiments. In section 4.4.1, we will solve a pressure equation with both natural and essential boundary conditions. A mixed formulation will be given for this example, together with the numerical solutions posed on a simple geometry. In section 4.4.2, the Poisson equation is solved and compared with its analytical solution for a particular set of input data. For this problem, we also estimate the rate of convergence of the mixed FEM. In section 4.4.3, we look at the aluminium DC-casting problem. This is a non-trivial problem defined on a highly irregular geometry. Graphs illustrating the numerical results are included at the end of the section.

4.4.1 Solving the pressure equation on a square domain

We want to solve the pressure equation

$$\begin{aligned} \mathbf{v} &= -K\nabla p \qquad \text{in } \Omega, \\ -\nabla\cdot\mathbf{v} &= 1 \qquad \text{in } \Omega, \end{aligned}$$

where $\Omega = [0,1]\times[0,1]$. If we indicate the side $x_1 = 1$ as Γ_1, side $x_2 = 1$ as Γ_2, side $x_1 = 0$ as Γ_3 and side $x_2 = 0$ as Γ_4, cf. Figure 3.1 in Section 3.3.1, the boundary conditions are

$$\begin{aligned} p &= 1 \qquad \text{on } \Gamma_1, \\ \mathbf{v}\cdot\mathbf{n} &= 0 \qquad \text{on } \Gamma_2, \\ p &= 0 \qquad \text{on } \Gamma_3, \\ \mathbf{v}\cdot\mathbf{n} &= 0 \qquad \text{on } \Gamma_4. \end{aligned}$$

The permeability function $K(x_1,x_2) = x_1+1$ reduces the problem to one spatial dimension, for which the analytical solution is

$$\begin{aligned} p(x_1,x_2) &= x_1, \\ \mathbf{v}(x_1,x_2) &= (-x_1-1,0). \end{aligned}$$

We define the space

$$H_0(\mathrm{div};\Omega) = \{\mathbf{u}\in H(\mathrm{div};\Omega);\ \mathbf{u}\cdot\mathbf{n} = 0 \ \text{ on } \Gamma_2,\Gamma_4\}\,.$$

The corresponding weak formulation is then given by (4.6) with $\lambda\equiv K$, $\mathbf{E}\equiv(0,0)^T$, $f\equiv 1$, $\partial\Omega_1 = \Gamma_1\cup\Gamma_3$ and $\partial\Omega_2 = \Gamma_2\cup\Gamma_4$, i.e.,

Find a pair of functions $(p,\mathbf{v})\in L^2(\Omega)\times H_0(\mathrm{div};\Omega)$ such that

$$\begin{aligned} \int_\Omega \mathbf{u}\cdot\frac{\mathbf{v}}{K}\,dx - \int_\Omega p\nabla\cdot\mathbf{u}\,dx &= -\int_{\Gamma_1}\mathbf{u}\cdot\mathbf{n}\,ds \qquad \forall\mathbf{u}\in H_0(\mathrm{div};\Omega), \\ -\int_\Omega q\nabla\cdot\mathbf{v}\,dx &= \int_\Omega q\,dx \qquad \forall q\in L^2(\Omega). \end{aligned}$$

Discretization of the problem implies the definition of two finite-dimensional subspaces Q_h and U_h, such that

$$Q_h \subset L^2(\Omega); \qquad U_h \subset H_0(\mathrm{div};\Omega);$$

The problem then becomes

Find a pair of functions $(p, \mathbf{v}) \in Q_h \times U_h$ such that

$$\begin{aligned} \int_\Omega \mathbf{u}_h \cdot \frac{\mathbf{v}_h}{K}\, dx - \int_\Omega p_h \nabla \cdot \mathbf{u}_h\, dx &= -\int_{\Gamma_1} \mathbf{u}_h \cdot \mathbf{n}\, ds \qquad \forall \mathbf{u}_h \in U_h, \\ -\int_\Omega q_h \nabla \cdot \mathbf{v}_h\, dx &= \int_\Omega q_h\, dx \qquad \forall q_h \in Q_h. \end{aligned}$$

As in section 4.2.6, we denote the trial functions for Q_h (pressure) and U_h (velocity) by ϕ_i and $\boldsymbol{\psi}_i = (\psi_{1,i}, \psi_{2,i})$, respectively. In the case of quadrilateral mixed elements, the element contributions will be the same as in (4.16) with $\lambda \equiv K$, $\mathbf{E} \equiv (0,0)^T$, $f \equiv 1$ and appropriate boundary indicators.

In this particular case, where each element is a square, we obtain $\boldsymbol{\psi}_1 \cdot \boldsymbol{\psi}_3 = \boldsymbol{\psi}_2 \cdot \boldsymbol{\psi}_4 = 0$, and $\int_{\Omega^e} \phi \nabla \cdot \boldsymbol{\psi} dx = 2$ for constant mesh size h in both directions.

Since the numerical solutions for both the pressure and the velocity are only continuous on each element, and discontinuous on the complete domain, it is convenient to evaluate these entities at the centroid of each element. Using blocked-center finite elements, the numerical method should be able to compute the analytical solution in the nodes. This is confirmed by the actual computer runs. When dividing the unit square with a 4x4 partition, we compute the solution with maximum errors as $4.16 \cdot 10^{-16}$ and $(4.44 \cdot 10^{-16}, 6.66 \cdot 10^{-16})$ for the pressure and the velocity, respectively.

4.4.2 Estimating the rate of convergence for the Poisson equation

Consider the Poisson equation

$$-\Delta p = 2\pi^2 \sin(\pi x_1) \sin(\pi x_2) \qquad \text{in } \Omega,$$

where $\Omega = [0,1] \times [0,1]$. By enforcing homogeneous Dirichlet boundary conditions, i.e., $p = 0$, on the whole boundary, the analytical solution is

$$\begin{aligned} p(x_1, x_2) &= \sin(\pi x_1) \sin(\pi x_2), \\ \mathbf{v}(x_1, x_2) &= -\nabla p = (-\pi \cos(\pi x_1) \sin(\pi x_2),\ -\pi \sin(\pi x_1) \cos(\pi x_2)). \end{aligned}$$

Clearly, this is a special case of the general problem discussed in Section 4.1 and 4.2 with $\lambda \equiv 1$, $f = 2\pi^2 \sin(\pi x_1) \sin(\pi x_2)$ and $\partial\Omega_2 = \emptyset$.

The study of the rate of convergence of the mixed method is associated with different choices of norms. Let $|\mathbf{v}| = \sqrt{v_1^2 + v_2^2}$ be the vector norm, we then define discrete norms for the scalar-valued function p and the vector-valued function $\mathbf{v}$ in 2D as

$$\begin{aligned} \|p\|_{l^\infty(\Omega)} &= \max_{i=1}^{N} |p_i|, \\ \|\mathbf{v}\|_{(l^\infty(\Omega))^2} &= \max_{i=1}^{N} |\mathbf{v}_i|, \end{aligned}$$

and

$$\begin{aligned} \|p\|_{l^2(\Omega)} &= \left[\frac{1}{N}\sum_{i=1}^{N}(p_i)^2\right]^{\frac{1}{2}}, \\ \|\mathbf{v}\|_{(l^2(\Omega))^2} &= \left[\frac{1}{N}\sum_{i=1}^{N}|\mathbf{v}_i|^2\right]^{\frac{1}{2}}. \end{aligned}$$

Here N is the number of elements in the grid, while p_i and $\mathbf{v}_i$ denote the values evaluated at the centroid of element number i. For the continuous norms, in addition to the corresponding norms associated with the spaces, we define

$$\begin{aligned} \|p\|_{L^2(\Omega)} &= \left[\int_\Omega p^2\,dx\right]^{\frac{1}{2}}, \\ \|\mathbf{v}\|_{(L^2(\Omega))^2} &= \left[\int_\Omega |\mathbf{v}|^2\,dx\right]^{\frac{1}{2}}. \end{aligned}$$

According to Weiser and Wheeler [27], the mixed FEM has its super rate of convergence at the nodes. In this particular case, when we use $H^0(\Omega)$ and $H(\mathrm{div};\Omega)$ for pressure and velocity, the rate of convergence for the discrete norms, which evaluates only nodal values, will be 2, while the continuous norms will have 1 as their rate of convergence, cf. (4.13).

Let p_h and $\mathbf{v}_h$ denote the numerical solutions obtained by the mixed FEM, the errors and the estimated rates of convergence of the numerical solutions measured in these norms are given in Tables 4.1 and 4.2. The results agree very well with the expectation.

4.4.3 The aluminium DC-casting problem – the pressure equation posed on an irregular geometry

The mixed formulation of the aluminium DC-casting problem is given in Section 4.2. In this section, we present the numerical solutions graphically, cf. Figures 4.6-4.10. These results are obtained by computation on the grid of block-centered finite elements, which takes a 5×5 partition on each of the 9 super elements, with the linear system solved by Gauss Elimination. If we compare these results with those obtained by a conforming finite element method in Section 3.3.3 (cf. Figures 3.3-3.8), we observe that these two solutions are very close to each other. Figures of the critical region for pressure and velocity solutions are also given. In order to present the velocity solutions more clearly, we again include a plot of streamlines in the critical region.

For the explanation of the numerical results we refer to Section 3.3.3.

The errors:			
h(s)	$\|p-p_h\|_{L^2(\Omega)}$	$\|p-p_h\|_{l^2(\Omega)}$	$\|p-p_h\|_{l^\infty(\Omega)}$
1/2	2.94614e-01	8.87665e-02	8.87665e-02
1/4	1.56928e-01	2.48895e-02	4.24891e-02
1/8	7.97315e-02	6.37567e-03	1.22660e-02
1/16	4.00261e-02	1.60328e-03	3.17575e-03
1/32	2.00331e-02	4.01402e-04	8.00870e-04
1/64	1.00191e-02	1.00387e-04	2.00653e-04
The rates of convergence:			
$1/2 \sim 1/4$	0.9087	1.8345	1.0629
$1/4 \sim 1/8$	0.9769	1.9649	1.7924
$1/8 \sim 1/16$	0.9942	1.9916	1.9495
$1/16 \sim 1/32$	0.9986	1.9979	1.9875
$1/32 \sim 1/64$	0.9996	1.9995	1.9969

Table 4.1: The errors and the rates of convergence of the numerical solutions for the pressure.

The errors:				
h(s)	$\|\mathbf{v}-\mathbf{v}_h\|_{(L^2(\Omega))^2}$	$\|\mathbf{v}-\mathbf{v}_h\|_{(l^2(\Omega))^2}$	$\|\mathbf{v}-\mathbf{v}_h\|_{(l^\infty(\Omega))^2}$	$\|\mathbf{v}-\mathbf{v}_h\|_{H(\mathrm{div};\Omega)}$
1/2	9.83419e-01	4.76725e-01	4.76725e-01	6.19303e+00
1/4	4.99654e-01	1.15383e-01	1.41315e-01	3.17706e+00
1/8	2.51298e-01	2.86215e-02	3.89668e-02	1.59877e+00
1/16	1.25847e-01	7.14155e-03	1.00031e-02	8.00670e-01
1/32	6.29486e-02	1.78453e-03	2.51763e-03	4.00496e-01
1/64	3.14755e-02	4.46078e-04	6.30470e-04	2.00268e-01
The rates of convergence:				
$1/2 \sim 1/4$	0.9769	2.0467	1.7542	0.9630
$1/4 \sim 1/8$	0.9915	2.0113	1.8586	0.9907
$1/8 \sim 1/16$	0.9977	2.0028	1.9618	0.9977
$1/16 \sim 1/32$	0.9994	2.0007	1.9903	0.9994
$1/32 \sim 1/64$	0.9999	2.0002	1.9976	0.9999

Table 4.2: The errors and the rates of convergence of the numerical solutions for the velocity.

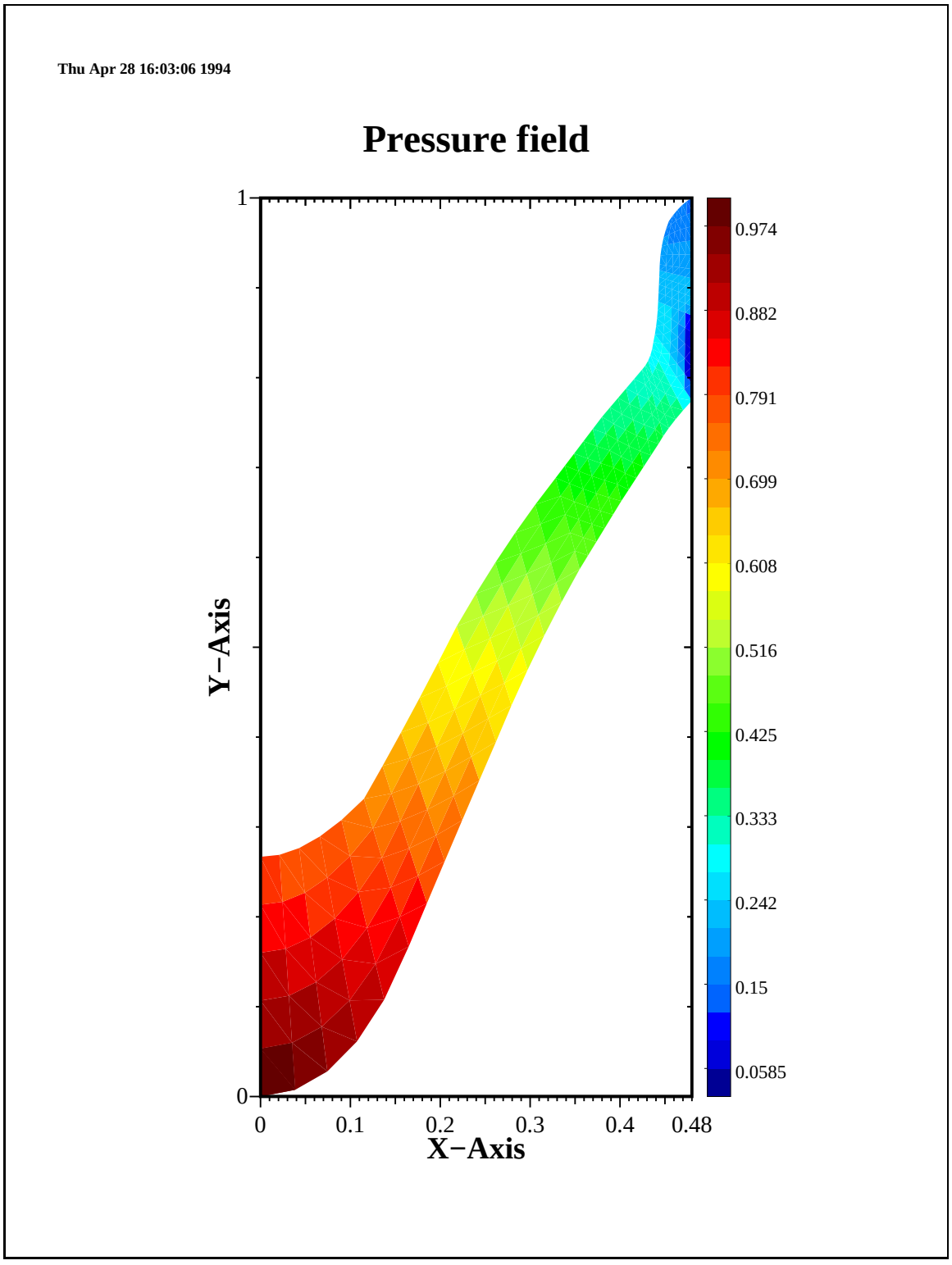


Figure 4.6: Plot of the mixed solution of the pressure P.

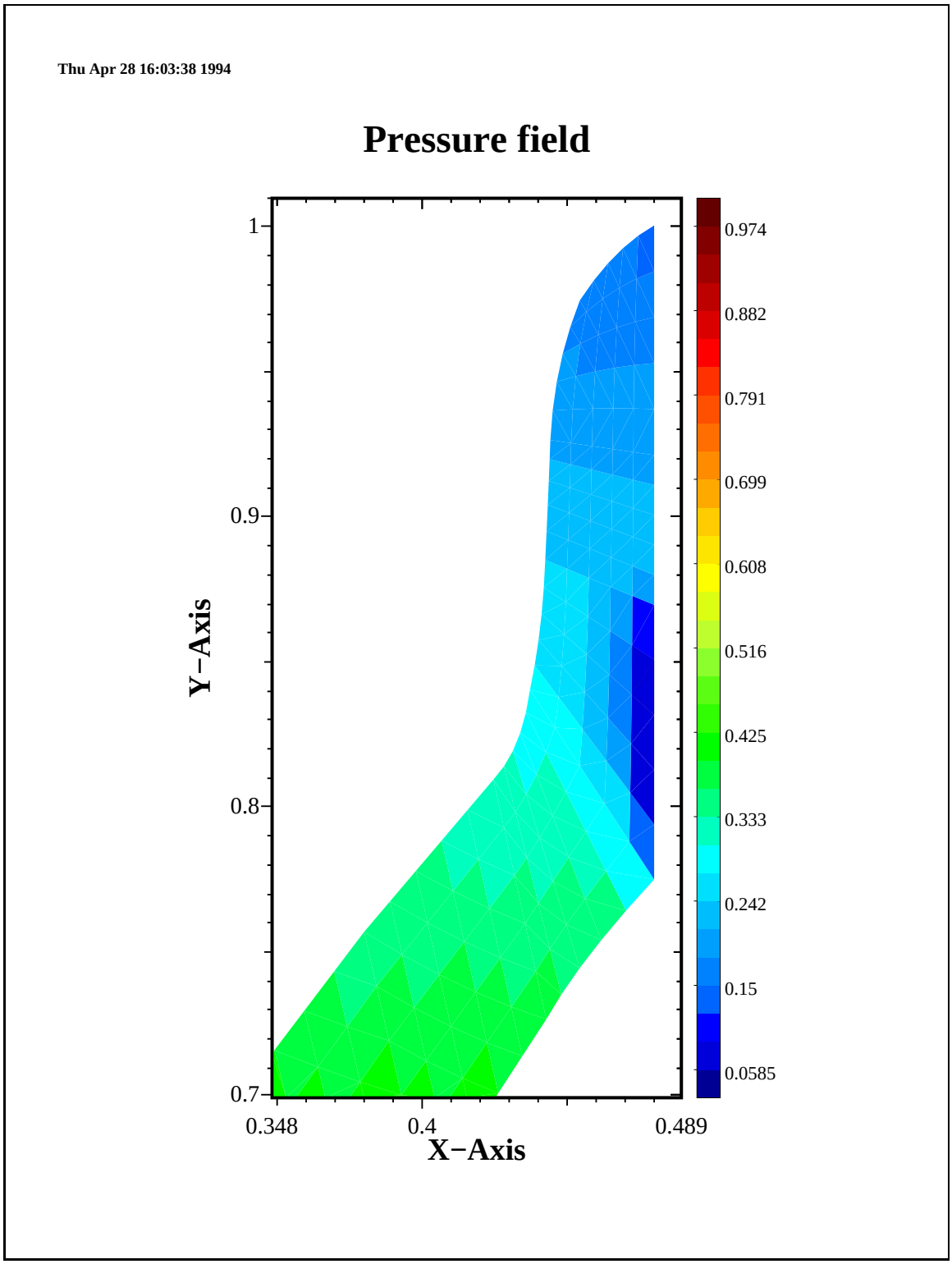


Figure 4.7: Plot of the mixed solution of P in the critical region.

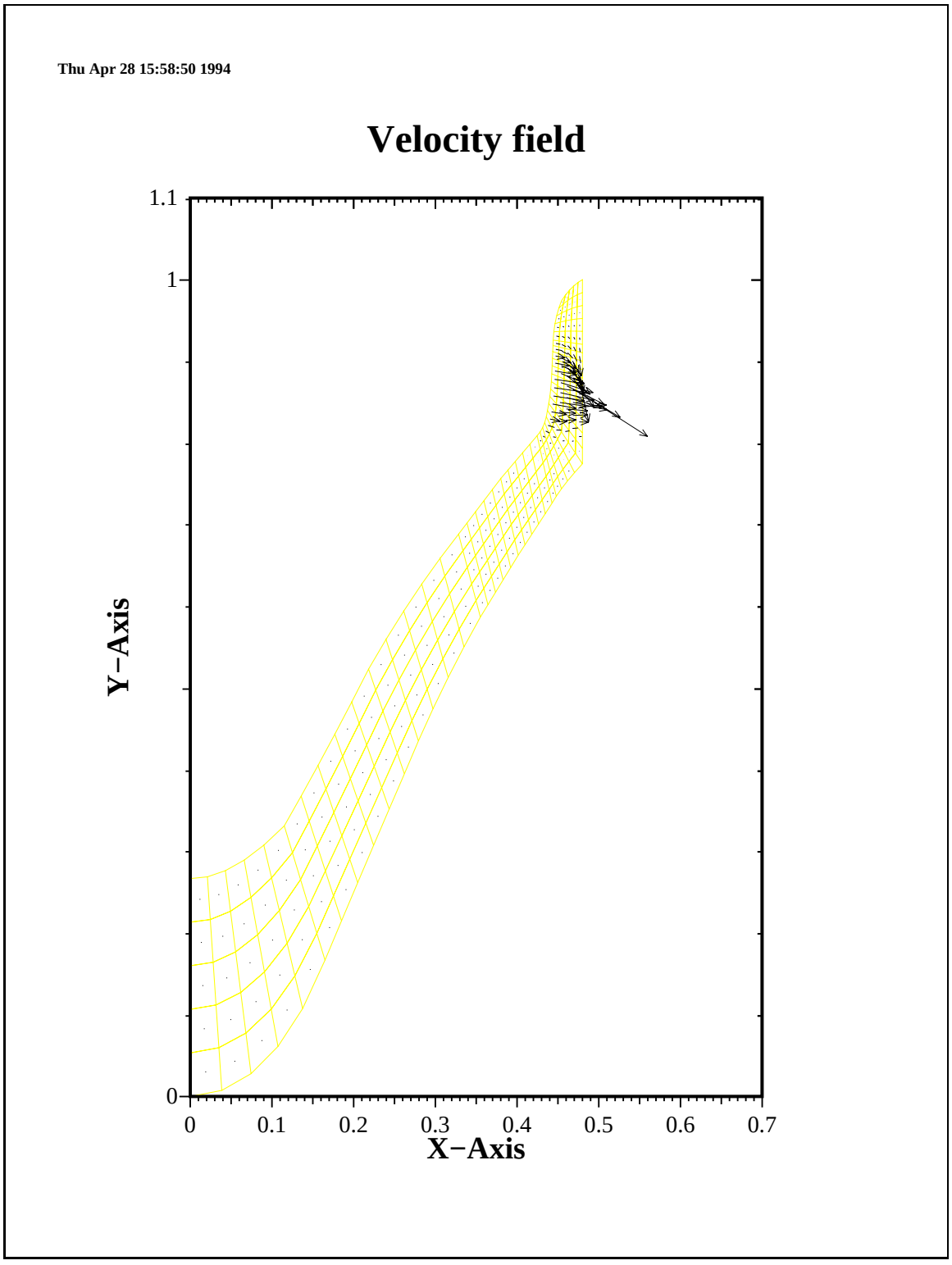


Figure 4.8: Plot of the mixed solution of the relative superficial velocity $\mathbf{v}$.

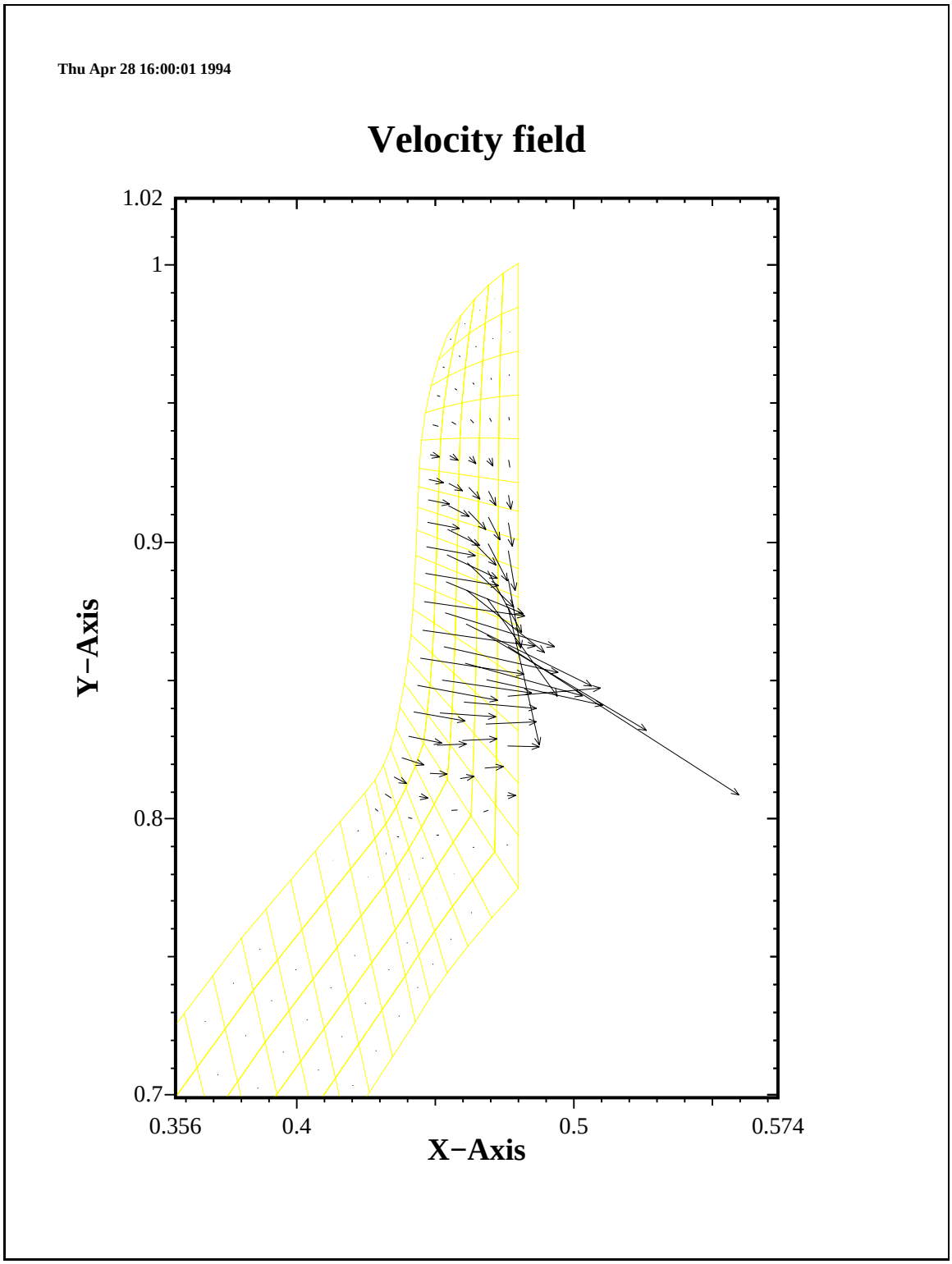


Figure 4.9: Plot of the mixed solution of **v** in the critical region.

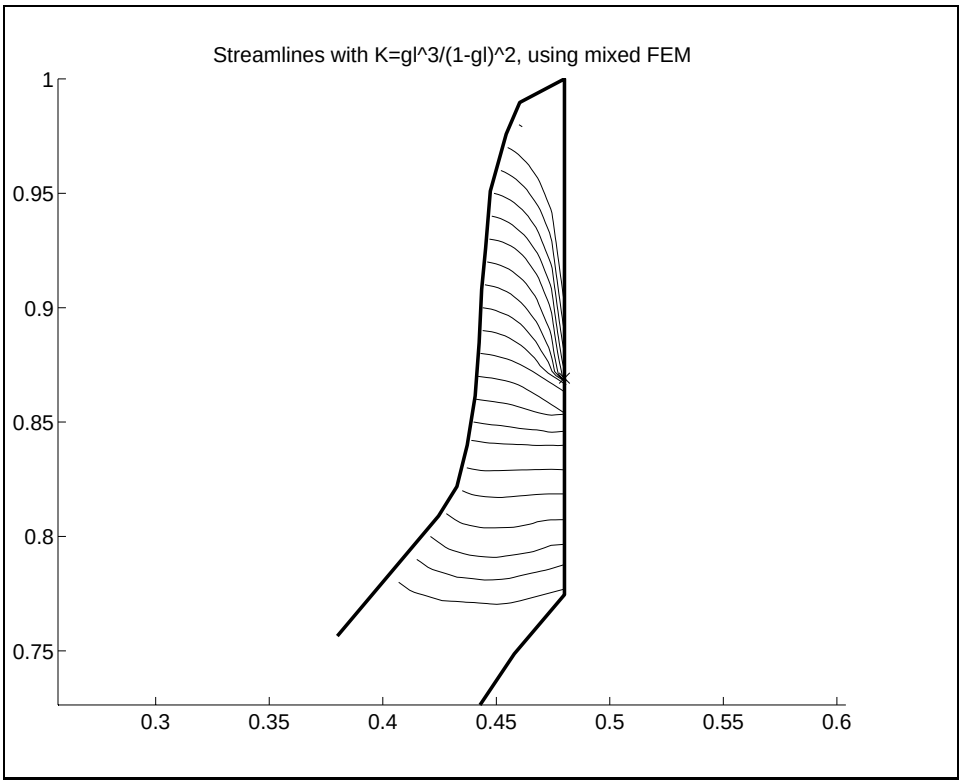
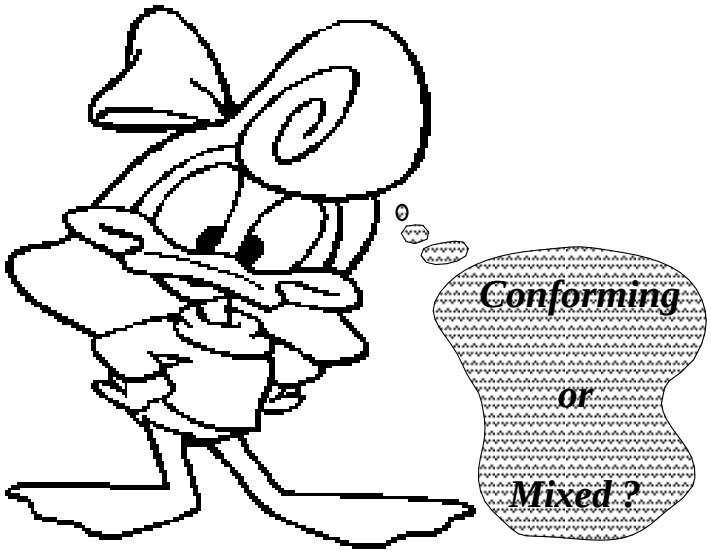


Figure 4.10: Plot of streamlines of in the critical region, mixed FEM.

Chapter 5

Comparing Conforming and Mixed FEM



From the discussions in the previous Chapters, we observe that both the conforming and the mixed FEM can produce numerical results which approximate the analytical solutions quite well. However, the differences between them are obvious. Consider the solution of the pressure in a general elliptic problem. In the conforming FEM, if we use piecewise bilinear trial functions, the rate of convergence is 2 in L^2-norm and 1 in H^1-norm, while in the mixed FEM, we obtain 1 as rate of convergence in L^2-norm and 2 in discrete l^2-norm with piecewise constant trial functions, provided that the solution p is "smooth" enough.

Based on similar observations, we compare these two methods in this Chapter. In Section 5.1, we give a brief comparison between the formulations, the rates of convergence, the linear systems and some physical aspects of the two methods. Further, we study the rates of convergence with some numerical experiments in Section 5.2, and end this Chapter with the concluding remarks.

5.1 The Relationship Between Conforming and Mixed FEM

5.1.1 The weak formulations

Consider the general elliptic boundary value problem (1.1) presented in Chapter 1

$$-\nabla\cdot(\lambda(\nabla p+\mathbf{E}))=f \qquad \text{in } \Omega\in\mathbb{R}^2, \tag{5.1}$$

subject to the boundary conditions

$$p = g_p \qquad \text{on } \partial\Omega_1, \tag{5.2}$$
$$\mathbf{v}\cdot\mathbf{n} = g_v \qquad \text{on } \partial\Omega_2, \tag{5.3}$$

where $\mathbf{v}=-\lambda\,(\nabla p+\mathbf{E})$.

Recall the weak formulation of the problem in the conforming FEM (cf. 3.4): Find $p\in H_g^1(\Omega)$ such that

$$\int_\Omega \nabla q\cdot(\lambda\nabla p)\,dx=\int_\Omega fq\,dx-\int_\Omega\nabla q\cdot(\lambda\mathbf{E})\,dx-\int_{\partial\Omega_2}g_v q\,ds \qquad \forall q\in H_0^1(\Omega), \tag{5.4}$$

where

$$H_0^1(\Omega) = \{p : p\in H^1(\Omega);\, p=0 \text{ on } \partial\Omega_1\},$$
$$H_g^1(\Omega) = \{p : p\in H^1(\Omega);\, p=g_p \text{ on } \partial\Omega_1\},$$

and the corresponding weak formulation in the mixed FEM (cf. (4.6)): Find a pair of functions $(p,\mathbf{v})\in L^2(\Omega)\times\tilde{H}(\mathrm{div};\Omega)$ such that:

$$\int_\Omega \mathbf{u}\cdot\left(\frac{1}{\lambda}\mathbf{v}\right)dx,-\int_\Omega p\nabla\cdot\mathbf{u}\,dx = -\int_\Omega\mathbf{u}\cdot\mathbf{E}\,dx-\int_{\partial\Omega_1}g_p(\mathbf{u}\cdot\mathbf{n})\,ds \qquad \forall\mathbf{u}\in H_0(\mathrm{div};\Omega), \tag{5.5}$$

$$-\int_\Omega q\nabla\cdot\mathbf{v}\,dx = -\int_\Omega fq\,dx \qquad \forall q\in L^2(\Omega), \tag{5.6}$$

where

$$\tilde{H}(\mathrm{div};\Omega) = \{\mathbf{u} \in H(\mathrm{div};\Omega);\ \mathbf{u}\cdot\mathbf{n} = g_v \ \text{ on } \partial\Omega_2\}.$$

The weak formulations of the two methods appear to be equivalent since it seems as if one can obtain the corresponding conforming formulation (4.6) by eliminating $\mathbf{v}$ in the mixed formulation defined by (5.5) and (5.6). Unfortunately, the out look is not correct.

The most essential difference lies in the nature of spaces for the solutions. We can observe it by examining the pressure solution from the two methods. In the conforming FEM, the term "∇p" appears in the formulation, and we hereby require $\nabla p \in L^2(\Omega)$. Hence the smallest space for the pressure solution is $H^1(\Omega)$. Similarly, $L^2(\Omega)$ is the smallest space for p in the mixed FEM since only p appears in the formulation. Consequently, the spaces for velocity solutions are also different. Moreover, we have

$$\text{"pressure space"} \ \subset \ \text{"velocity component space"}$$

in the conforming FEM, and

$$\text{"velocity component space"} \ \subset \ \text{"pressure space"}$$

in the mixed FEM. This means that no matter how we choose the spaces, the two solutions will never be equivalent.

It is easy to see that this also applies to the discrete problems (3.13) and (4.7). Therefore, the mixed formulation problem is a *non-conforming displacement* model for solving the elliptic problem (5.1), cf. Raviart and Thomas [19]. For other non-conforming methods based on hybrid models, we refer to Roberts and Thomas [20].

5.1.2 The rates of convergence

Generally, if two different numerical solutions both converge towards the exact solution in a specific norm, then the numerical solutions must also converge towards each other in the same norm. Let p be the analytical solution, while p^c and p^m are the numerical solutions of the conforming and the mixed FEM respectively, we have

$$\|p^c - p^m\| = \|(p^c - p) + (p^m - p)\| \le \|p - p^c\| + \|p - p^m\|. \tag{5.7}$$

It is clear that when both p^c and p^m converge towards p in $\|\cdot\|$, they will at the same time converge towards each other in the same norm. However, the rate of convergence for $\|p^c - p^m\|$ is then dependent on both the rates of convergence for the two methods.

We take L^2-norm as an example. As the standard error estimate indicates, (cf. (3.14) and (4.13)) when the conditions in the Lax-Milgram Theorem are satisfied and the analytical solution is "smooth" enough, we have

$$\|p - p^c\|_{L^2(\Omega)} = \mathcal{O}(h^2) \quad \text{and} \quad \|p - p^m\|_{L^2(\Omega)} = \mathcal{O}(h),$$

provided that we use piecewise linear trial functions for $p^c \in H^1(\Omega)$ and piecewise constant trial functions for $p^m \in H^0(\Omega)$. Then (5.7) leads to

$$\|p^c - p^m\|_{L^2(\Omega)} \le \mathcal{O}(h^2) + \mathcal{O}(h) = \mathcal{O}(h),$$

which means that the rate of convergence for $\|p^c - p^m\|_{L^2(\Omega)}$ will be at least 1. The case for the solution of $\mathbf{v}$ can be studied in the same way. In the conforming FEM, $\mathbf{v}$ is calculated by taking gradient of the conforming finite element solution of p, which indicates that its rate of convergence will be of one order lower than that for the pressure solution, i.e.,

$$\|\mathbf{v} - \mathbf{v}^c\|_{(L^2(\Omega))^2} = \mathcal{O}(h),$$

where $\mathbf{v}$ is the analytical solution and $\mathbf{v}^c$ is the conforming solution. In the mixed FEM, $\mathbf{v}$ is also a primary unknown of the formulation, so it has the same rate of convergence as p. With $\mathbf{v}^m \in H(\mathrm{div};\Omega)$ being the corresponding mixed solution, we have

$$\|\mathbf{v} - \mathbf{v}^m\|_{(L^2(\Omega))^2} = \mathcal{O}(h).$$

It is then obvious that

$$\|\mathbf{v}^c - \mathbf{v}^m\|_{(L^2(\Omega))^2} \le \|\mathbf{v} - \mathbf{v}^c\|_{(L^2(\Omega))^2} + \|\mathbf{v} - \mathbf{v}^m\|_{(L^2(\Omega))^2} = \mathcal{O}(h) + \mathcal{O}(h) = \mathcal{O}(h).$$

The rates of convergence in the discrete norms can be studied similarly, so we drop the detail here. By introducing the notations

$$\begin{aligned} e_p^c = p - p^c \quad &, \quad e_{\mathbf{V}}^c = \mathbf{v} - \mathbf{v}^c, \\ e_p^m = p - p^m \quad &, \quad e_{\mathbf{V}}^m = \mathbf{v} - \mathbf{v}^m, \\ e_p = p^m - p^c \quad &, \quad e_{\mathbf{V}} = \mathbf{v}^m - \mathbf{v}^c, \end{aligned}$$

we summarize some of our conclusions in Table 5.1.

norm	e_p^c	e_p^m	e_p	$e_{\mathbf{V}}^c$	$e_{\mathbf{V}}^m$	$e_{\mathbf{V}}$
L^2-norm	$\mathcal{O}(h^2)$	$\mathcal{O}(h)$	$\mathcal{O}(h)$	$\mathcal{O}(h)$	$\mathcal{O}(h)$	$\mathcal{O}(h)$
l^2-norm(Discrete)	$\mathcal{O}(h^2)$	$\mathcal{O}(h^2)$	$\mathcal{O}(h^2)$	$\mathcal{O}(h^2)$	$\mathcal{O}(h^2)$	$\mathcal{O}(h^2)$
l^∞-norm((Discrete)	$\mathcal{O}(h^2)$	$\mathcal{O}(h^2)$	$\mathcal{O}(h^2)$	$\mathcal{O}(h^2)$	$\mathcal{O}(h^2)$	$\mathcal{O}(h^2)$

Table 5.1: Summary of the convergence.

Generally, if we use polynomials of degree m as the trial functions for p in the conforming FEM, the standard error estimate gives:

$$\|p - p_h^c\|_{L^2(\Omega)} \le Ch^{m+1}|p|_{H^{m+1}(\Omega)}, \qquad \|p - p_h^c\|_{H^1(\Omega)} \le Ch^m|p|_{H^{m+1}(\Omega)},$$

provided that the solution p is "smooth" enough, i.e., $p \in H^{m+1}(\Omega)$, cf. Johnson [12]. In addition, since $\mathbf{v}$ is obtained by taking the gradient of solution p, its accuracy is weakened:

$$\|\mathbf{v} - \mathbf{v}_h^c\|_{(L^2(\Omega))^2} \le Ch^m|p|_{H^{m+1}(\Omega)}.$$

In the mixed FEM, while polynomials of degree m are used as the trial functions for p and the trial functions $\boldsymbol{\psi}$ for $\mathbf{v}$ satisfy $\nabla \cdot \boldsymbol{\psi} = P_m$, we have the standard error estimate as (cf. Falk and Osborn [10] and Ewing and Wheeler [3])

$$\|p - p_h^m\|_{L^2(\Omega)} \le Ch^{m+1},$$

and

$$\|\mathbf{v} - \mathbf{v}_h^m\|_{H(\mathrm{div};\Omega)} \le Ch^{m+1}, \quad \|\mathbf{v} - \mathbf{v}_h^m\|_{(L^2(\Omega))^2} \le Ch^{m+1},$$

provided that the solutions p and $\mathbf{v}$ are smooth enough. Here we have the same accuracy for both p and $\mathbf{v}$ solutions, which means that the rate of convergence for $\mathbf{v}$ solution is higher here than that in the conforming FEM. This indicates that the mixed FEM is superior to conforming FEM in case of seeking the velocity solution.

5.1.3 The linear system of equations

In both the methods, given a suitable triangulation of Ω and a corresponding set of trial functions, we end up with a system of linear algebraic equations. Remember for the conforming FEM, the stiffness matrix $A = (a_{ij})$ has the presentation

$$a_{ij} = \int_{\Omega} \nabla\phi_i \cdot (\lambda \nabla\phi_j)\, dx,$$

where ϕ_i and ϕ_j are the trial functions. It is easy to prove that this system is symmetric and positive definite. The system can thus be solved iteratively by the conjugate gradient method. However, the mixed FEM results in a symmetric and indefinite system as (cf. Section 4.2.6)

$$\begin{pmatrix} \mathbf{C} & \mathbf{B} \\ \mathbf{B}^T & 0 \end{pmatrix} \begin{pmatrix} \mathbf{v} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{R}_v \\ \mathbf{R}_p \end{pmatrix},$$

which is much more difficult to solve because of the following reasons: First, we have twice as many unknowns as in the conforming method. Second, when using an iterative solver such as the indefinite generalizations of the conjugate gradient method, we need a special preconditioner. (For more details about the iterative method for the indefinite system, we refer to Rusten and Winther [22] and [23].) This is hence regarded as a disadvantage of the mixed FEM.

5.1.4 Physical aspects

Generally, the mixed formulation follows the local mass conservation and gives a better physical interpretation than the conforming formulation. By choosing $\mathbf{v} \cdot \mathbf{n}$ at the middle point of a side as the degree of freedom for $\mathbf{v}$ in the mixed FEM, we demand that $\mathbf{v} \cdot \mathbf{n}$ remains continuous over the common side between two neighbour elements, which means that the amount of mass that flows out from one element is the same as that flows into the neighbour element through that common side.

However, the conforming FEM does not follow this kind of mass conservation. It is obvious that after using piecewise bilinear trial functions for p, we will get $\mathbf{v}$ that is only continuous in one element because $\mathbf{v}$ is obtained by taking gradient of the p solution. In other words, $\mathbf{v} \cdot \mathbf{n}$ will be discontinuous over the common side between two elements. Consequently, we lose the mass conservation in the discrete problem. Take square elements for instance, when piecewise bilinear functions is used for p, $v_1 = -\dfrac{\partial p}{\partial x_1}$ will be piecewise constant in x_1 and piecewise linear in x_2. Over a side that is vertical to x_1-axe, i.e., $\mathbf{n} = (1, 0)$, $\mathbf{v} \cdot \mathbf{n} = v_1$ will be clearly discontinuous there since v_1 is discontinuous in x_1 direction.

Furthermore, the equilibrium equation

$$\nabla \cdot \mathbf{v} = f \quad \text{in } \Omega$$

is not guaranteed to suffice in the conforming FEM since taking gradient $(\nabla\cdot)$ of a discontinuous function $\mathbf{v}$ will not result in a function $f \in L^2(\Omega)$. While in the mixed FEM, this equilibrium equation is included directly in the formulation.

5.2 Numerical Experiments

We continue our comparison by performing some numerical experiments. Consider the problem

$$-\nabla \cdot (\lambda \nabla p) = f \qquad \text{in } \Omega,$$

where $\Omega = [0,1] \times [0,1]$. The boundary conditions are

$$\begin{aligned} p &= g_p \qquad \text{on } \partial\Omega_1, \\ \mathbf{v} \cdot \mathbf{n} &= 0 \qquad \text{on } \partial\Omega_2, \end{aligned}$$

where $\mathbf{v} = -\lambda \nabla p$. Hence the conforming finite element formulation of the problem is: Find $p \in g_p + H_0^1(\Omega)$ such that

$$\int_\Omega \nabla q \cdot (\lambda \nabla p)\, dx = \int_\Omega f q \, dx$$

for all $q \in H_0^1(\Omega)$. The corresponding mixed formulation can then be: Find a pair of functions $(p, \mathbf{v}) \in L^2(\Omega) \times H_0(\mathrm{div}; \Omega)$ such that

$$\begin{aligned} \int_\Omega \mathbf{u} \cdot \left(\frac{1}{\lambda}\mathbf{v}\right) dx - \int_\Omega p \nabla \cdot \mathbf{u}\, dx &= -\textstyle\int_{\partial\Omega_1} g_p (\mathbf{u} \cdot \mathbf{n})\, ds \qquad \forall \mathbf{u} \in H_0(\mathrm{div}; \Omega), \\ -\int_\Omega q \nabla \cdot \mathbf{v}\, dx &= -\textstyle\int_\Omega f q \, dx \qquad \forall q \in L^2(\Omega). \end{aligned}$$

Here $H_0^1(\Omega)$, $L^2(\Omega)$ and $H_0(\mathrm{div}; \Omega)$ are the same as defined in the previous Chapters.

We study the rates of convergence by solving two cases with both methods:

♠ **Case 1:**

With

$$\lambda \equiv 1, \quad f(x_1, x_2) = 2\pi^2 \sin(\pi x_1) \sin(\pi x_2)$$

and

$$g_p \equiv 0, \quad \partial\Omega_1 = \partial\Omega, \quad \partial\Omega_2 = \emptyset,$$

we have the analytical solutions:

$$\begin{aligned} p &= \sin(\pi x_1) \sin(\pi x_2), \\ \mathbf{v} &= -\nabla p = \begin{bmatrix} -\pi \cos(\pi x_1) \sin(\pi x_2) \\ -\pi \sin(\pi x_1) \cos(\pi x_2) \end{bmatrix}. \end{aligned}$$

■

♠ **Case 2:**

With

$$\begin{aligned} \lambda(x_1, x_2) &= x_1 + 1, \\ f(x_1, x_2) &= \pi \sin(\pi x_2) \left[2\pi(1 + x_1)\cos(\pi x_1) + \sin(\pi x_1)\right], \end{aligned}$$

and the boundary conditions

$$\begin{aligned} g_p &\equiv 0 \qquad \text{on } \partial\Omega_1 = \Gamma_2 \cup \Gamma_4, \\ \mathbf{v} \cdot \mathbf{n} &= 0 \qquad \text{on } \partial\Omega_2 = \Gamma_1 \cup \Gamma_3, \end{aligned}$$

where we indicate the side $x_1 = 1$ as Γ_1, side $x_2 = 1$ as Γ_2, side $x_1 = 0$ as Γ_3 and side $x_2 = 0$ as Γ_4, (cf. Figure 3.1) we have the analytical solutions

$$\begin{aligned} p &= \cos(\pi x_1)\sin(\pi x_2), \\ \mathbf{v} &= -\lambda \nabla p = \begin{bmatrix} \pi(x_1+1)\sin(\pi x_1)\sin(\pi x_2) \\ -\pi(x_1+1)\cos(\pi x_1)\cos(\pi x_2) \end{bmatrix}. \end{aligned}$$

■

In the conforming FEM, we solve the small linear systems with Gauss Elimination and the large linear systems with conjugate gradient method. In the mixed FEM, Gauss Elimination is used to solved the linear system. The results are listed in Tables 5.2-5.3. We see that these results verify our conclusions in Table 5.1.

5.3 Concluding Remarks

So far, we can claim that both the conforming and the mixed finite element methods behave well in most of the numerical cases. Applied to the aluminium DC-casting model problem, both the methods give good results, and the two results are actually very close to each other even though the finite-dimensional subspaces are different. (Remember that we use piecewise constant trial functions for p in the mixed FEM and piecewise linear trial functions in the conforming FEM.) Furthermore, these two results converge towards each other as shown in the previous two Sections. However, the differences between the two methods are also obvious.

The main disadvantage of the conforming FEM is that only p is obtained when solving the variational problem. In many model problems, $\mathbf{v}$ has also important physical meanings, and sometimes $\mathbf{v}$ is of more interest than p. Using conforming FEM may thus cause inconvenience, because $\mathbf{v}$ can only be obtained by calculating the gradient of the p solution numerically. Moreover, the rate of convergence of solution $\mathbf{v}$ is reduced by 1 comparing with that of solution p. Therefore, the mixed FEM is superior in this aspect, not only because that both p and $\mathbf{v}$ are solved simultaneously but also that the error estimation gives same degree of accuracies for both p and $\mathbf{v}$.

One drawback for the mixed FEM is that it is trickier to implement than the standard conforming FEM. The mixed FEM uses non-isoparametric grid, this means that the trial function nodes are different from the geometry nodes, which is the case for both pressure-nodes and velocity-nodes. Besides, high-order trial functions are very complicated for $\mathbf{v}$, and the mixed FEM requires large amount of work on assembling of the stiffness matrix. Furthermore, when the corresponding linear system of equations is solved by a iterative method, one must take extra care of the preconditioning. The size of the solution vector in the mixed FEM is about twice as that in the corresponding conforming FEM, so the corresponding solving time is longer. Generally, when p is the only variable and $\mathbf{v}$ is not of interest, it is obviously unnecessary to ask for all the troubles to use the mixed FEM.

Considering all these factors, we can reach the following conclusions:

Errors for the conforming method, case 1:						
h(s)	$\|e_p^c\|_{L^2(\Omega)}$	$\|e_p^c\|_{l^2(\Omega)}$	$\|e_p^c\|_{l^\infty(\Omega)}$	$\|e_{\mathbf{v}}^c\|_{(L^2(\Omega))^2}$	$\|e_{\mathbf{v}}^c\|_{(l^2(\Omega))^2}$	$\|e_{\mathbf{v}}^c\|_{l^\infty(\Omega)}$
1/2	1.01255e-1	7.58799e-2	2.27640e-1	9.83242e-1	4.85297e-1	4.85297e-1
1/4	2.55247e-2	2.11847e-2	5.29617e-2	4.99654e-1	1.15518e-1	1.41480e-1
1/8	6.41312e-3	5.75546e-3	1.29498e-2	2.51298e-1	2.86236e-2	3.89697e-2
1/16	1.60558e-3	1.51480e-3	3.21895e-3	1.25847e-1	7.14159e-3	1.00032e-2
1/32	4.01545e-4	3.89613e-4	8.03577e-4	6.29486e-2	1.78453e-3	2.51763e-3
The rates of convergence for the conforming method:						
$1/2 \sim 1/4$	1.9880	1.8407	2.1037	0.9766	2.0708	1.7783
$1/4 \sim 1/8$	1.9928	1.8800	2.0320	0.9915	2.0128	1.8602
$1/8 \sim 1/16$	1.9979	1.9258	2.0083	0.9977	2.0029	1.9619
$1/16 \sim 1/32$	1.9995	1.9590	2.0021	0.9994	2.0007	1.9903
Errors for the mixed method:						
h(s)	$\|e_p^m\|_{L^2(\Omega)}$	$\|e_p^m\|_{l^2(\Omega)}$	$\|e_p^m\|_{l^\infty(\Omega)}$	$\|e_{\mathbf{v}}^m\|_{(L^2(\Omega))^2}$	$\|e_{\mathbf{v}}^m\|_{(l^2(\Omega))^2}$	$\|e_{\mathbf{v}}^m\|_{l^\infty(\Omega)}$
1/2	2.94614e-1	8.87665e-2	8.87665e-2	9.83419e-1	4.76725e-1	4.76725e-1
1/4	1.56928e-1	2.48895e-2	4.24891e-2	4.99654e-1	1.15383e-1	1.41315e-1
1/8	7.97315e-2	6.37567e-3	1.22660e-2	2.51298e-1	2.86215e-2	3.89668e-2
1/16	4.00261e-2	1.60328e-3	3.17575e-3	1.25847e-1	7.14155e-3	1.00031e-2
1/32	2.00331e-2	4.01402e-4	8.00870e-4	6.29486e-2	1.78453e-3	2.51763e-3
The rates of convergence for the mixed method:						
$1/2 \sim 1/4$	0.9087	1.8345	1.0629	0.9769	2.0467	1.7542
$1/4 \sim 1/8$	0.9769	1.9649	1.7924	0.9915	2.0113	1.8586
$1/8 \sim 1/16$	0.9942	1.9916	1.9495	0.9977	2.0028	1.9618
$1/16 \sim 1/32$	0.9986	1.9979	1.9875	0.9994	2.0007	1.9903
Errors between the two methods, case 1:						
h(s)	$\|e_p\|_{L^2(\Omega)}$	$\|e_p\|_{l^2(\Omega)}$	$\|e_p\|_{l^\infty(\Omega)}$	$\|e_{\mathbf{v}}\|_{(L^2(\Omega))^2}$	$\|e_{\mathbf{v}}\|_{(l^2(\Omega))^2}$	$\|e_{\mathbf{v}}\|_{l^\infty(\Omega)}$
1/2	2.90078e-1	1.04324e-1	1.04324e-1	1.42109e+0	8.57145e-3	8.57145e-3
1/4	1.56273e-1	2.57309e-2	4.39255e-2	7.12254e-1	1.35213e-4	1.65602e-4
1/8	7.96466e-2	6.42599e-3	1.23628e-2	3.56139e-1	2.10054e-6	2.85979e-6
1/16	4.00154e-2	1.60639e-3	3.18191e-3	1.78069e-1	3.31359e-8	5.24908e-8
1/32	2.00318e-2	4.01595e-4	8.01257e-4	8.90347e-2	7.50808e-10	1.76965e-9
The rates of convergence:						
$1/2 \sim 1/4$	0.8924	2.0195	1.2479	0.9965	5.9862	5.6937
$1/4 \sim 1/8$	0.9724	2.0015	1.8291	1.0000	6.0083	5.8557
$1/8 \sim 1/16$	0.9931	2.0001	1.9580	1.0000	5.9862	5.7677
$1/16 \sim 1/32$	0.9983	2.0000	1.9896	1.0000	5.4638	4.8905

Table 5.2: Comparing two methods, case 1.

Errors for the conforming method, case 2:						
h(s)	$\|e_p^c\|_{L^2(\Omega)}$	$\|e_p^c\|_{l^2(\Omega)}$	$\|e_p^c\|_{l^\infty(\Omega)}$	$\|e_{\mathbf{v}}^c\|_{(L^2(\Omega))^2}$	$\|e_{\mathbf{v}}^c\|_{(l^2(\Omega))^2}$	$\|e_{\mathbf{v}}^c\|_{l^\infty(\Omega)}$
1/2	9.99466e-2	1.10591e-1	2.54716e-1	1.50241e+0	7.44234e-1	8.91080e-1
1/4	2.51974e-2	2.66856e-2	6.09790e-2	7.69907e-1	1.79726e-1	2.76474e-1
1/8	6.33214e-3	6.59413e-3	1.50395e-2	3.87819e-1	4.46832e-2	8.00510e-2
1/16	1.58541e-3	1.64211e-3	3.74678e-3	1.94286e-1	1.11577e-2	2.11340e-2
1/32	3.96507e-4	4.09835e-4	9.35875e-4	9.71909e-2	2.78864e-3	5.39720e-3
The rates of convergence for the conforming method:						
1/2 ∼ 1/4	1.9879	2.0511	2.0625	0.9645	2.0500	1.6884
1/4 ∼ 1/8	1.9925	2.0168	2.0196	0.9893	2.0080	1.7882
1/8 ∼ 1/16	1.9978	2.0056	2.0050	0.9972	2.0017	1.9214
1/16 ∼ 1/32	1.9994	2.0024	2.0013	0.9993	2.0004	1.9693
Errors for the mixed method, case 2:						
h(s)	$\|e_p^m\|_{L^2(\Omega)}$	$\|e_p^m\|_{l^2(\Omega)}$	$\|e_p^m\|_{l^\infty(\Omega)}$	$\|e_{\mathbf{v}}^m\|_{(L^2(\Omega))^2}$	$\|e_{\mathbf{v}}^m\|_{(l^2(\Omega))^2}$	$\|e_{\mathbf{v}}^m\|_{l^\infty(\Omega)}$
1/2	2.95522e-1	8.55666e-2	1.03601e-1	1.55590e+0	8.33122e-1	1.12169e+0
1/4	1.57067e-1	2.39629e-2	5.04161e-2	7.74776e-1	1.99635e-1	3.28217e-1
1/8	7.97498e-2	6.13877e-3	1.49405e-2	3.88384e-1	4.94131e-2	9.42195e-2
1/16	4.00284e-2	1.54377e-3	3.93866e-3	1.94356e-1	1.23229e-2	2.49472e-2
1/32	2.00334e-2	3.86508e-4	1.00378e-3	9.71995e-2	3.07885e-3	6.39054e-3
Convergence rates for the mixed method:						
1/2 ∼ 1/4	0.9119	1.8362	1.0391	1.0059	2.0612	1.7730
1/4 ∼ 1/8	0.9778	1.9648	1.7547	0.9963	2.0144	1.8006
1/8 ∼ 1/16	0.9945	1.9915	1.9235	0.9988	2.0036	1.9171
1/16 ∼ 1/32	0.9986	1.9979	1.9723	0.9997	2.0009	1.9649
Errors between the two methods, case 2:						
h(s)	$\|e_p\|_{L^2(\Omega)}$	$\|e_p\|_{l^2(\Omega)}$	$\|e_p\|_{l^\infty(\Omega)}$	$\|e_{\mathbf{v}}\|_{(L^2(\Omega))^2}$	$\|e_{\mathbf{v}}\|_{(l^2(\Omega))^2}$	$\|e_{\mathbf{v}}\|_{l^\infty(\Omega)}$
1/2	2.93599e-1	1.10613e-1	1.29331e-1	2.19036e+0	2.57829e-1	2.73555e-1
1/4	1.56739e-1	2.76600e-2	5.53576e-2	1.09882e+0	5.42392e-2	8.85062e-2
1/8	7.97054e-2	6.93143e-3	1.59373e-2	5.49754e-1	1.29033e-2	2.39510e-2
1/16	4.00227e-2	1.73420e-3	4.16861e-3	2.74925e-1	3.18475e-3	6.20179e-3
1/32	2.00327e-2	4.33640e-4	1.05993e-3	1.37469e-1	7.93621e-4	1.57786e-3
The rates of convergence:						
1/2 ∼ 1/4	0.9055	1.9996	1.2242	0.9952	2.2490	1.6280
1/4 ∼ 1/8	0.9756	1.9966	1.7964	0.9991	2.0716	1.8857
1/8 ∼ 1/16	0.9939	1.9989	1.9348	0.9997	2.0185	1.9493
1/16 ∼ 1/32	0.9985	1.9997	1.9756	0.9999	2.0047	1.9747

Table 5.3: Comparing two methods, case 2.

- For those models where we are only interested in the p solution, use the conforming FEM;
- The mixed FEM is recommended when $\mathbf{v}$ also has important physical meanings.

The summary of the discussion is listed in Table 5.4.

The conforming FEM		The mixed FEM	
Easy to implement	★	Hard to implement	▼
Easy to solve the linear system	★	Hard to solve the linear system	▼
Accurate pressure solution	★	Accurate pressure solution	★
No direct access to $\mathbf{v}$	▼	Direct access to $\mathbf{v}$	★
Lower accuracy in $\mathbf{v}$	▼	Higher accuracy in $\mathbf{v}$	★
No local mass conservation	▼	Local mass conservation	★
Not satisfy the equilibrium equation	▼	Satisfy the equilibrium equation	★

★: Advantage
▼: Disadvantage

Table 5.4: Summary of the comparison of the conforming and the mixed methods.

Chapter 6

Study of Singularities

6.1 The Appearance of the Singularities

In this Chapter we will focus on the difficulties connected to the lack of the V-ellipticity and the continuity in the aluminium DC-casting model. Recall that in the DC-casting model, we have the permeability K which is a scalar-valued function of the volume fraction of liquid g_l, and is usually modelled by the Kozeny-Carman relation

$$K = \frac{g_l^3}{(1-g_l)^2},$$

where g_l is a scalar-valued function which physically varies from 0 on boundary Γ_1 to 1 on boundary Γ_4. However, there are two problems:

(1) When $g_l = 0$, then $K = 0$. If we choose a g_l that goes towards 0 on Γ_1 linearly, the permeability K will go towards 0 on Γ_1 by order of x_2^3 for a given x_1. Recall the conditions in Lax-Milgram Theorem, cf Section 3.1.2, we loose the V-ellipticity, i.e.,

$$a(u,u) \geq \alpha \|u\|_{H^1(\Omega)} \quad \forall u \in H^1(\Omega),$$

for some positive constant α which is independent of u. Therefore the solution might have singularity on the boundary Γ_1 subject to the boundary condition $\mathbf{v} \cdot \mathbf{n} = 0$, where $K \to 0$.

(2) Moreover, as $g_l = 1$, we have $K = \infty$, which means $\|K\|_{L^\infty(\Omega)} = \infty$. When we choose a g_l that goes towards infinite on Γ_4 linearly, the permeability K will go towards infinite on Γ_4 by order of $(1-x_2)^{-2}$ for a given x_1 (cf Figure 2.3 for a graphical interpretation). Hence the continuity condition on $a(.,.)$ in the Lax-Milgram Theorem is not satisfied, i.e., we do not have

$$a(u,v) \leq \gamma \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} \quad \forall u, v \in H^1(\Omega)$$

for some positive constant γ, since $\gamma = \|K/A\|_{L^\infty(\Omega)}$, cf Section 3.1.2. We also have the problem of presenting the infinite values of K on this boundary segment in the numerical method.

All in all, any of the two situations above can lead to that we fail the requirements in the Lax-Milgram Theorem, which means that we miss the well-posedness of the problem. This again indicates that we are no longer guaranteed with the error estimate for standard elliptic problems.

The reason that we include two small positive values ε_1 and ε_2 in our DC-casting problem is clear: we want to avoid the occurrences of the two unpleasant situations. However, we did employ a strategy on choosing the two small positive values in order to approximate the actual model. The idea is like the following: we start with $\varepsilon_1 = \varepsilon_2 = 0.1$ and gradually reduce them towards zero by 0.01, 0.001, $\cdots$. To our surprise, we still got reasonable numerical solutions even when we set both ε_1 and ε_2 to zero. Unfortunately, the study on the rate of convergence is difficult to carry out due to the fact that the analytical solution is not available.

Therefore, a "simplified" problem on the same solution domain with several sets of parameters is tested for the purpose of studying the rates of convergence in these problems. We define the "simplified" problem as:

- First, we choose a known function to be the analytical solution, where a sinusoid function can be a natural choice.[1]

- Second, we select a function for λ so that it is zero on some boundary segments where $\mathbf{v} \cdot \mathbf{n} = 0$. We can use a polynomial, or the $K(g_l)$ function in the DC-casting model.

- Finally, we fit the right-hand side function f according to the analytical solution and the λ function we have chosen, and we set the boundary conditions so that the "simplified" problem is complete.

Consequently, the study of the rates of convergence for this "simplified" problem is possible since we know the analytical solution. We got the following observation: for problems with "smooth" analytical solutions, i.e., functions in $H^2(\Omega)$ in our case, the rates of convergence agree with the standard error estimate for elliptic problems even if not all of the conditions in the Lax-Milgram Theorem are satisfied.

It should be emphasized that the conditions in the Lax-Milgram Theorem are sufficient but not necessary conditions for the standard error estimate. This means that the problems which agree with the standard rates of convergence do not have to satisfy all the conditions. (Our "simplified" problem is an example.) For complicated problems such as our DC-casting process, it is hard to predict in which space the solution is located, so the study becomes even more difficult. Since the complete study of such topics require huge amount of background knowledge, it is therefore not possible to give a very detailed discussion in a Cand. Scient. thesis. A very similar problem has been studied by Le Roux [13] in more details, where the author concludes that such problem has a unique solution in the weighted Sobolev space. We refer to Le Roux [13] for more detailed discussions.

In order to simplify the study of the singularities, we set up in the following Section a one-dimensional model problem, where the analytical solution can be easily obtained, and the corresponding space where the solution is located will also be known to us. We use the conforming FEM to solve some relevant cases with numerical experiments.

6.2 A One-dimensional Model problem

Consider the 1D model

$$-(\lambda u')' = f \quad \text{in } I \in [0,1], \tag{6.1}$$

with the boundary conditions

$$[\lambda u'](0) = 0, \quad u(1) = 1.$$

In analogy with the model of aluminium DC-casting, we consider the function λ satisfying the requirements $\lambda(0) = 0$ and $\lambda(x) > 0$ for $0 < x \leq 1$.

[1] Other analytical functions have also been tried, such as polynomials, exponential functions, etc.

6.2.1 FEM formulation

The weak formulation for the 1D problem is: Find $u \in H_g^1(I)$ such that

$$a(u,v) = L(v) \quad \forall v \in H_0^1(I),$$

where

$$\begin{aligned} a(u,v) &= \int_I v'(\lambda u')\,dx, \\ L(v) &= \int_I fv\,dx, \end{aligned}$$

and the coo responding linear subspace $H_0^1(I)$ is defined as

$$H_0^1(I) = \{u :\ u \in H^1(I),\ u(1) = 0\},$$

with the linear variety $H_g^1(I)$ defined as

$$H_g^1(I) = \{u :\ u \in H^1(I),\ u(1) = 1\}.$$

For the discrete problem, we define the finite-dimensional subspace and the variety as

$$V_h(I) \in H_0^1(I), \quad V_{h,g}(I) \in H_g^1(I),$$

so the problem is: Find $u_h \in V_{h,g}(I)$ such that

$$a(u_h, v) = L(v), \quad \forall v \in V_h(I). \tag{6.2}$$

The exact solution to (6.1) can be obtained simply by taking integration as

$$\lambda u' = -\int_0^x f(y)\,dy,$$

which brings

$$u' = \frac{-\int_0^x f(y)\,dy}{\lambda(x)} = G(x).$$

and

$$u = \int_0^x G(z)\,dz + C,$$

where C is a constant determined by the boundary condition $u(1) = 1$. Note that the space that the solution u is located in is dependent on the properties of the functions f and λ. In this problem, the V-ellipticity condition of the Lax-Milgram Theorem is not satisfied, (cf Section 3.1.2,) hence we do not have a well-posed variational problem. To see how the numerical solutions perform in the different situations, we will look at several numerical experiments in the next sub Section.

6.2.2 Numerical experiments

Before we go directly to the numerical experiments,[2] we describe first the numerical integration scheme that solves one of the difficulties in the numerical method of the DC-casting model.

[2]The numerical programming of the 1D model is very simple, so it is implemented in ***Matlab***. The code is not included.

Numerical integration Recall that in the DC-casting model, we have $\lim_{g_l \to 1} K(g_l) = \infty$ on the boundary Γ_4. This brings the problem of presenting the infinite values in the numerical method. Fortunately, using numerical integration with Gaussian quadrature scheme to compute the stiffness matrix and the righthand side vector helps us to avoid this difficulty. In the Gaussian quadrature scheme, cf Figure 6.1, the evaluating points are actually located inside the element for both 1D elements and 2D box elements. (For more details about numerical integration, we refer to Johnson [12] and Zienkiewicz [28].) So we never need to evaluate the value of K on the boundary. Since the K is only infinite on the boundary but finite inside the solutions domain, we avoid the difficulty in presenting the infinite values. However, when the grid is very fine, we will meet very large values of K, which may be so large that it exceeds the capacity of the data machine. Note that even though we avoid presenting infinite values in the numerical method, we still miss the continuity of $a(.,.)$, thus the variational problem is not well-posed according to the Lax-Milgram Theorem.

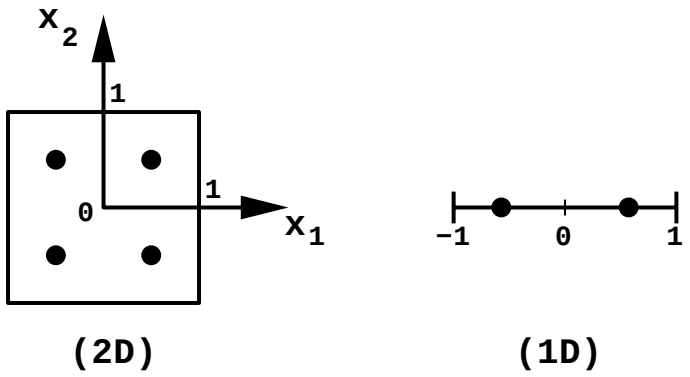


Figure 6.1: Numerical integration in 2D and 1D: The Gaussian quadrature scheme.

Case 6.1: We study the problem in (6.1) with $f \equiv 1$ and $\lambda = x^p$, i.e.,

$$-(x^p\, u')' = 1 \quad \text{in } I \in [0,1], \quad u(1) = 1,\ [x^p\, u'](0) = 0.$$

The exact solution can be obtained by integration as

$$u'(x) = -x^{1-p},$$

and thus

$$u = \begin{cases} 1 + \dfrac{1 - x^{2-p}}{2-p}, & p \neq 2, \\ 1 - \ln(x), & p = 2. \end{cases}$$

It is easy to check which space u is located in. Notice that $u''(x) = -(1-p)x^{-p}$, so we have

$$\int_0^1 (u'')^2\, dx = (1-p)^2 \int_0^1 x^{-2p}\, dx = \frac{(1-p)^2}{1-2p}\,[x^{1-2p}]_0^1,$$

while

$$\int_0^1 \left(u'(x)\right)^2\, dx = \int_0^1 x^{2-2p}\, dx = \frac{1}{3-2p}\,[x^{3-2p}]_0^1.$$

For the solution u, when $p \neq 2$, we have

$$\int_0^1 (u(x) - C)^2 \, dx = \int_0^1 \frac{x^{4-2p}}{(2-p)^2} \, dx = \frac{1}{(2-p)^2(5-2p)} \left[x^{5-2p}\right]_0^1,$$

where $C = \frac{3-p}{2-p}$. When $p = 2$, we have

$$\int_0^1 (u-1)^2 \, dx = \int_0^2 \ln^2(x) \, dx = \Big[2x - 2x\ln(x) + x\ln^2(x)\Big]_0^1 = 2.$$

Next are some remarks about the u solution:

- ♡ For $p < 0$, $u'' \in L^2(I)$, we have $u \in H^2(I)$. Since $\lim_{x\to 0} \lambda(x) = \infty$, we have the difficulty in presenting the infinite value in the numerical method. However, this difficulty is avoided by using numerical integration with Gaussian quadrature scheme.
- ♡ For $0 \leq p < 0.5$, $u'' \in L^2(I)$, we have $u \in H^2(I)$.
- ♡ For $0.5 \leq p \leq 1$, $u'' \notin L^2(I)$, $u' \in L^2(I)$, we have $u \in H^1(I)$, where u' is bounded in I.
- ♡ For $1 < p < 1.5$, we still have $u \in H^1(I)$, but $\lim_{x\to 0} u'(x) = \infty$, i.e., the solution has singularity at $x = 0$.
- ♡ For $2.5 > p \geq 1.5$, we do not have $u' \in L^2(I)$, but $u \in L^2(I)$.
- ♡ For $p \geq 2.5$, we have $u \notin L^2(I)$.

The errors and the rates of convergence are listed in Tables 6.1 and 6.2, where $\|e\|_{H^1(I)}$ is the error in H^1-norm and $a \sim \|.\|_{H^1}$ denotes the rates of convergence in H^1-norm, etc. We see clearly from these tables that the standard error estimate no longer holds when $p > 0.5$. With $0.5 \leq p < 1.5$, the solutions converge, but with much lower rates of convergence. For $p > 1.5$, the solutions do not converge in H^1-norm. (The negative values of a can not be called the rates of convergence, because they actually indicate divergence.) Observe that when $p = 1.5$, the rates of convergence in H^1-norm are zero, and when $p = 2.5$, the rates of convergence in L^2-norm are zero. Actually, $p = 1.5$ and $p = 2.5$ are the turning points where the solutions move out of space $H^1(I)$ and $L^2(\Omega)$, and it is not surprising that something special happens here. A graph of the rates of convergence as functions of p is included in Figure 6.2, with the summary on rate of convergence being listed in Table 6.3.

When the solution u is "smooth" enough, i.e., $u \in H^2(I)$ here, we still have the standard rates of convergence even if not all of the conditions of the Lax-Milgram theorem are satisfied. This accords with our observation in the "simplified" problem. In fact, this result can be proven analytically by considering an auxiliary problem $((\lambda + \varepsilon)v')' = f$ where $\varepsilon > 0$, and we have $u = \lim_{\varepsilon\to 0} v$. Note that the auxiliary problem is strictly elliptic and well-posed. It can be shown that v goes towards u continuously as $\varepsilon \to 0$, thus the problem for u is also well-posed with the standard error estimate for the numerical solution. We omit the the details for the proof and consider this observation as a general one.

h	$\|e\|_{H^1(I)}$	$a \sim \|.\|_{H^1}$	$\|e\|_{H^1(I)}$	$a \sim \|.\|_{H^1}$	$\|e\|_{L^2(I)}$	$a \sim \|.\|_{L^2}$
$p = 0.4$:						
5.000e-01	1.1552e-01	$\sim$	1.1476e-01	$\sim$	1.3251e-02	$\sim$
2.500e-01	6.1517e-02	0.9091	6.1419e-02	0.9018	3.4791e-03	1.9293
1.250e-01	3.2389e-02	0.9255	3.2377e-02	0.9237	8.9718e-04	1.9552
6.250e-02	1.6884e-02	0.9399	1.6882e-02	0.9395	2.2882e-04	1.9712
3.125e-02	8.7323e-03	0.9512	8.7321e-03	0.9511	5.7957e-05	1.9811
1.562e-02	4.4889e-03	0.9600	4.4889e-03	0.9600	1.4616e-05	1.9875
7.812e-03	2.2966e-03	0.9669	2.2966e-03	0.9669	3.6753e-06	1.9916
3.906e-03	1.1705e-03	0.9724	1.1705e-03	0.9723	9.2252e-07	1.9942
1.953e-03	5.9476e-04	0.9768	5.9476e-04	0.9768	2.3127e-07	1.9960
9.766e-04	3.0146e-04	0.9804	3.0146e-04	0.9804	5.7930e-08	1.9972
$p = 0.5$:						
5.000e-01	1.0958e-01	$\sim$	1.0904e-01	$\sim$	1.0807e-02	$\sim$
2.500e-01	6.0446e-02	0.8582	6.0376e-02	0.8528	2.8899e-03	1.9029
1.250e-01	3.2876e-02	0.8786	3.2868e-02	0.8773	7.5555e-04	1.9354
6.250e-02	1.7675e-02	0.8954	1.7674e-02	0.8951	1.9478e-04	1.9557
3.125e-02	9.4164e-03	0.9085	9.4163e-03	0.9084	4.9764e-05	1.9686
1.562e-02	4.9810e-03	0.9187	4.9810e-03	0.9187	1.2640e-05	1.9771
7.812e-03	2.6199e-03	0.9270	2.6199e-03	0.9270	3.1979e-06	1.9828
3.906e-03	1.3716e-03	0.9337	1.3716e-03	0.9337	8.0690e-07	1.9867
1.953e-03	7.1527e-04	0.9393	7.1527e-04	0.9393	2.0322e-07	1.9893
9.766e-04	3.7179e-04	0.9440	3.7179e-04	0.9440	5.1116e-08	1.9912
4.883e-04	1.9272e-04	0.9480	1.9272e-04	0.9480	1.2837e-08	1.9935
2.441e-04	9.9652e-05	0.9515	9.9652e-05	0.9515	3.1778e-09	2.0142
$p = 0.6$:						
5.000e-01	1.0068e-01	$\sim$	1.0032e-01	$\sim$	8.4969e-03	$\sim$
2.500e-01	5.7936e-02	0.7972	5.7889e-02	0.7932	2.3286e-03	1.8675
1.250e-01	3.2844e-02	0.8189	3.2838e-02	0.8179	6.2170e-04	1.9052
6.250e-02	1.8404e-02	0.8356	1.8404e-02	0.8354	1.6336e-04	1.9282
3.125e-02	1.0224e-02	0.8481	1.0224e-02	0.8481	4.2505e-05	1.9423
1.562e-02	5.6420e-03	0.8577	5.6419e-03	0.8576	1.0993e-05	1.9510
7.812e-03	3.0976e-03	0.8651	3.0976e-03	0.8651	2.8331e-06	1.9562
3.906e-03	1.6938e-03	0.8709	1.6938e-03	0.8709	7.2872e-07	1.9589
$p = 0.9$:						
5.000e-01	3.9718e-02	$\sim$	3.9652e-02	$\sim$	2.2912e-03	$\sim$
2.500e-01	2.6844e-02	0.5652	2.6833e-02	0.5634	7.4503e-04	1.6207
1.250e-01	1.7958e-02	0.5800	1.7956e-02	0.5795	2.3818e-04	1.6452
6.250e-02	1.1942e-02	0.5886	1.1942e-02	0.5885	7.6020e-05	1.6476
3.125e-02	7.9144e-03	0.5935	7.9143e-03	0.5935	2.4381e-05	1.6406
1.562e-02	5.2350e-03	0.5963	5.2350e-03	0.5963	7.8719e-06	1.6310
7.812e-03	3.4589e-03	0.5979	3.4589e-03	0.5979	2.5573e-06	1.6221
3.906e-03	2.2840e-03	0.5988	2.2840e-03	0.5988	8.3486e-07	1.6150

Table 6.1: The errors and the rates of convergence for case 6.1 (part 1).

h	$\|e\|_{H^1(I)}$	$a \sim \|.\|_{H^1}$	$\|e\|_{H^1(I)}$	$a \sim \|.\|_{H^1}$	$\|e\|_{L^2(I)}$	$a \sim \|.\|_{L^2}$
$p = 1.4$:						
1.250e-01	3.1742e-01	$\sim$	3.1739e-01	$\sim$	4.8847e-03	$\sim$
6.250e-02	2.9623e-01	0.0997	2.9622e-01	0.0996	2.3006e-03	1.0863
3.125e-02	2.7641e-01	0.0999	2.7641e-01	0.0999	1.0773e-03	1.0947
1.562e-02	2.5791e-01	0.1000	2.5791e-01	0.1000	5.0328e-04	1.0979
7.812e-03	2.4064e-01	0.1000	2.4064e-01	0.1000	2.3492e-04	1.0992
3.906e-03	2.2452e-01	0.1000	2.2452e-01	0.1000	1.0962e-04	1.0997
$p = 1.5$:						
1.250e-01	5.3882e-01	$\sim$	5.3876e-01	$\sim$	8.5095e-03	$\sim$
6.250e-02	5.3886e-01	-0.0001	5.3884e-01	-0.0002	4.2902e-03	0.9880
3.125e-02	5.3887e-01	0.0000	5.3886e-01	-0.0001	2.1516e-03	0.9956
1.562e-02	5.3887e-01	0.0000	5.3887e-01	0.0000	1.0770e-03	0.9984
7.812e-03	5.3887e-01	0.0000	5.3887e-01	0.0000	5.3869e-04	0.9994
3.906e-03	5.3887e-01	0.0000	5.3887e-01	0.0000	2.6938e-04	0.9998
$p = 1.6$:						
1.250e-01	8.7892e-01	$\sim$	8.7881e-01	$\sim$	1.4146e-02	$\sim$
6.250e-02	9.4199e-01	-0.1000	9.4196e-01	-0.1001	7.6330e-03	0.8901
3.125e-02	1.0096e+00	-0.1000	1.0096e+00	-0.1000	4.1000e-03	0.8966
1.562e-02	1.0820e+00	-0.1000	1.0820e+00	-0.1000	2.1988e-03	0.8989
7.812e-03	1.1597e+00	-0.1000	1.1597e+00	-0.1000	1.1786e-03	0.8996
3.906e-03	1.2429e+00	-0.1000	1.2429e+00	-0.1000	6.3167e-04	0.8999
$p = 2.4$:						
1.250e-01	2.4250e+01	$\sim$	2.4246e+01	$\sim$	3.8653e-01	$\sim$
6.250e-02	4.5247e+01	-0.8999	4.5245e+01	-0.9000	3.6097e-01	0.0987
3.125e-02	8.4432e+01	-0.9000	8.4431e+01	-0.9000	3.3685e-01	0.0998
1.562e-02	1.5755e+02	-0.9000	1.5755e+02	-0.9000	3.1431e-01	0.1000
7.812e-03	2.9401e+02	-0.9000	2.9401e+02	-0.9000	2.9326e-01	0.1000
3.906e-03	5.4863e+02	-0.9000	5.4863e+02	-0.9000	2.7362e-01	0.1000
$p = 2.5$:						
1.250e-01	3.5428e+01	$\sim$	3.5424e+01	$\sim$	5.5734e-01	$\sim$
6.250e-02	7.0850e+01	-0.9999	7.0848e+01	-1.0000	5.5772e-01	-0.0010
3.125e-02	1.4170e+02	-1.0000	1.4170e+02	-1.0000	5.5779e-01	-0.0002
1.562e-02	2.8339e+02	-1.0000	2.8339e+02	-1.0000	5.5780e-01	0.0000
7.812e-03	5.6678e+02	-1.0000	5.6678e+02	-1.0000	5.5780e-01	0.0000
3.906e-03	1.1336e+03	-1.0000	1.1336e+03	-1.0000	5.5780e-01	0.0000
$p = 2.6$:						
1.250e-01	5.1550e+01	$\sim$	5.1544e+01	$\sim$	7.9914e-01	$\sim$
6.250e-02	1.1049e+02	-1.0999	1.1049e+02	-1.1000	8.5694e-01	-0.1007
3.125e-02	2.3683e+02	-1.1000	2.3683e+02	-1.1000	9.1853e-01	-0.1001
1.562e-02	5.0766e+02	-1.1000	5.0766e+02	-1.1000	9.8447e-01	-0.1000
7.812e-03	1.0882e+03	-1.1000	1.0882e+03	-1.1000	1.0551e+00	-0.1000
3.906e-03	2.3326e+03	-1.1000	2.3326e+03	-1.1000	1.1309e+00	-0.1000

Table 6.2: The errors and the rates of convergence for case 6.1 (part 2).

p value	H^1-norm	L^2-norm	u
$0 \le p \le 0.5$	$a = 1$	$a = 2$	$u \in H^2(I)$
$0.5 < p < 1.5$	$0 < a < 1$	$1 < a < 2$	$u \in H^1(I)$
$p = 1.5$	$a = 0$	$a = 1$	$u \in L^2(I)$
$1.5 < p < 2.5$	$-1 < a < 0$	$0 < a < 1$	$u \in L^2(I)$
$p = 2.5$	$a = -1$	$a = 0$	$u \notin L^2(I)$
$p > 2.5$	$a < -1$	$a < 0$	$u \notin L^2(I)$

Table 6.3: Summary of the rates of convergence a in Case 6.1.

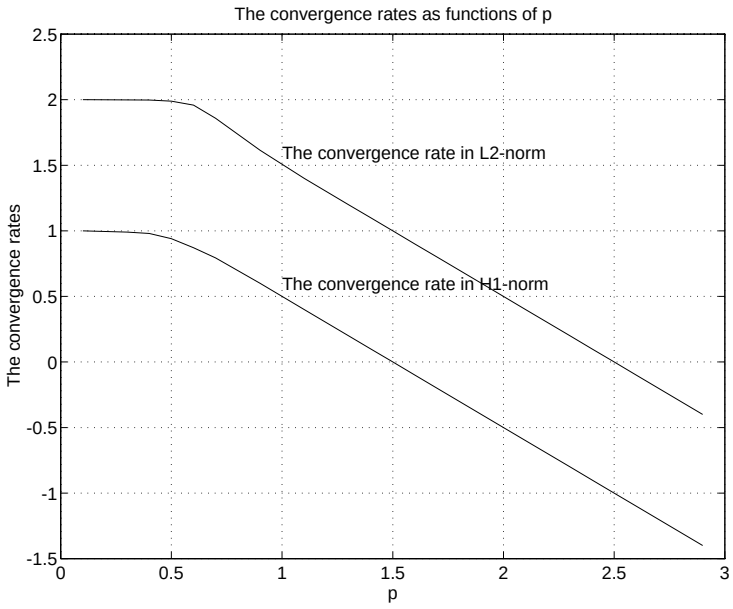


Figure 6.2: The plot of the rates of convergence as functions of p in Case 6.1.

Case 6.2: Here we treat the problem (6.1) with $f \equiv 1$ and

$$\lambda(x) = \frac{(x + e_1)^3}{(1 + e_2 - x)^2}.$$

Notice that this function resembles the function $K(g_l)$ in the aluminium DC-casting problem when g_l is chosen to be linear function in x_2 for a given x_1. (In this case g_l resembles x, and e_1, e_2 resemble ε_1 and ε_2, respectively.) The plot of the function $\lambda(x)$ for some choices of (e_1, e_2) is included in Figure 6.3.

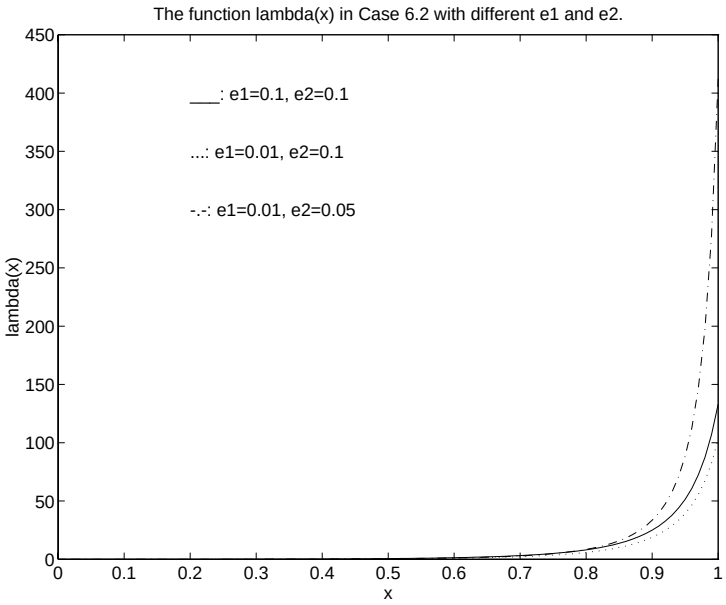


Figure 6.3: The function $\lambda(x)$ in Case 6.2.

It can be shown that the exact solution is

$$u'(x) = -\frac{x(1 + e_2 - x)^2}{(x + e_1)^3},$$

and

$$\begin{aligned} u(x) \quad = \quad & -x - \frac{e_1(1 + e_1 + e_2)^2}{2(e_1 + x)^2} + \frac{1 + 4e_1 + 3e_1^2 + 2e_2 + 4e_1e_2 + e_2^2}{e_1 + x} \\ & + (2 + 3e_1 + 2e_2)\ln(e_1 + x) + C, \end{aligned}$$

where C is a constant to be determined by the boundary condition $u(1) = 1$. The plot of the exact solutions for different choices of e_1 and e_2 is given in Figure 6.4. Note that with $e_2 >> 0$ and $e_1 = 0$, this problem is similar to the problem in Case 6.1 with $p = 3$.

We tested this problem with different choices of e_1 and e_2, the errors and the rates of convergence are listed in Table 6.4. The results show that, with $e_2 = 0$ and e_1 not too small, the rates of convergence are not affected. But the value of e_1 can not be too small if an accurate numerical solution is required. We also notice that the numerical solution does not converge with $e_1 = 0$. This is caused by the singularity in the solution. And the reader must be notified that these results are only valid for this particular problem.

h	$\|e\|_{H^1(I)}$	$a \sim \|.\|_{H^1}$	$\|e\|_{H^1(I)}$	$a \sim \|.\|_{H^1}$	$\|e\|_{L^2(I)}$	$a \sim \|.\|_{L^2}$
$e_1 = e_2 = 0.1$:						
6.250e-02	7.5624e-01	$\sim$	7.5581e-01	$\sim$	2.5355e-02	$\sim$
3.125e-02	6.7192e-01	0.1705	6.7187e-01	0.1699	8.7133e-03	1.5410
1.562e-02	3.9922e-01	0.7511	3.9921e-01	0.7510	2.4690e-03	1.8193
7.812e-03	2.0916e-01	0.9326	2.0916e-01	0.9326	6.4043e-04	1.9468
3.906e-03	1.0584e-01	0.9827	1.0584e-01	0.9827	1.6166e-04	1.9861
1.953e-03	5.3079e-02	0.9956	5.3079e-02	0.9956	4.0515e-05	1.9965
9.766e-04	2.6560e-02	0.9989	2.6560e-02	0.9989	1.0135e-05	1.9991
$e_1 = e_2 = 0.01$:						
3.125e-02	9.2238e+01	$\sim$	9.2236e+01	$\sim$	5.9711e-01	$\sim$
1.562e-02	3.2729e+01	1.4948	3.2729e+01	1.4948	2.1969e-01	1.4425
7.812e-03	2.0382e+01	0.6833	2.0382e+01	0.6833	8.6015e-02	1.3528
3.906e-03	2.1165e+01	-0.0544	2.1165e+01	-0.0544	3.2029e-02	1.4252
1.953e-03	1.3325e+01	0.6675	1.3325e+01	0.6675	9.5513e-03	1.7456
9.766e-04	7.1064e+00	0.9070	7.1064e+00	0.9070	2.5227e-03	1.9207
4.883e-04	3.6132e+00	0.976	3.6132e+00	0.9758	6.4003e-04	1.9788
2.441e-04	1.8143e+00	0.994	1.8143e+00	0.9939	1.6061e-04	1.9946
$e_1 = e_2 = 0.001$:						
6.250e-02	5.0679e+02	$\sim$	5.0678e+02	$\sim$	3.8244e+00	$\sim$
3.125e-02	1.1928e+03	-1.2349	1.1928e+03	-1.2349	4.6729e+00	-0.2891
1.562e-02	2.3559e+03	-0.9820	2.3559e+03	-0.9820	4.9404e+00	-0.0803
7.812e-03	3.4850e+03	-0.5648	3.4850e+03	-0.5648	4.1068e+00	0.2666
3.906e-03	3.2912e+03	0.0825	3.2912e+03	0.0825	2.3848e+00	0.7841
1.953e-03	1.5546e+03	1.0820	1.5546e+03	1.0820	9.4114e-01	1.3414
9.766e-04	5.4745e+02	1.5058	5.4745e+02	1.5058	3.5298e-01	1.4148
4.883e-04	7.1015e+02	-0.3754	7.1015e+02	-0.3754	1.3954e-01	1.3389
2.441e-04	4.9632e+02	0.5169	4.9632e+02	0.5169	4.4413e-02	1.6517
$e_1 = 0.1$, $e_2 = 0$:						
6.250e-02	6.1722e-01	$\sim$	6.1686e-01	$\sim$	2.1104e-02	$\sim$
3.125e-02	5.5222e-01	0.1605	5.5218e-01	0.1598	7.2459e-03	1.5422
1.562e-02	3.2878e-01	0.7481	3.2877e-01	0.7481	2.0532e-03	1.8193
7.812e-03	1.7234e-01	0.9318	1.7234e-01	0.9318	5.3260e-04	1.9468
3.906e-03	8.7223e-02	0.9825	8.7223e-02	0.9825	1.3444e-04	1.9860
1.953e-03	4.3745e-02	0.9956	4.3745e-02	0.9956	3.3694e-05	1.9965
$e_1 = 0$, $e_2 = 0.1$:						
5.000e-01	2.8456e+01	$\sim$	2.8411e+01	$\sim$	1.6105e+00	$\sim$
2.500e-01	8.8208e+01	-1.6322	8.8172e+01	-1.6339	2.5483e+00	-0.6621
1.250e-01	2.6032e+02	-1.5613	2.6030e+02	-1.5618	3.7686e+00	-0.5645
6.250e-02	7.5159e+02	-1.5296	7.5157e+02	-1.5297	5.4382e+00	-0.5291
$e_1 = 0$, $e_2 = 0$:						
5.000e-01	2.3055e+01	$\sim$	2.3018e+01	$\sim$	1.2967e+00	$\sim$
2.500e-01	7.2258e+01	-1.648	7.2228e+01	-1.6498	2.0858e+00	-0.6857
1.250e-01	2.1425e+02	-1.568	2.1422e+02	-1.5685	3.1016e+00	-0.5724
6.250e-02	6.1988e+02	-1.533	6.1986e+02	-1.5328	4.4856e+00	-0.5323

Table 6.4: The errors and the rates of convergence for case 6.2.

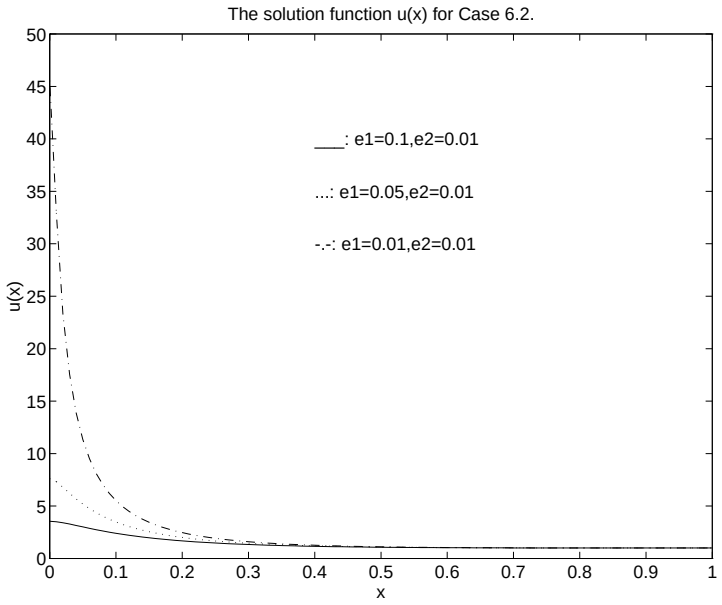


Figure 6.4: The exact solution $u(x)$ of the Case 6.2.

6.3 A Two-dimensional Model: The Aluminium DC-casting Problem

The difficulty in the estimation of the rates of convergence in many complicated problems lies in the lack of an analytical solution. The traditional way is to compute a solution on a "very" fine grid and treat it as the exact one. The disadvantage of this method is the extremely long computation time. We will here introduce a new method which effectively estimates the rates of convergence for those problems whose exact solutions can not be obtained by analytical studies. This new method is applied to our aluminium DC-casting model problem with the results being presented. The numerical method for the two-dimensional model is implemented in `DIFFPACK`.

6.3.1 A method for estimating the rate of convergence

Let u denote the analytical solution of a problem which can not be obtained analytically, and let u_h be the numerical solution from a certain numerical method, where h is the mesh grid size. We hereby estimate the rate of convergence in a certain norm $\|\cdot\|$ for an ideal problem, i.e., both the mathematical and numerical problem are well-posed, and the numerical solution converges towards the exact solution as $h \to 0$ in $\|\cdot\|$. In addition, we denote the error $e = e(h) = u - u_h$, which is a function of the mesh size h.

Assume that the rate of convergence for $\|e(h)\|$ is a, then we have the following relation

$$\|e(h)\| = C\,h^a, \tag{6.3}$$

for some constant C independent of h. Since u is not available, $e(h)$ is actually unknown to us. We want to find a way of estimating the value of a with the information that can be

obtained after some simple calculation. Hereby we define an assistant function $E(h_1, h_2)$ as

$$E(h_1, h_2) = u_{h_1} - u_{h_2} = -(u - u_{h_1}) + (u - u_{h_2}) = -e(h_1) + e(h_2).$$

By choosing $h_1 = 2h_2 = 2h$, we define $\tilde{e}(h) = E(2h, h)$, then

$$\tilde{e}(h) = -e(2h) + e(h).$$

Furthermore, after assuming $\|e(2h)\| \geq \|e(h)\|$ which is quite natural for ideal problems, we have

$$\|e(2h)\| - \|e(h)\| \leq \|\tilde{e}(h)\| \leq \|e(2h)\| + \|e(h)\|.$$

By equation (6.3), we can reach the following important inequality relation:

$$C(2h)^a - Ch^2 \leq \|\tilde{e}(h)\| \leq C(2h)^a + Ch^a. \tag{6.4}$$

When $h \to 0$, both the left and right side of the inequality go towards 0, then we have $\|\tilde{e}(h)\| \to 0$, which indicates that $\|\tilde{e}(h)\|$ converges. In order to discover that at what rate $\|\tilde{e}(h)\|$ converges, we assume that

$$\|\tilde{e}(h)\| = \tilde{C}h^{\tilde{a}},$$

and set it into (6.4). After dividing then with h^a, we will have

$$C(2^a - 1) \leq \tilde{C}h^{\tilde{a}-a} \leq C(2^a + 1).$$

It is not difficult to see that we need $\tilde{a} - a = 0$, i.e., $\tilde{a} = a$ for the above inequality to hold as $h \to 0$. ($\tilde{a} - a = 0 \quad \Rightarrow \quad h^{\tilde{a}-a} = 1.$) This means that we can obtain the rate of convergence of $\|e(h)\|$ by simply estimating that of $\|\tilde{e}(h)\|$.

The implementation of estimating the rate of convergence of $\|\tilde{e}(h)\|$ should not be difficult, and the computation time is also tolerable. We apply this method to several simple problems with known exact solutions, and this method produce the correct rates of convergence as expected. In the next sub Section, we present a numerical test of the method on the Poisson equation.

6.3.2 Testing of the method with the Poisson equation

Consider the Poisson equation

$$-\Delta p = 2\pi^2 \sin(\pi x_1)\sin(\pi x_2) \qquad \text{in } \Omega,$$

where $\Omega = [0, 1] \times [0, 1]$. Applying homogeneous Dirichlet boundary conditions, i.e., $p = 0$, on the whole boundary will produce the analytical solution as

$$p(x_1, x_2) = \sin(\pi x_1)\sin(\pi x_2).$$

This problem will have the standard rates of convergence, i.e., when using piecewise linear trial functions for p, the rate of convergence in H^1- and L^2-norm will be 1 and 2 respectively, cf (3.14). We solve this problem with the conforming finite element method, and estimate the rates of convergence with the method introduced in the preceding sub Section. The data are listed in Table 6.5. And we see that the method gives the correct rates of convergence.

partitions	$\|\tilde{e}\|_{H^1(\Omega)}$	$\tilde{a} \sim \|\tilde{e}\|_{H^1(\Omega)}$	$\|\tilde{e}\|_{L^2(\Omega)}$	$\tilde{a} \sim \|\tilde{e}\|_{L^2(\Omega)}$
(1x1)∼(2x2)	2.04607e+00	∼	4.09213e-01	∼
(2x2)∼(4x4)	8.67274e-01	1.2383	9.86359e-02	2.0527
(4x4)∼(8x8)	4.34440e-01	0.9973	2.47781e-02	1.9930
(8x8)∼(16x16)	2.17839e-01	0.9959	6.21381e-03	1.9955
(16x16)∼(32x32)	1.09012e-01	0.9988	1.55488e-03	1.9987

Table 6.5: The rates of convergence of the Poisson equation estimated by the new method.

6.3.3 The rates of convergence of the aluminium DC-casting problem

Recall the aluminium DC-casting problem (cf. Chapter 2)

$$\nabla \cdot \left(\frac{K}{A} (\nabla p + E) \right) = 0,$$

subject to the boundary conditions

$$\begin{aligned} p &= g_p \qquad \text{on } \partial\Omega_a, \\ \mathbf{v} \cdot \mathbf{n} &= 0 \qquad \text{on } \partial\Omega_b, \end{aligned}$$

where $K = 0$ on $\Gamma_1 \subset \partial\Omega_b$.

Generally, two-dimensional elliptic boundary value problems with homogeneous righthand side functions will have non-trial solutions with non-homogeneous boundary conditions. And the fact with our DC-casting model is that the exact solution is by no means available due to the irregular geometry. However, after having studied some one-dimensional problems, we can intuitively expect our 2D problem, which has a homogeneous right side function and the mobility $\lambda = 0$ only on some boundary segments where $\mathbf{v} \cdot \mathbf{n} = 0$, to have solution in the space $H^1(\Omega)$ and avoid singularity in this situation. Nevertheless, we can not tell whether the solution is in $H^2(\Omega)$, and neither can we predict the exact rates of convergence. Our prediction is that the numerical solution will converge in H^1- and L^2-norm and the rate of convergence is weaker than that of standard elliptic problems.

When we solve the DC-casting problem with conforming FEM, we use band structured matrix for small systems, and the linear system is solved by Gauss Elimination without pivoting provided that the band of the stiffness matrix is reduced using the Puttonen's method. For large systems, we use an iterative solver with the following strategy:

- We use a sparse structured matrix;
- The iterative method is the conjugate gradient method;
- The preconditioning type is RILU;
- The RILU relaxation parameter is 0.9;
- The convergence stop criterion is the maximum absolute value of the residual, i.e., the discrete l^∞-norm;
- The maximum error in the convergence test, i.e., the tolerance is 10^{-9}.

The rates of convergence of our aluminium DC-casting problem for different choices of ε_1 and ε_2 are obtained by using the method introduced in sub Section 6.3.1. The results are listed in Table 6.6. As described in the discussion on the implementation for the conforming finite element method, the solution domain is divided into 9 super elements. The numbers in the column "partition" in Table 6.6 indicate further partitions in each super element. The details of the implementation are not described in this thesis.

The results in Table 6.6 show that the solution of this problem converges even when we let $\varepsilon_1 = \varepsilon_2 = 0$, but at lower rate than the standard elliptic problems. This can be caused by either the distorted elements, or the fact that the solution is not in space $H^2(\Omega)$. It verifies our prediction. Furthermore, the rates of convergence vary very little for different choices of ε_1 and ε_2. Comparing the graphic results of $\varepsilon_1 = \varepsilon_2 = 0$ with those of $\varepsilon_1 = \varepsilon_2 = 0.1$, one can actually "see" no difference, so they are not included in this thesis.

Now we can conclude that it is safe to set $\varepsilon_1 = \varepsilon_2 = 0$ in our problem, provided we choose the volume fraction of liquid g_l to be a linear function in x_2 for a given x_1. In other words, letting the K function be 0 on one part of the boundary or go towards infinity on another part of the boundary will not cause loss of much accuracy in the numerical solutions. However, the reader should be aware that the result is dependent on the performance of g_l, i.e., how fast K goes towards zero or infinity on boundary Γ_1 or Γ_4. It means that the numerical results could be different for a different choice of g_l. Nevertheless, the function g_l used in our model is only an approximation to the real function which is actually an unknown in the complete model problem, this means that we may not be able to judge our approximation before the whole system is solved.

partitions	$\|\tilde{e}\|_{H^1(\Omega)}$	$\tilde{a} \sim \|\tilde{e}\|_{H^1(\Omega)}$	$\|\tilde{e}\|_{L^2(\Omega)}$	$\tilde{a} \sim \|\tilde{e}\|_{L^2(\Omega)}$
$\varepsilon_1 = \varepsilon_2 = 0.1$:				
(1x1)∼(2x2)	4.8382e-01	∼	6.4126e-03	∼
(2x2)∼(4x4)	2.7462e-01	0.8170	1.5895e-03	2.0123
(4x4)∼(8x8)	1.5315e-01	0.8425	5.7771e-04	1.4602
(8x8)∼(16x16)	8.8667e-02	0.7885	2.0493e-04	1.4952
(16x16)∼(32x32)	4.9033e-02	0.8546	7.6539e-05	1.4209
$\varepsilon_1 = \varepsilon_2 = 0.01$:				
(1x1)∼(2x2)	5.0392e-01	∼	6.6579e-03	∼
(2x2)∼(4x4)	3.0718e-01	0.7141	2.1239e-03	1.6484
(4x4)∼(8x8)	1.8191e-01	0.7559	9.6634e-04	1.1361
(8x8)∼(16x16)	1.1161e-01	0.7048	4.1728e-04	1.2115
(16x16)∼(32x32)	6.5716e-02	0.7641	1.6619e-04	1.3282
$\varepsilon_1 = \varepsilon_2 = 0.001$:				
(1x1)∼(2x2)	5.0492e-01	∼	6.6751e-03	∼
(2x2)∼(4x4)	3.0341e-01	0.7348	2.1945e-03	1.6049
(4x4)∼(8x8)	1.8712e-01	0.6973	1.0376e-03	1.0806
(8x8)∼(16x16)	1.0985e-01	0.7684	4.7991e-04	1.1124
(16x16)∼(32x32)	6.6759e-02	0.7185	2.1378e-04	1.1666
$\varepsilon_1 = 0; \varepsilon_2 = 0.01$:				
(1x1)∼(2x2)	5.0442e-01	∼	6.6636e-03	∼
(2x2)∼(4x4)	3.0356e-01	0.7326	2.1989e-03	1.5995
(4x4)∼(8x8)	1.8778e-01	0.6929	1.0448e-03	1.0736
(8x8)∼(16x16)	1.1073e-01	0.7620	4.8710e-04	1.1009
(16x16)∼(32x32)	6.7703e-02	0.7098	2.2031e-04	1.1447
$\varepsilon_1 = 0.01; \varepsilon_2 = 0$:				
(1x1)∼(2x2)	5.0456e-01	∼	6.6719e-03	∼
(2x2)∼(4x4)	3.0778e-01	0.7131	2.1280e-03	1.6486
(4x4)∼(8x8)	1.8236e-01	0.7551	9.6781e-04	1.1367
(8x8)∼(16x16)	1.1190e-01	0.7046	4.1818e-04	1.2106
(16x16)∼(32x32)	6.5896e-02	0.7639	1.6669e-04	1.3270
$\varepsilon_1 = 0; \varepsilon_2 = 0$:				
(1x1)∼(2x2)	5.0502e-01	∼	6.6770e-03	∼
(2x2)∼(4x4)	3.0409e-01	0.7318	2.2022e-03	1.6003
(4x4)∼(8x8)	1.8812e-01	0.6928	1.0455e-03	1.0748
(8x8)∼(16x16)	1.1092e-01	0.7621	4.8739e-04	1.1010
(16x16)∼(32x32)	6.7800e-02	0.7102	2.2043e-04	1.1448

Table 6.6: The errors and the rates of convergence for the DC-casting problem.

Bibliography

[1] F. Brezzi: *On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers*, RAIRO Numer. Anal., 8 (1974), pp. 129-151.

[2] A. M. Bruaset: *Object-oriented design of preconditioned iterative methods.* (To appear in the proceedings of the Colorado Conference on Iterative Methods'94.)

[3] R. E. Ewing and M. F. Wheeler: *Computational aspects of mixed finite element methods,* Numerical Methods for Scientific Computing (R. S. Stepleman, ed.), North-Holland, Amsterdam, 1983, pp. 163-172.

[4] M. C. Flemings: *Solidification processing,* McGraw-Hill, 1974.

[5] H. P. Langtangen: *Diffpack: Software for partial differential equations.* (To appear in the proceedings of OONSKI'94).

[6] H. P. Langtangen: *Programming Finite Element Methods in DIFFPACK. SINTEF-report*, 1992.

[7] H. P. Langtangen: *Finite Element Programming with Classes in DIFFPACK. SINTEF-report*, 1992.

[8] H. P. Langtangen: *Getting Started with C++ and DIFFPACK. SINTEF-report*, 1992.

[9] H. P. Langtangen and G. Pedersen: *Finite Element Preprocessors in DIFFPACK. SINTEF-report*, 1993.

[10] R.S. Falk and J.E. Osborn: *Error estimates for mixed methods*, R.A.I.R.O., 14 (1980), pp. 249-277.

[11] E. Haug, A. Mo and H. J. Thevik: *Macrosegregation near a cast surface caused by exudation and solidification shrinkage,* SINTEF-report, No. STF33 A94028, 1994. (Preprint of a paper submitted to the International Journal of Heat and Mass Transfer.)

[12] Claes Johnson: *Numerical solution of partial differential equations by the finite element method*, Claes Johnson and Studentlitteratur, Lund, 1987.

[13] M.-N. Le Roux: *A Mixed Finite Element Method for a Weighted Elliptic Problem,* R.A.I.R.O. Numerical Analysis, vol 16, no 3, 1982, pp. 243-273.

[14] T. B. Massalski (editor): *Binary alloy phase diagrams,* volume 1, ASM International, second edition, 1990.

[15] A. Mo: *Mathematical modelling of surface segregation in aluminium DC casting caused by exudation*, Int J. Heat Mass Transfer, Vol. 36, No. 18, pp. 4335-4340, 1993.

[16] J. Ni and C. Beckermann: *A volumn-averaged two-phase model for transport phenomena during solidification,* Metallurgical Transactions, 22B, pp. 349-361, 1991.

[17] C. C. Paige and M. A. Saunders: *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617-629.

[18] R. D. Pehkle, A. Jeyarajan and H. Wada: *Summary of Thermal Properties for Casting Alloys and Mold Materials,* Technical Report PB83-211003, National Technical Information Service, 1983.

[19] P. A. Raviart and J. M. Thomas: *A mixed finite element method for 2-nd order elliptic problems*, in Mathematical Aspects of Finite Element Methods, Lecture Notes in Mathematics 606, I. Galligani and E. Magenes, eds., Springer-Verlag, Berlin, New York, 1977, pp. 295-315.

[20] J. E. Roberts and J. -M. Thomas: *Mixed and hybrid method*, in Handbook of Numerical Analysis, Vol. II, P. G. Ciarlet and J. L. Lions, eds., Elsevier Science Publishers B.V. (North-Holland), 1991.

[21] T. F. Russel and M. F. Wheeler: *Finite element and finite difference methods for continuous flow in porous media*, in The Mathematics of Reservoir Simulation, R. E. Ewing, ed., Society for Industrial and Applied Mathematics, Philadelphia, PA, 1983.

[22] T. Rusten and R. Winther: *A preconditioned iterative method for saddlepoint problems*, SIAM J. Matrix Anal. Appl. Vol. 13, No. 3, pp. 887-904, July 1992.

[23] T. Rusten and R. Winther: *Substructure Preconditioners for Elliptic Saddle Point Problems,* Mathematics of Computation, Vol. 60, No. 201, pp. 23-48, 1993.

[24] W. Shen and A. M. Bruaset: *Mixed Finite Element Solution of Elliptic Boundary Value Problems; A Case Study Based on* `DIFFPACK`, SINTEF report, No. STF33 A94018, 1994.

[25] C. J. Smithells: *Metals Reference Book,* Butterworths, London and Boston, fifth eqition, 1976.

[26] B. Stroustrup: *The C++ Programming Language*, 2nd edition. Addison-Wesley, 1992.

[27] A. Weiser and M. F. Wheeler: *On convergence of block-centered finite differences for elliptic problems*, SIAM J. Numer. Anal., 25 (1988), pp. 351-375.

[28] O. C. Zienkiewicz and K. Morgan: *Finite elements and approximation*, University of Wales, Swansea, United Kingdom, John Wiley & Sons, Inc., 1983.

.

.