



FOR REFERENCE ONLY

Corrections

page 9	line 20/21	...and a residual ...
page 10	line 4	...elements is required ...
page 16	line 21	...smoother problems. Also ...
page 25	line 13	...this situation, the ...
page 39	line 6	...with a trace ...
page 42	line 5	replace $\tilde{r}_i^o$ by $\tilde{r}_o^h$
page 57	line 12	replace (4.3.5) by (4.3.5a)
page 57	line 22	...fracture problem will ...
page 58	Fig. 4.3	replace α by $2\pi - \alpha$
page 59	line 24	...assumed throughout Chapter ...
page 60	line 15	...modification of the ...
page 62	line 3	...error bounds will ...
page 64	Fig. 4.4	replace α by $2\pi - \alpha$
page 75	line 10	replace $ z _{1,1,\Omega}$ by $ z _{2,1,\Omega}$
page 76	line 18	...of u^h are ...
page 94	line 25	...global singular decomposition...
page 110	line 7	...the quantities derived ...
page 114	line 12	...diagram, Fig. 7.4, the ...
page 121	line 15	...using analogues to ...
page 123	line 18	...The situation never ...
page 125	line 7	replace $u(x)^p$ by $u(y)^2 dx dy$
page 125	line 12	...[Kufner et al. 1977]): ...
page 127	line 6	...the Second Symposium ...
page 127	line 12	combined effect of ...
page 128	line 22	...element approximations on ...

**Numerical Methods for Problems
in
Reservoir Simulation**

**Richard Jones
Imperial College
PhD**

September 1988

Abstract

The purpose of the work presented in this thesis is to investigate practical numerical methods for various reservoir simulation problems. The main focus is on finite element methods for approximating the reservoir pressure field in two situations where singularities occur. Mathematically these are elliptic problems with respectively source and reentrant corner singularities. Also a third problem is investigated, approximation of a system of conservation laws which represents a model enhanced oil recovery procedure. This situation is considered only in one dimension with the feature of interest being the inherent discontinuous solutions and their approximation via finite differences.

For the source singularity problem the standard singularity subtraction idea is formulated as a local subtraction technique. The method is then fully and rigorously analysed in the presence of variational crimes with optimal error bounds being obtained. To approximate the solutions of a general single reentrant corner problem a more practical version of the Dual Singular Function Method is developed and its error behaviour rigorously analysed. This is then generalized to allow approximation of a model fractured well problem. Finally a second order Roe's scheme is generalized for the conservation law system and compared with an exact solution derived for the problem.

To my parents and my Catherines

Acknowledgements

I would like to thank my supervisor Dr. John Barrett for his mathematical guidance and his patience and effort in helping me to write this thesis. In these respects I am also indebted to my original and 'executive' supervisor, Professor Charles Elliott. In addition I gratefully acknowledge the financial support of BP Exploration and the practical experience I gained with them and through my industrial supervisor Dr. Kevin Parrott.

Also on a lighter but no less relevant note I would like to thank my social, cultural, ethical and moral supervisors (in alphabetical order): Nick, Heather, Caroline, Sue, Ronan, Simon, Shaun, Joe, Colin, Dave, Ade and my fellow mathematicians especially Susan and cell-mates Nic and Robert.

Contents

I	Introduction	4
1	Mathematical Modelling of Oil Reservoirs	5
1.1	The Physical Problem	5
1.2	The Defining Equations and Pressure Fields	6
1.3	Numerical Methods	9
1.4	Chemical Flooding	11
II	Finite Elements for Elliptic Problems with Singularities	13
2	Source Modelling Methods for Elliptic Problems	14
2.1	Introduction	14
2.2	Fundamental Solutions and Finite Element Approximations .	15
2.3	Singularity Subtraction, Mixed Methods and Mesh Refinement	17
2.4	A Local Subtraction Technique	23
2.5	Gradient Superconvergence	25
3	Local Singularity Subtraction for the Pressure Equation	26
3.1	Introduction	26
3.2	The Locally Subtracted Pressure Problem	28
3.3	Practical Finite Element Approximation	31
3.4	The H^1 Error Estimate	34
3.5	The L_2 Error Estimate	39
3.6	Local Subtraction in the General Case	46

4	Fracture Problems and Domain Induced Singularities	49
4.1	Introduction	49
4.2	The Pressure Problem with an Infinite Permeability Well Region	50
4.3	A Model Fracture Problem and Approximate Solutions	54
4.4	Domain Induced Singularities and their Approximation . . .	57
4.5	A Practical Derivation of the DSFM	63
5	A Dual Singular Function Method Without Cutoff Functions	69
5.1	Introduction	69
5.2	Stress Intensity Factor Approximation and Basic Error Analysis	70
5.3	Enhanced F.E. Approximations	76
5.4	Optimal Approximations	79
5.5	General Applicability of a DSFM	84
6	Application of the DSFM to Two Test Problems	89
6.1	Introduction	89
6.2	A Simple Test Problem	90
6.3	Approximation of the Model Fracture Problem	92
III	A Chemical Flood Model	100
7	A Modified Roe's Scheme for a Parabolically Degenerate System of Hyperbolic Conservation Laws	101
7.1	Introduction	101
7.2	The Physical Problem and Modelling Equations	102
7.3	Analysis of the Model Equations and Solution of the Riemann Problem	104
7.4	Numerical Methods for the Chemical Flood System	115
7.5	Numerical Experiments and Implementation of Methods . . .	120
	Notation and Basic Results	125
	Bibliography	126

List of Figures

3.1	Two-well reservoir with discontinuous permeabilities	27
4.1	Reservoir section with vanishing conductivity region	51
4.2	The model fracture problem domain	55
4.3	The general single reentrant corner domain	58
4.4	The decomposed reentrant corner domain	64
6.1	The simple slit domain	90
6.2	The idealised fracture domain and computational subdomains	95
7.1	Riemann invariants in the (u, f) plane	111
7.2	Riemann invariants in the (u, c) plane	112
7.3	Solution path in the (u, c) plane	113
7.4	Solution path in the (u, f) plane	114
7.5	The Minmod average of two sets of vectors	119
7.6	Solutions on a 40 point grid - $u_r = 0.35$	122
7.7	Solutions on a 100 point grid - $u_r = 0.35$	122
7.8	Solutions on a 200 point grid - $u_r = 0.35$	123
7.9	Solutions on a 100 point grid - $u_r = 0.70$	124
7.10	Solutions on a 200 point grid - $u_r = 0.70$	124

List of Tables

6.1	Numerical Results for the Simple Crack Problem	92
6.2	Basic Stress Intensity Factor Estimates for the Fracture Problem	99

Part I

Introduction

Chapter 1

Mathematical Modelling of Oil Reservoirs

1.1 The Physical Problem

Oil has become of great importance in the modern industrial economy and this has resulted in an increased desire for knowledge about many aspects of the processes involved in its recovery. Information is required on basic matters such as the total amount of oil recoverable and on complex questions, for example regarding effects and efficiencies of enhanced recovery techniques or quantifying the benefits of fracturing wells. One approach used to study such matters is the formulation of mathematical models derived from physical laws governing oil reservoirs. These are then analysed to provide trends or details of system elements which can subsequently be of use in predicting attributes of theoretical proposals or as guidance, along with observed data, to indicate features of the system or its behaviour under practical conditions.

An oil reservoir is a complex physical system comprising in general a highly variable fluid mixture (of oils, gases and water) and its ill-defined containment (a stratified and fractured porous medium). Equations of state and motion for these fluids in porous media then provide the mathematical system which obviously will also be highly complicated. In fact approximations of first the physical and then the resulting mathematical system are

required before any solution can be attempted. It is the latter area which will be addressed in this thesis.

Certain physical aspects of reservoir modelling will be considered in the next section, focusing on the importance of the pressure equation and then deriving mathematical features of its solution. In the third section a general discussion of numerical methods for approximation of the pressure equation is presented, this being the main subject tackled in the thesis. Finally a complete simple enhanced recovery procedure is outlined in section 1.4, prior to being modelled and analysed in the final chapter, and relevant mathematical and numerical solution features briefly discussed.

1.2 The Defining Equations and Pressure Fields

A fundamental equation in mathematical modelling of flow in a porous medium is Darcy's law which in general form gives the velocity $\mathbf{v}$ thus:

$$\mathbf{v} = -\sigma(x, c)(\nabla p - \alpha(c)\nabla z). \quad (1.2.1)$$

Here σ is the fluid mobility (ratio of rock permeability to fluid viscosity) such that $0 < \sigma_l \leq \sigma \leq \sigma_u$, c its concentration, α fluid density, z a depth function (accounting for gravitational effects) and p the unknown pressure. This last quantity is defined using an equation of state which in many reservoir situations can be taken as the incompressibility condition i.e.

$$\nabla \cdot \mathbf{v} = q \quad (1.2.2)$$

where q represents fluid injection or production. To complete the description boundary conditions are needed; for example a Neumann condition

$$\mathbf{v} \cdot \boldsymbol{\nu} = g \quad \text{on } \partial\Omega, \quad (1.2.3)$$

where Ω is the region representing the reservoir and $\boldsymbol{\nu}$ its boundary's outward unit normal, prescribes the normal flow into Ω .

Equations (1.2.1–3) can be combined to yield an elliptic boundary value problem for the pressure thus:

$$\begin{aligned} -\nabla \cdot (\sigma \nabla p) &= f & \text{in } \Omega \\ \sigma \frac{\partial p}{\partial \boldsymbol{\nu}} &= g & \text{on } \partial\Omega \end{aligned} \quad (1.2.4)$$

where for consistency $\int_{\Omega} f + \int_{\partial\Omega} g = 0$. This consistency condition ensures the existence of a solution and from the conditions on σ the equation will be elliptic thus uniqueness is given up to an additive constant (which will be removed if the pressure is specified at any point). The equation's right hand side f is comprised of source/sink and gravitational terms (independent of p though the latter can be a function of c) and g is the Neumann data i.e. represents boundary flow conditions. In general g equals zero on outer boundaries of Ω (representing no out-flow from the reservoir) but may be non-zero within Ω . This latter condition will result from flow between the reservoir and fractures acting as injection or production wells. This will give rise to the total flux around the boundary being specified as the wells' fluid transfer rate.

Using (1.2.1) with a solution to (1.2.4) a reservoir flow field may be defined and from there to obtain information on fluid saturations mass conservation equations must be considered. Taking simple miscible displacement as an example of this the result is the following general parabolic equation [Peaceman 1977]

$$\phi \frac{\partial C}{\partial t} + \mathbf{v} \cdot \nabla C - \nabla \cdot (D \nabla C) = h(x, C) \quad \text{in } \Omega \quad (1.2.5)$$

where C is the oil saturation, $\phi(x)$ rock porosity, D a diffusion matrix and the right hand side represents fluid injection and production. For a complete description of the unknowns associated boundary and initial conditions are required. The former has the same form as (1.2.3) and the latter specifies an initial saturations i.e.

$$\begin{aligned} (D \nabla C) \cdot \nu &= d & \text{on } \partial\Omega \\ C(x, 0) &= C_o(x) & x \in \Omega. \end{aligned} \quad (1.2.6)$$

The main bulk of this thesis is concerned with the analysis of numerical methods for the accurate solution of (1.2.4). In particular since the velocity is defined by D'arcy's law (1.2.1) accurate pressure gradients are required for the solution of (1.2.5). Details of the important reservoir features that are associated with the pressure equation will now be given and then followed by a description of their effect on the solution.

Fundamental to (1.2.4) are the injection and production wells which drive the equation. They are usually manifest as holes of very small radius (in comparison to reservoir dimensions). Fluid transfer across the well boundary then occurs at a high rate which mathematically is equivalent to having a point source or sink at the well centre. The source/sink strength is then the well's total fluid flux (i.e. total outflow/inflow) and will give rise to delta functions within the right hand side of (1.2.2). This gives rise to an (apparent) infinite pressure at the well centre (i.e. mathematically the point which represents the well).

Another type of well occurs when a rock fracture (often artificially created) acts as a fluid transfer agent. Here the well size is no longer comparatively negligible and the appropriate mathematical model is now a prescribed total flux. The fracture appears as an internal boundary and mathematically the in/outflow is represented by the total flux condition:

$$\int_{\Gamma} \sigma \frac{\partial p}{\partial \nu} = q$$

where Γ is the fracture's boundary and q the well production/injection rate. In this case the pressure is constant on Γ (i.e. in the well) and an unknown quantity in the initial formulation though this is derived when a solution is constructed (as will be demonstrated in section 4.3).

Also of importance in (1.2.4) will be any discontinuity in rock or fluid properties. This will result in non-smooth behaviour of the pressure, for example a jump in σ is equivalent to an interface within the reservoir and usually a jump in the pressure gradient. A second situation may also occur which is the presence of fractures within the reservoir body. These will either act as flow barriers or infinite conductivity channels and as with the fracture-well they are modelled as internal boundaries with appropriate flow conditions.

Returning to the pressure equation it is noted that in the case of a standard source-modelled well locally (1.2.4) will have the form:

$$-\nabla \cdot (\sigma \nabla p) = q_o \delta(x - x_o) \quad (1.2.7)$$

where q_o is the well's injection rate. p may thus be written as

$$p = p_s + p_r \quad (1.2.8)$$

where p_s is singular and solves:

$$-\nabla \cdot (\sigma_o \nabla p) = q_o \delta(x - x_o) \quad (1.2.9)$$

with $\sigma_o = \sigma(x_o)$. p_s is then the Green's function (or fundamental solution) of this Laplace operator i.e. (in two-dimensions):

$$p_s = -\frac{q_o}{2\pi\sigma_o} \log |x - x_o| \quad (1.2.10)$$

and p_r satisfies

$$-\nabla \cdot (\sigma \nabla p_r) = \nabla \cdot ((\sigma - \sigma_o) \nabla p_s) = \nabla \sigma \cdot \nabla p_s \quad (1.2.11)$$

so is non-singular as the zero in $\sigma - \sigma_o$ at x_o cancels the singularity in ∇p_s .

Singular behaviour in the pressure/flow fields is also seen when modelling fractures. This results from domain irregularity which is due to the fractures being represented as slits in the plane i.e. the resulting internal boundary has reentrant corners of 2π at each end. Now it is well-known that when solving elliptic equations in polygonal domains with reentrant corners singularities are induced in the solutions and in this case the pressure will locally have the form

$$p = p_r + kr^{\frac{1}{2}} \sin \frac{1}{2}\theta \quad (1.2.12)$$

where $p_r \in H^2$ locally, k is unknown and (r, θ) are centred on the corner. In fact in general p may be decomposed locally into a singular expansion and residual of arbitrary smoothness given a sufficient number of singularities. Finally in the situation of discontinuous σ , the interface set up will induce a jump in the pressure gradient and thus also solution irregularity.

1.3 Numerical Methods

The basic requirements of a numerical method is to deliver accurate approximations to a problem under a practical implementation. Interpretation of

this criterion depends heavily on the problem concerned with the domain of definition and the equation system type both having large influences. In the case of the pressure problems analysed here their two-dimensional nature suggests that the flexibility of finite elements are required. However in the one-dimensional chemical flood this is unnecessary and problems with solution uniqueness (due to the system's hyperbolicity) imply that finite differences are more suitable. This latter case will be discussed further in section 1.4 but first problems associated with the approximation of pressure solutions will be considered.

In addition to their flexibility finite elements are also a natural choice for approximation of (1.2.4) with it being elliptic. However an implementation of the basic method will lead to poor results as the piecewise polynomials used are unable to approximate well singularities present in the pressure solution. Strategies for overcoming this problem are then the main focus of this thesis.

There are two main approaches to deal with these singularities, the first being to make use of a priori knowledge of the singularity form and the second to use mesh refinement around singularities so they may be more accurately approximated. The former is used in the methods to be analysed in this thesis with the basic idea being to include the singularities in the approximation space. In the source singularity case this becomes a straightforward reformulation of the problem to one with the singularity removed, as outlined in section 1.2. This is described in 2.3, along with other finite element methods used for the problem, and then used to derive an approximation technique which is subsequently analysed in Chapter 3. For the geometric singularity situation singularity amplitudes are unknown but a singular expansion is available and this has lead to more involved methods. These are the singular function method (SFM) as proposed by Babuška, Mejlík and Vítas [Babuška et al. 1961] and Blum and Dobrowolski's dual singular function method (DSFM) [Blum and Dobrowolski 1982] which also gives information on the unknown amplitudes. This latter technique, described in Chapter 4, is then used as the basis of a new, simpler and more practical formulation which is analysed in Chapter 5.

1.4 Chemical Flooding

The final topic to be considered is a full simple model simulation derived from an enhanced recovery process known as chemical flooding. Here water is injected to maintain a pressure gradient across the reservoir (as with standard water flooding) with it also containing chemicals to lower the water/oil interfacial tension. This greatly increases oil mobility and thus improves the efficiency of the recovery. The actual problem analysed is one-dimensional with incompressible flow which leads to a constant pressure field and contains two unknowns to be solved for, oil saturation and chemical concentration. The former quantity furnishes information on process efficiency. The physical derivation of this system is given in Chapter 7 and then from knowledge of the underlying mathematical structure a finite difference technique is constructed to approximate the unknowns. Finally to evaluate the method's performance an analytic solution is constructed with which the results of numerical experiments performed are then compared.

With no pressure field calculation necessary the problem is purely one of mass transport for the two variables. The resultant system is then first order and non-linear i.e.

$$\begin{aligned}\frac{\partial u}{\partial t}(x, t) + \frac{\partial f}{\partial x}(u, v) &= 0 \\ \frac{\partial v}{\partial t}(x, t) + \frac{\partial g}{\partial x}(u, v) &= 0\end{aligned}\tag{1.4.1}$$

where f and g are flux functions derived from fluid and reservoir properties. This is a system of hyperbolic conservation laws with the relevant solutions having none of the smoothness or uniqueness properties associated with elliptic equations. In fact an infinite number exist and they are discontinuous. There is however a unique physical solution.

Determining a physical solution from the mathematical description is achieved via a rule known as the entropy condition. Using characteristics an analytic solution may be constructed and implicit in this process is that it must have evolved along characteristic lines from the initial data. The entropy condition, which puts a constraint on solution discontinuities, is then deduced from this evolution requirement and provides uniqueness for

the physical problem. These ideas are discussed in Chapter 7.

When choosing an approximation technique for this problem three basic criteria were considered. Obviously capturing of the unique physical solution is essential and then the quality of approximations obtained is important, and thirdly the basic conservation property which motivates the equations was thought desirable. To satisfy the physically derived criteria a conservative scheme with a discrete entropy condition was chosen. Then when assessing the quality of approximation the factors of accuracy and stability of numerical solutions had to be considered. Simple first order schemes would give a low order of approximation, especially when trying to model discontinuities, whereas second order schemes provide potentially greater accuracy though produce unphysical oscillatory solutions in the vicinity of discontinuities. A way of combining the advantages of these methods is then to have a scheme which is monotone i.e. non-oscillatory and second order wherever possible. In this class of so-called TVD (total variation diminishing) schemes Roe's method [Roe 1981a] satisfies the physical criteria and thus was chosen as the basis for a numerical technique.

Finally because the flood model leads to (1.4.1) with a non-convex flux function there are resulting complexities which require special treatment. With this and the above criteria in mind a relevant generalization of Roe's scheme to two component systems is proposed and numerical experiments undertaken are described to illustrate its performance. Details of the formulation of this method are given in section 7.4 and its implementation with the numerical results given in section 7.5.

Part II

Finite Elements for Elliptic Problems with Singularities

Chapter 2

Source Modelling Methods for Elliptic Problems

2.1 Introduction

In Chapter 1 a simple mathematical formulation for reservoir pressure fields was briefly described, the main idea introduced being to consider wells as point sources or sinks (of pressure). As mentioned the mathematical representation of either of these features is a delta function in the inhomogeneous (right hand side) term of the defining elliptic equation (1.2.4), which then gives rise to a solution with singularities at each source/sink. The derivation of this singular behaviour is given immediately below and then resultant problems in its finite element approximation are described. The following section 2.3 reviews methods which have been developed to overcome these problems and in 2.4 motivation for a local subtraction technique is given along with an outline of its implementation. The chapter then ends with a summary of gradient superconvergence, later used to obtain enhanced accuracy flow rates, and describes an extension of this necessary for the class of elliptic problems being considered.

2.2 Fundamental Solutions and Finite Element Approximations

The following elliptic boundary value problem for the pressure p is considered:

$$\begin{aligned} -\nabla \cdot (\sigma \nabla p) &= f & \text{in } \Omega \\ \frac{\partial p}{\partial \nu} &= 0 & \text{on } \partial\Omega \end{aligned} \quad (2.2.1a)$$

$$p = p_u \quad \text{at } x = x_u \neq x_i \quad (2.2.1b)$$

where $0 < \sigma_l \leq \sigma \leq \sigma_u$ for ellipticity and:

$$f = \sum_i q_i \delta(x - x_i), \quad \text{with } \sum_i q_i = 0 \quad (2.2.2)$$

i.e. the reservoir Ω contains wells of strength q_i centred at points x_i . The requirement that $\sum_i q_i = 0$ is to ensure solution existence. Uniqueness is then guaranteed by the extra condition (2.2.1b) which can be replaced by an integral constraint as will be done in Chapter 3.

Now from (1.2.7–10) it is clear that p may be decomposed into regular and singular parts p_r and p_s respectively thus:

$$p = p_r + p_s \quad \text{where } p_s = \sum_i p_i \quad (2.2.3)$$

with p_i given as follows:

$$\begin{aligned} -\nabla^2 p_i &= \frac{q_i}{\sigma_i} \delta(x - x_i) & \text{in } \Omega \\ \Rightarrow p_i &= -\frac{q_i}{2\pi\sigma_i} \log |x - x_i| \end{aligned} \quad (2.2.4)$$

where $\sigma_i = \sigma(x_i)$. An equation analogous to (1.2.11) then defines p_r i.e., noting (2.2.1),

$$\begin{aligned} -\nabla \cdot (\sigma \nabla p_r) &= \sum_i \nabla \cdot ((\sigma - \sigma_i) \nabla p_i) = \sum_i \nabla \sigma \cdot \nabla p_i & \text{in } \Omega \\ \frac{\partial p_r}{\partial \nu} &= -\sum_i \frac{\partial p_i}{\partial \nu} & \text{on } \partial\Omega. \end{aligned} \quad (2.2.5)$$

This then demonstrates the presence of log singularities in a source-modelled well pressure equation and that they may be explicitly calculated.

Having obtained a formulation of singularities involved in the solution of (2.2.1) a finite element approximation to p is now defined with the influence of these singularities on relevant error bounds being discussed. A weak formulation of (2.2.1) defines $p \in W^{1,1}(\Omega)$ to be the function satisfying (2.2.1b) and:

$$(\sigma \nabla p, \nabla v) = (f, v) \quad \forall v \in W^{1,\infty}(\Omega) \quad (2.2.6)$$

where $(\cdot, \cdot)$ represents the usual L_2 inner product over Ω . Now in view of (2.2.2) a finite element approximation to p may now be defined as $p^h \in S^h(\Omega)$ satisfying (2.2.1b) and:

$$(\sigma \nabla p^h, \nabla \chi) = \sum_i q_i \chi(x_i) \quad \forall \chi \in S^h(\Omega) \quad (2.2.7)$$

where $S^h(\Omega) \subset W^{1,\infty}(\Omega)$ is a space of piecewise polynomials of degree $k-1$ ($k \geq 2$) defined over a partition of Ω (a more complete definition will be given in Chapter 3). For this approximation the L_2 error is given by Scott [Scott 1973] as follows:

$$\|p - p^h\|_{0,\Omega} \leq ch \quad (2.2.8)$$

independent of the order of polynomials in $S^h(\Omega)$. This result is sharp though a long way from the potentially optimal ch^k convergence available in sufficiently regular problems. Thus even for the basic linear elements this convergence is a power of h below that possible in smoother problems. Also this problem becomes worse when considering ∇p^h which is necessary to derive the flow field. Ewing and Wheeler [Ewing and Wheeler 1983] demonstrate that this quantity does not converge at all as is reasonable noting $p \in H^{1-\epsilon}(\Omega)$ and thus $\nabla p \in H^{-\epsilon}(\Omega)$.

Now if p^h is considered in a subdomain of Ω , Ω_s say, which contains no singularities then given sufficient regularity of σ , Nitsche and Schatz [Nitsche and Schatz 1974] show that optimal L_2 and H^1 error bounds would be obtained i.e.

$$\|p - p^h\|_{0,\Omega_s} \leq ch^k \quad (2.2.9)$$

$$\|p - p^h\|_{1,\Omega_s} \leq ch^{k-1}. \quad (2.2.10)$$

This demonstrates the severe loss of accuracy resulting from log singularities present in p and that the worst approximation performance occurs in well regions. These observations show that it is worthwhile seeking improvements in the treatment of source singularities.

2.3 Singularity Subtraction, Mixed Methods and Mesh Refinement

A universally applicable first step in improving approximations to equations of the form (2.2.6) is to subtract out the solution's principal singular behaviour. From (2.2.3) and (2.2.4) the right hand side of (2.2.6) may be rewritten thus:

$$(f, v) = \sum_i (-\nabla \cdot (\sigma_i \nabla p_i), v) = \sum_i \left\{ (\sigma_i \nabla p_i, \nabla v) - \left\langle \sigma_i \frac{\partial p_i}{\partial \nu}, v \right\rangle \right\} \quad (2.3.1)$$

where $\langle \cdot, \cdot \rangle$ is the usual L_2 inner product defined over the boundary of Ω , $\partial\Omega$, with outward unit normal ν . From this a new problem may be defined for the non-singular part of p , p_r , by replacing (f, v) in (2.2.6) so:

$$(\sigma \nabla p, \nabla v) = \sum_i (\sigma_i \nabla p_i, \nabla v) - \sum_i \left\langle \sigma_i \frac{\partial p_i}{\partial \nu}, v \right\rangle \quad (2.3.2)$$

and so from (2.2.1b), (2.2.5) and (2.3.2) p_r is defined by

$$p_r = p_u - \sum_i p_i(x_u) \quad \text{at } x = x_u \quad (2.3.3a)$$

and

$$(\sigma \nabla p_r, \nabla v) = \sum_i (\nabla \sigma \cdot \nabla p_i, v) - \sum_i \left\langle \sigma \frac{\partial p_i}{\partial \nu}, v \right\rangle \quad (2.3.3b)$$

$$= - \sum_i ((\sigma - \sigma_i) \nabla p_i, \nabla v) - \sum_i \left\langle \sigma_i \frac{\partial p_i}{\partial \nu}, v \right\rangle \quad (2.3.3c)$$

$$\forall v \in H^1(\Omega)$$

i.e. a weak form of (2.2.5). The right hand side of (2.3.3) is now non-singular as $\sigma - \sigma_i$ is zero at the $|x - x_i|^{-1}$ singularity of ∇p_i and p_r is thus

contained in $H^1(\Omega)$, in fact with sufficient regularity of σ (i.e. $\sigma \in C^1(\Omega)$) $p_r \in H^{2-\epsilon}(\Omega)$ but no smoother unless $\sigma - \sigma_i$ has a zero of order greater than one at x_i . This maximum regularity of p_r can be seen directly from (2.2.5) as with $p_i \in H^{1-\epsilon}(\Omega)$ the right hand side $(\nabla \sigma \cdot \nabla p_i) \in H^{-\epsilon}(\Omega)$. These restrictions will obviously not apply if σ is constant in the well regions for then the singularity will be completely cancelled as $\sigma - \sigma_i$ will have a zero of arbitrary order at x_i .

A Galerkin approximation to (2.3.3) can be directly defined using the notation associated with (2.2.7), i.e. p_r^h satisfying (2.3.3a) and:

$$(\sigma \nabla p_r^h, \nabla \chi) = - \sum_i ((\sigma - \sigma_i) \nabla p_i, \nabla \chi) - \sum_i \left\langle \sigma_i \frac{\partial p_i}{\partial \nu}, \chi \right\rangle \quad (2.3.4)$$

$$\forall \chi \in S^h(\Omega)$$

with an approximation to p immediately obtainable noting (2.2.3) thus:

$$p^h = p_r^h + \sum_i p_i. \quad (2.3.5)$$

The error in p^h is now simply that in p_r^h which, assuming $p_r \in H^{2-\epsilon}(\Omega)$, is given by Ewing and Wheeler [Ewing and Wheeler 1981] as:

$$\|p - p^h\|_{0,\Omega} = \|p_r - p_r^h\|_{0,\Omega} \leq ch^{2-\epsilon}. \quad (2.3.6)$$

A gradient approximation then follows from (2.3.5) similarly i.e.:

$$\nabla p^h = \nabla p_r^h + \sum_i \nabla p_i, \quad (2.3.7)$$

and here $\nabla p_r \in H^{1-\epsilon}(\Omega)$ with the subsequent error bounded from [Ewing and Wheeler 1981] thus:

$$\|\nabla(p - p^h)\|_{0,\Omega} = \|\nabla(p_r - p_r^h)\|_{0,\Omega} \leq ch^{1-\epsilon} \quad (2.3.8)$$

as expected.

This example of singularity subtraction shows that a great improvement can be effected using the technique here gaining nearly an extra power of h in L_2 and H^1 error bounds and, more importantly, convergent gradients.

These of course occur at the expense of having to evaluate integrals over Ω and $\partial\Omega$. The boundary integral will present no problem, a standard quadrature will give sufficient accuracy, though for the integral over Ω some care must be taken with any numerical integration in the well regions (where $(\sigma - \sigma_i)\nabla p_i \in H^{1-\epsilon}(\Omega)$). Also if the alternative right hand side formulation deriveable from (2.3.3b) is used then its singularity at the well will be even greater ($\nabla\sigma.\nabla p_i \in H^{-\epsilon}(\Omega)$) thus requiring more effort for accurate numerical integration. However the method will now give a convergent flow field and with piecewise linear basis functions the given errors are almost optimal.

A possible limitation on the use of this subtraction idea still exists as the smoothness condition applied to σ is now too restrictive. A jump in some reservoir or fluid property could cause σ to be discontinuous in which case p_r would reduce to its minimal smoothness i.e. $p_r \in H^1(\Omega)$ with the above error bounds not holding. However as was indicated in 1.2 this problem may be easily overcome. Solution irregularity will only occur at the actual jump in σ and so if the discontinuity line is fitted by the mesh the approximating polynomials will not be required to follow the irregularity. As shown by Barrett and Elliott [Barrett and Elliott 1985] this will allow optimal convergence if isoparametric fitting is used and thus will obviously solve the problem in this case.

Now turning to other methods for this pressure problem a standard way of improving approximation of singular solutions when using the finite element method is by having a mesh refined in such a way that more elements are available to approximate the singular behaviour. Eriksson [Eriksson 1985a] has developed an a priori adaptive refinement technique which delivers optimal L_1 estimates for approximations to the solution of (2.2.6) and its gradient with $f = \delta(x)$ and σ smooth. This of course does not imply L_2 convergence of these quantities, and no results are given for this norm, but optimal local L_∞ estimates are also derived for the approximation away from the singularity. A further improvement is then gained when the singularity is subtracted allowing the refinement conditions to be less restrictive while still delivering these optimal estimates [Eriksson 1985b].

Another technique used for this pressure problem is the mixed finite

element method [Raviart and Thomas 1977] and is specifically applied as fluid velocities derived from pressure gradients are a crucial input to the (oil) concentration equation (as described in 1.2). Darcy's law (1.2.1) and the equation of state (1.2.2) represent a coupled system for pressure and velocity which can be approximated simultaneously using this method (see Douglas, Ewing and Wheeler [Douglas et al. 1983] for a full treatment). If this system is smooth mixed finite elements allow the same order of approximation for both unknowns i.e. k if $k - 1^{\text{th}}$ order polynomials are used for the pressure approximation and k^{th} order for velocity. This obviously compares well with standard finite elements that use $k - 1$ order polynomials to give order k and $k - 1$ convergence respectively for pressure and velocity. Another advantage is that direct velocity approximations are obtained via this method. In general, though velocity and pressure are directly related, they may behave very differently, the former usually varying slowly whereas the latter (especially with large changes in σ) may fluctuate much more rapidly and thus be more difficult to approximate. Therefore in a situation where the velocity field is derived from pressure approximations pressure solution inaccuracies will be passed on and the idea in a mixed finite element formulation is then that the direct approximation property will minimise this problem.

Now obviously in a source singularity problem the pressure/velocity solutions are not smooth and again some subtraction will be required to give equations regular enough to provide convergent solutions. This is shown in the analysis of Douglas, Ewing and Wheeler [Douglas et al. 1983] who reformulate the original equations:

$$\mathbf{v} = -\sigma \nabla p, \quad \nabla \cdot \mathbf{v} = f$$

thus:

$$\mathbf{v}_r = -\sigma \nabla p - \mathbf{v}_s, \quad \nabla \cdot \mathbf{v} = 0 \quad (2.3.9)$$

where, using the same right hand side as in (2.2.1):

$$\mathbf{v}_s = \sum_i \frac{q_i}{2\pi} \nabla \log|x - x_i|. \quad (2.3.10)$$

This leads to a variational form:

$$\begin{aligned} \left(\frac{1}{\sigma}\mathbf{v}_r, \mathbf{u}\right) - (p, \nabla \cdot \mathbf{u}) &= \left(\frac{1}{\sigma}\mathbf{v}_s, \mathbf{u}\right) & \forall \mathbf{u} \in U \\ (\nabla \cdot \mathbf{v}_r, w) &= 0 & \forall w \in W \end{aligned} \quad (2.3.11)$$

where $W = L_2(\Omega)$ and $U \equiv \{\mathbf{u} \in W^2|\nabla \cdot \mathbf{u} \in W \text{ and } \mathbf{u} \cdot \nu = 0 \text{ on } \partial\Omega\}$ and a condition of the form (2.2.1b) is applied to p to ensure uniqueness. (2.3.11) is then approximated via the mixed finite element method using Raviart and Thomas spaces W_h and U_h , which are discretizations of W and U satisfying $\text{div}U_h \subset W_h$, thus defining approximate solutions p^h and $\mathbf{v}^h$ with:

$$\begin{aligned} \left(\frac{1}{\sigma}\mathbf{v}_r^h, \mathbf{u}\right) - (p^h, \nabla \cdot \mathbf{u}) &= \left(\frac{1}{\sigma}\mathbf{v}_s, \mathbf{u}\right) & \forall \mathbf{u} \in U_h \\ (\nabla \cdot \mathbf{v}_r^h, w) &= 0 & \forall w \in W_h \end{aligned} \quad (2.3.12)$$

where p^h will satisfy the same uniqueness condition as p and $\mathbf{v}^h = \mathbf{v}_r^h + \mathbf{v}_s$. The error in these approximations is given by:

$$\|p - p^h\|_W + \|\mathbf{v} - \mathbf{v}^h\|_U \leq c \left(\inf_{\phi \in W_h} \|p - \phi\|_W + \inf_{\mathbf{u} \in U_h} \|\mathbf{v}_r - \mathbf{u}\|_U \right). \quad (2.3.13)$$

This error equation demonstrates that the velocity subtraction performed to give (2.3.11) is essential otherwise no convergence would result in general with $v \in H^{-\epsilon}(\Omega)$. In this case the velocity equation being approximated is $\nabla \cdot \mathbf{v}_r = 0$, i.e. so $\mathbf{v}_r$ is smooth and the resultant bound derived from (2.3.13) will be limited by the smoothness of p . This limits the convergence of p^h to the $O(h)$ bound derived in (2.2.8) i.e. so the right hand side of (2.3.13) may be replaced by ch thus giving the same result for the velocity as the standard finite element method with subtraction though obviously not the enhanced pressure convergence. Obviously the reason for the pressure solution being so poor is that no subtraction was performed on its singularity which would be an obvious next step, though as yet this possibility has not been analysed.

Having now presented the various techniques involving finite elements applicable to the source singularity problem a brief comparison of relative advantages along with conclusions drawn therefrom will be given to close the section. The finite element method with global subtraction of singularities

is obviously very straightforward to implement being simply the application of a standard technique to a modified problem and with piecewise linears produces near optimal L_2 and H^1 error results (higher order basis functions are obviously unnecessary in this case providing no increase in accuracy unless some way can be found of subtracting out the higher order singularities, though this appears to be a complex problem). Eriksson's mesh refinement with subtraction will obviously give similar L_2 and H^1 error results (being simply a slight variation on standard finite elements). Also for general elements it has the advantage of optimal L_∞ results away from singularities (as well as converging at the optimal rate in $L_1(\Omega)$ and $W^{1,1}(\Omega)$) though this is not necessarily of great importance and there is the added complexity of an unflexible mesh structure. Mixed methods, as developed so far, will deliver velocity approximations to the same order of accuracy as standard finite elements (with respectively velocity and pressure subtractions of course) with the advantage of not having to use an intermediate step which practically could introduce further errors (i.e. if p is badly behaved). However with this technique pressure approximations are poor though it might be possible to incorporate a pressure singularity subtraction which would potentially give increased accuracy as described. Also mixed finite element methods are less straightforward to implement and computationally more expensive than standard finite elements and are obviously not such a well-known technique. Finally to apply a mixed finite element method for problems with discontinuous σ Douglas, Ewing and Wheeler [Douglas et al. 1983] indicate that some form of local subtraction is required before this may be done. This possibility is in fact investigated in the following section in conjunction with standard finite element methods and in fact leads to valuable simplifications in the approximation (full details are given in Chapter 3).

In conclusion all this discussion is based on the assumption that σ varies in the vicinity of singularities such that $\sigma - \sigma_i$ has only a zero of order one at x_i . If this zero is of higher order or σ is constant in the well region then a singularity subtracted pressure solution will be a great deal smoother or infinitely smooth (now depending on regularity of σ and boundaries) and could easily allow optimal error estimates to be attained for higher order

elements. In this case mixed finite element methods may perform better than standard finite elements giving an extra power of h in velocity convergence (see Russel and Wheeler [Russel and Wheeler 1983]) though of course at the expense of greater complexity and computational cost. Finally it should be noted that the ideas of gradient superconvergence [Levine 1985] may be applied to post process standard linear finite element results to obtain similar convergence, this being described in section 2.5.

2.4 A Local Subtraction Technique

In this chapter so far it has been seen that some form of singularity subtraction is necessary for obtaining reasonable approximations to the pressure problem containing source terms. The technique described for pressure approximations with finite elements requires global subtraction of these singularities though of course this only needs to be done where the singular behaviour is actually exhibited. On this basis subtraction should only take place in the well regions leaving the problem unmodified away from the singularities. This idea is then the basis of local subtraction and a general description of the technique will now be given along with an application for situations involving discontinuous σ .

As just outlined a reformulation of the source singularity problem is being sought where a local subtraction is used to eliminate the singular behaviour. Thus define p_o such that

$$\begin{aligned} p_o &= p - p_i && \in \Omega_i \\ p_o &= p && \in \Omega_o \end{aligned} \tag{2.4.1}$$

with p , p_i as given by (2.2.1) and (2.2.4), Ω_i containing x_i but no other singular point and where

$$\overline{\Omega}_o \cup (\cup_i \overline{\Omega}_i) = \overline{\Omega}. \tag{2.4.2}$$

This now allows an analogy of (2.3.3) to be constructed, defining p_o . (2.3.1)

may be rewritten thus:

$$\begin{aligned} (f, v) &= (f, v)_{\cup_i \Omega_i} = \sum_i (-\nabla \cdot (\sigma_i \nabla p_i), v)_{\Omega_i} \\ &= \sum_i \left\{ (\sigma_i \nabla p_i, \nabla v)_{\Omega_i} - \left\langle \sigma_i \frac{\partial p_i}{\partial \nu_i}, v \right\rangle_{\partial \Omega_i} \right\} \end{aligned} \quad (2.4.3)$$

where $(\cdot, \cdot)_{\Omega_i}$ and $\langle \cdot, \cdot \rangle_{\partial \Omega_i}$ are the standard L_2 inner products over Ω_i and $\partial \Omega_i$ respectively and ν_i is the outward normal on $\partial \Omega_i$. The equivalent of (2.3.2) is then

$$(\sigma \nabla p, \nabla v) = \sum_i (\sigma_i \nabla p_i, \nabla v)_{\Omega_i} - \left\langle \sigma_i \frac{\partial p_i}{\partial \nu_i}, v \right\rangle_{\partial \Omega_i} \quad (2.4.4)$$

and on splitting the left hand side thus:

$$(\sigma \nabla p, \nabla v) = (\sigma \nabla p, \nabla v)_{\Omega_o} + \sum_i (\sigma \nabla p, \nabla v)_{\Omega_i} \quad (2.4.5)$$

gives, noting (2.4.1), the local subtraction analogy of (2.3.3c) i.e.

$$(\sigma \nabla p_o, \nabla v) = - \sum_i ((\sigma - \sigma_i) \nabla p_i, \nabla v)_{\Omega_i} - \left\langle \sigma_i \frac{\partial p_i}{\partial \nu_i}, v \right\rangle_{\partial \Omega_i} \quad (2.4.6)$$

$$\forall v \in H^1(\Omega).$$

So (2.4.6) now gives a mathematical definition of the basic local subtraction technique. As required a function p_o has been defined which is non-singular in Ω , is also equal to p away from the singular points and obtaining p from p_o is obvious from (2.4.1). This has resulted in a right hand side integral to be performed over each subregion Ω_i and a line integral over the associated boundary of the flux (i.e. the source term in this formulation). It is thus now clear that computationally this is a much more convenient formulation of the problem. An approximation to (2.4.6) may be defined via a similar procedure to (2.3.4), noting that Ω is now subdivided (the details being given in section 3.3), with the resulting right hand side being much simpler. Integration of each singular function takes place over only one small region or boundary as opposed to integrating the full sum over the whole region and boundary $\partial \Omega$ as in (2.3.4).

Having now obtained an alternative subtraction technique the other problem raised at the beginning of this section may be tackled namely allowing for discontinuous σ . As previously described this is a standard smooth interface problem which can be dealt with by using a finite element mesh fitting this internal boundary as given by Barrett and Elliott [Barrett and Elliott 1985]. However in most physical situations the interface will surround a well, via a change in well-region rock properties or with an injected fluid setting up a front, and thus some subtraction will be required in addition. Here an obvious and straightforward way to deal with this is to define the subtraction region's boundary to be the interface which will tackle both problems simultaneously. In fact, to demonstrate local subtraction and give a rigorous error analysis for finite element solutions it is convenient to consider a problem incorporating this situation the details of which will be given in Chapter 3.

2.5 Gradient Superconvergence

As mentioned at the end of section 2.2 under certain conditions a singularity subtracted pressure problem may be smooth enough to allow optimal error estimates when using finite element and mixed methods in which case the former will not be able to deliver such good velocity approximations. However if the basic linear elements are used then velocities may be recovered from the pressure solutions in such a way that an $O(h^2)$ piecewise linear approximation is produced. This idea is known as gradient superconvergence and was developed by Levine [Levine 1985] for elliptic boundary value problems with Dirichlet data. The pressure problem being analysed in this thesis is of the same class though with Neumann boundary data for which the original analysis does not hold. However if the approximations are only considered away from the boundaries of the region the phenomenon will still be observed. Thus for all important areas of flow including the well regions this technique is applicable and will greatly enhance velocity field approximations.

Chapter 3

Local Singularity Subtraction for the Pressure Equation

3.1 Introduction

In this chapter a finite element approximation of the following model is analysed:

$$\begin{aligned} -\nabla \cdot (\sigma \nabla p) &= q\delta(x - x_1) - q\delta(x - x_2) && \text{in } \Omega \subset \mathbb{R}^2 \\ \frac{\partial p}{\partial \nu} &= 0 && \text{on } \partial\Omega \text{ and } \int_{\Omega} p = \lambda, \end{aligned} \tag{3.1.1}$$

where ν is an outward normal on $\partial\Omega$, the integral condition in (3.1.1) is imposed to ensure uniqueness of the pressure solution and injection/production rates are equal for consistency. Also as described in 2.4 the medium discontinuity being treated here is where σ is discontinuous across the boundary of a subregion of Ω enclosing a well, and in this model problem each well region will be taken to exhibit this behaviour. Ω is thus split into three subdomains such that

$$\Omega \equiv \Omega_o \cup \Omega_1 \cup \Omega_2 \cup \partial\Omega_1 \cup \partial\Omega_2 \tag{3.1.2}$$

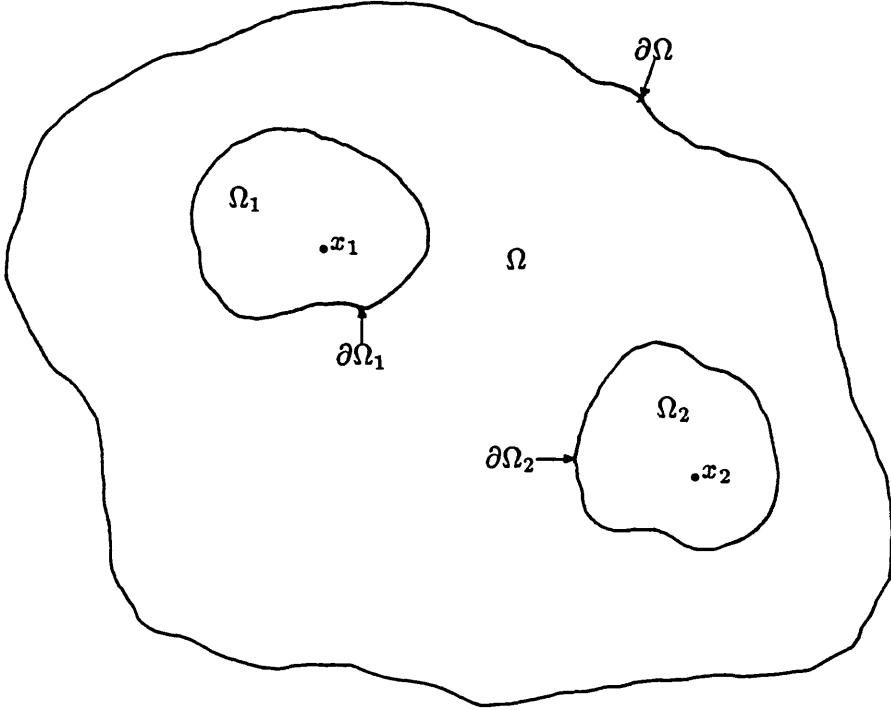


Figure 3.1: Two-well reservoir with discontinuous permeabilities

where $\partial\Omega_1$ and $\partial\Omega_2$ are the boundaries of Ω_1 and Ω_2 respectively, all of which are contained within Ω (see Fig. 3.1). Thus the singular points:

$$x_i \in \Omega_i \quad i = 1, 2 \quad (3.1.3)$$

and σ is taken to be constant (and positive for ellipticity) in the three subdomains, i.e.

$$\sigma = \sigma_i \quad \text{in } \Omega_i \quad i = 0 \rightarrow 2, \quad (3.1.4)$$

for simplicity as will be described below.

This model problem was chosen with σ piecewise constant in order to show the full potential of local subtraction when used with general order elements in a fully practical manner. As described in the previous chapter, with σ constant local subtraction will lead to a smooth problem thus allowing potentially optimal error bounds to be derived (as shown in sections 3.4

and 3.5). First however a full exposition of the problem will be given and its solution smoothness investigated (section 3.2). This is then followed by a full definition of the practical finite element method to be used (with allowance for variational crimes i.e. inexact domain representation and numerical integration) and its application to this problem where relevant approximation existence and uniqueness results will be proven (section 3.3). Finally to close the chapter, in section 3.6 the method's applicability for general σ will be described and error behaviour in this case discussed.

3.2 The Locally Subtracted Pressure Problem

A differential form of the pressure problem has already been given in (3.1.1) though for approximation the weak form is needed and, using the notation of section 2.3, this is to find $p \in W^{1,1}(\Omega)$ such that

$$\begin{aligned} (\sigma \nabla p, \nabla v) &= q(v(x_1) - v(x_2)) \quad \forall v \in W^{1,\infty}(\Omega) \\ \int_{\Omega} p &= \lambda. \end{aligned} \quad (3.2.1)$$

Now due to the nature of σ it is useful to split the inner product to correspond with the subdomains Ω_i ($i = 0 \rightarrow 2$) to which end the operator $a_i(u, v)$ is defined thus:

$$a_i(u, v) = (\sigma_i \nabla u, \nabla v)_{\Omega_i} \quad i = 0 \rightarrow 2 \quad (3.2.2)$$

and so (3.2.1) becomes:

$$\sum_{i=0}^2 a_i(p, v) = q(v(x_1) - v(x_2)). \quad (3.2.3)$$

Following the derivation of (2.3.3) this problem may have its source singularities removed by replacing the right hand side of (3.2.3) which may be written thus:

$$q(v(x_1) - v(x_2)) = \sum_{i=1}^2 \left\{ a_i(p_i, v) - \left\langle \sigma_i \frac{\partial p_i}{\partial \nu_i}, v \right\rangle_{\partial \Omega_i} \right\} \quad (3.2.4)$$

where ν_i is the outward normal on $\partial\Omega_i$ and p_i the singularity associated with x_i :

$$p_i = (-1)^i \frac{q}{2\pi\sigma_i} \log|x - x_i| \quad i = 1, 2. \quad (3.2.5)$$

So if r_i is defined thus:

$$\begin{aligned} r_i &= p - p_i & \text{in } \Omega_i & \quad i = 1, 2 \\ r_o &= p & \text{in } \Omega_o \end{aligned} \quad (3.2.6)$$

then combining (3.2.3) and (3.2.4) gives:

$$\sum_{i=0}^2 a_i(r_i, v) = - \sum_{i=1}^2 \left\langle \sigma_i \frac{\partial p_i}{\partial \nu_i}, v \right\rangle_{\partial\Omega_i} \quad \forall v \in H^1(\Omega). \quad (3.2.7)$$

(3.2.7) is then analogous to (2.4.6) for this problem, the right hand side integrals over Ω_i have disappeared due to σ being piecewise constant so no residual higher order singular behaviour exists and thus the subtracted pressure $\{r_i\}_{i=0}^2$ will be smooth (precise regularity conditions being given at the end of this section). Thus if μ is defined as follows:

$$\mu = \lambda - \sum_{i=1}^2 \int_{\Omega_i} p_i \quad (3.2.8)$$

the problem (3.2.1) with singularities subtracted is to find $r_i \in H^1(\Omega_i)$ ($i = 0 \rightarrow 2$) such that

$$r_i = r_o - p_i \quad \text{on } \partial\Omega_i \quad i = 1, 2 \quad (3.2.9a)$$

and

$$\begin{aligned} \sum_{i=0}^2 a_i(r_i, v) &= - \sum_{i=1}^2 \left\langle \sigma_i \frac{\partial p_i}{\partial \nu_i}, v \right\rangle_{\partial\Omega_i} \quad \forall v \in H^1(\Omega) \\ \sum_{i=0}^2 \int_{\Omega_i} r_i &= \mu. \end{aligned} \quad (3.2.9b)$$

Having now defined, in (3.2.9), a weak formulation of (3.1.1) with its source terms removed a formal treatment of the resultant solution's regularity is required before any analysis of errors in its approximation may be performed. To do this, global subtraction is used to define a residual and

two singular functions over Ω whose regularity can be determined and from which the locally subtracted pressure solutions $\{r_i\}_{i=0}^2$ may be obtained.

Initially write

$$R = p - \sum_{i=1}^2 p_i \quad \text{in } \Omega \quad (3.2.10)$$

which thus satisfies

$$\begin{aligned} -\nabla \cdot (\sigma \nabla R) &= 0 & \text{in } \Omega \\ \frac{\partial R}{\partial \nu} &= - \sum_{i=1}^2 \frac{\partial p_i}{\partial \nu} & \text{on } \partial \Omega \\ \int_{\Omega} R &= \mu' = \lambda - \sum_{i=1}^2 \int_{\Omega} p_i. \end{aligned} \quad (3.2.11)$$

(3.2.11) is then a standard Neumann problem with no singularities present and thus, due to ellipticity and the integral condition, is solved in a weak sense by a unique function $R \in H^1(\Omega)$. Then using (3.2.10) the functions $\{r_i\}_{i=0}^2$ now follow directly thus:

$$\begin{aligned} r_o &= R + \sum_{i=1}^2 p_i & \text{in } \Omega_o \\ r_1 &= R + p_2 & \text{in } \Omega_1 \\ r_2 &= R + p_1 & \text{in } \Omega_2 \end{aligned} \quad (3.2.12)$$

and regularity of $\{r_i\}_{i=0}^2$ will follow immediately from that of its constituent functions as defined here.

To determine the regularity of R in the relevant subdomains $\{\Omega_i\}_{i=0}^2$ note that its defining equation (3.2.11) is an interface problem with the right hand side and boundary data smooth on each Ω_i , $\partial \Omega_i$ respectively and with σ also smooth in the subdomains. From the analysis of Bramble and Schatz [Bramble and Schatz 1970] with these functions respectively in $H^k(\Omega_i)$, $H^{k+\frac{3}{2}}(\partial \Omega_i)$ and $C^{k,1}(\Omega_i)$ (where $C^{k,1}$ denotes functions whose k^{th} derivate is Lipschitz continuous), $R \in H^{k+2}(\Omega_i)$ provided that $\partial \Omega, \partial \Omega_i \in C^\infty$ ($i = 1, 2$). The first three conditions are obviously satisfied for the given problem and though this still leaves a condition on subregion boundaries

to be satisfied. Following the analysis of Grisvard [Grisvard 1976] for the solutions of elliptic problems this boundary regularity condition may be weakened to require $\partial\Omega, \partial\Omega_i \in C^{k+1,1}$ ($i = 1, 2$) for R to be in the stated space. Finally noting that p_1 is smooth in $\Omega_o \cup \Omega_2$ and p_2 smooth in $\Omega_o \cup \Omega_1$ it may be concluded that $r_i \in H^{k+2}(\Omega_i)$ ($i = 0 \rightarrow 2$) given $\partial\Omega, \Gamma_i \in C^{k+1,1}$ ($i = 0 \rightarrow 2$) and in the following a condition on boundary smoothness of this form will be assumed to give whatever regularity of $\{r_i\}_{i=0}^2$ is required for the error analysis performed.

3.3 Practical Finite Element Approximation

Having now defined a singularity subtracted problem from (3.1.1) the process of approximating its solution may now be considered. To this end a description will now be given of the finite element technique to be used which, to make it fully practical, includes isoparametric fitting of the subdomains boundaries $\partial\Omega_i$ ($i = 1, 2$). Boundary fitting is essential for optimal approximations and a detailed exposition of this and the resultant method will be given. Also the question of approximation existence and uniqueness is addressed with a rigorous proof being given to close the section.

To construct a finite element approximation to (3.2.7) the region Ω must first be partitioned and done in a way so that the internal regions Ω_i ($i = 1, 2$) are isoparametrically fitted. This gives rise to regions Ω_i^h ($i = 0 \rightarrow 2$) and associated boundaries $\partial\Omega_i^h$ ($i = 1, 2$) approximating Ω_i and $\partial\Omega_i$ respectively, and here it is noted that for ease of exposition the boundary $\partial\Omega$ is assumed to be fitted exactly. As a result the partitions T^h and T_i^h containing finite elements τ are defined, where $\bar{\Omega} = \bigcup_{\tau \in T^h} \bar{\tau}$ and $T_i^h = \{\tau : \tau \cap \Omega_i^h \neq \emptyset\}$, along with the boundary sets B_i^h where:

$$B_i^h = \{\tau \in T_j^h, j = 0, i; \underline{m}(\tau \cap \partial\Omega_i^h) \neq 0 \text{ in } R^1\} \quad i = 1, 2. \quad (3.3.1)$$

Also, associated with the partition T^h is a finite dimensional subspace $S^h \subset C^0(\bar{\Omega})$ of order $k \geq 2$ with $\forall \chi \in S^h, \chi|_{\tau} \in H^2(\tau) \forall \tau \in T^h$. Using this subspace an interpolation operator $\Pi_h : C^0(\bar{\Omega}) \rightarrow S^h$ is defined and with these simplicial Lagrangian isoparametric elements (Ciarlet and

Raviart [Ciarlet and Raviart 1972]) the following is satisfied,

$$|w - \Pi_h w|_{m,\tau} \leq ch^{k-m} \|w\|_{k,\tau} \quad \forall w \in H^k(\tau), \quad \forall \tau \in T^h \quad (3.3.2)$$

where $k \geq m \geq 0, k \geq 2$.

As just described the main importance of using isoparametric fitting of the regions Ω_i ($i = 0 \rightarrow 2$) is to obtain optimal approximations to the boundaries $\partial\Omega_i^h$ ($i = 1, 2$). This means that if $\partial\Omega_i^h \cap \bar{\tau}$ is described by $\psi_\tau^h(X_\tau)$ (i.e. $X_\tau \in \Delta_\tau$, a local coordinate interval) and $\partial\Omega_i \cap \bar{\tau}$ by $\psi_\tau(X_\tau)$, where $\psi_\tau, \psi_\tau^h \in C^{1,1}(\Delta_\tau)$ and vanish at the vertices of Δ_τ then:

$$\|\psi_\tau - \psi_\tau^h\|_{m,\infty,\Delta_\tau} \leq ch^{k-m}, \quad m = 0, 1 \quad \forall \tau \in B_i^h \quad (3.3.3a)$$

$$\|\nabla \psi_\tau\|_{0,\infty,\Delta_\tau} \leq ch, \quad \|\nabla \psi_\tau^h\|_{0,\infty,\Delta_\tau} \leq ch \quad \forall \tau \in B_i^h \quad (3.3.3b)$$

The power of this boundary approximation also allows a bound to be placed on the norms of functions defined in the skin between $\partial\Omega_i$ and $\partial\Omega_i^h$ i.e. for $u \in H^1(\Omega_i \cup \Omega_i^h)$

$$\|u\|_{0,\Omega_i \setminus \Omega_i^h} \leq c \left(h^k |u|_{1,\Omega_i \setminus \Omega_i^h} + h^{\frac{k}{2}} \|u\|_{0,\partial\Omega_i} \right) \quad (3.3.4a)$$

$$\|u\|_{0,\Omega_i^h \setminus \Omega_i} \leq c \left(h^k |u|_{1,\Omega_i^h \setminus \Omega_i} + h^{\frac{k}{2}} \|u\|_{0,\partial\Omega_i^h} \right) \quad (3.3.4b)$$

from [Barrett and Elliott 1987].

A result of fitting the regions is that the Ω_i^h do not necessarily lie within their respective Ω_i ($i = 0 \rightarrow 2$). Thus to define functions within Ω_i^h , regions $\tilde{\Omega}_i$ (with smooth boundaries $\partial\tilde{\Omega}_i$) are introduced where $\Omega_i \cup \Omega_i^h \subset \tilde{\Omega}_i$ and the relevant functions and operators are extended into these regions. In the given problem with σ_i constant these extensions are obvious and so in the $\tilde{\Omega}_i$ differential operators $\tilde{A}_i$ are defined identical to A_i thus:

$$\tilde{A}_i u \equiv -\nabla \cdot (\sigma_i \nabla) \equiv A_i u \quad i = 0 \rightarrow 2 \quad (3.3.5a)$$

from which integral operators for the approximation are given as:

$$\begin{aligned} a_i^h(u, v) &\equiv (\sigma_i \nabla u, \nabla v)_{\Omega_i^h} = (\tilde{A}_i u, v)_{\Omega_i^h} + \left\langle \sigma_i \frac{\partial u}{\partial \nu_i^h}, v \right\rangle_{\partial\Omega_i^h} \quad i = 1, 2 \\ a_o^h(u, v) &\equiv (\sigma_o \nabla u, \nabla v)_{\Omega_o^h} = (\tilde{A}_o u, v)_{\Omega_o^h} + \left\langle \sigma_o \frac{\partial u}{\partial \nu_o}, v \right\rangle_{\partial\Omega_o} \\ &\quad - \sum_{i=1}^2 \left\langle \sigma_o \frac{\partial u}{\partial \nu_i^h}, v \right\rangle_{\partial\Omega_i^h} \end{aligned} \quad (3.3.5b)$$

with ν_i^h the outward normal on $\partial\Omega_i^h$. To extend functions into these new regions the extension operator $E_i : H^s(\Omega_i) \rightarrow H^s(\tilde{\Omega}_i)$ ($i = 0 \rightarrow 2$) is defined so that:

$$E_i w_i = w_i \text{ on } \Omega_i, \quad \|E_i w_i\|_{s, \Omega_i} \leq c \|w_i\|_{s, \Omega_i} \quad \forall w_i \in H^s(\Omega_i), s \geq 1 \quad (3.3.6)$$

where c is independent if w and $E_i w_i$ is denoted by $\tilde{w}_i$ (see Kufner, John and Fučík [Kufner et al. 1977]).

These definitions now allow the formulation of a finite element approximation to (3.2.9): find $r_i^h \in S^h(\Omega_i^h)$ ($i = 0 \rightarrow 2$) such that

$$\begin{aligned} \sum_{i=0}^2 a_i^h(r_i^h, \chi) &= - \sum_{i=1}^2 \left\langle \sigma_i \frac{\partial p_i}{\partial \nu_i^h}, \chi \right\rangle_{\partial\Omega_i^h} \quad \forall \chi \in S^h \subset H^1(\Omega) \\ \sum_{i=0}^2 \int_{\Omega_i^h} r_i^h &= \mu, \end{aligned} \quad (3.3.7)$$

$$\text{with } r_i^h = r_o^h - \Pi_h p_i \quad \text{on } \partial\Omega_i^h \quad (i = 1, 2). \quad (3.3.8)$$

$S^h(\Omega_i^h)$ is the finite dimensional subspace of $H^1(\Omega_i)$ analogous to S^h and the $r_i^h \in S^h(\Omega_i^h)$ are constrained on the boundaries $\partial\Omega_i^h$ to behave like their continuous counterparts in (3.2.9a). (3.3.7–8) is then an obvious approximation of (3.2.9). It is noted that the integrals on the left hand side of (3.3.7) can be evaluated exactly using isoparametric integration. Although the right hand side integral similarly treated will contain an error for ease of exposition it is assumed that all integrals in (3.3.7) are evaluated exactly.

Having now formulated the problem, the existence and uniqueness of a solution is required before proceeding to prove its powers of approximation. With (3.3.7–8) a finite-dimensional system the existence of a solution will follow from uniqueness. Assume $\{q_i^h\}_{i=0}^2$ and $\{r_i^h\}_{i=0}^2$ are both solutions of (3.3.7) then:

$$\sum_{i=0}^2 a_i^h(q_i^h - r_i^h, \chi) = 0 \quad \forall \chi \in S^h. \quad (3.3.9)$$

Now defining $d^h = d_i^h = q_i^h - r_i^h$ in Ω_i^h , then as

$$q_i^h - q_o^h = \Pi_h s_i = r_i^h - r_o^h \quad \text{on } \partial\Omega_i^h \quad (i = 1, 2)$$

d^h is continuous across $\partial\Omega_i$ and thus $d^h \in S^h$ i.e. setting $\chi = d^h$ in (3.3.9) gives:

$$\sum_{i=0}^2 a_i^h(d_i^h, d_i^h) = 0. \quad (3.3.10)$$

Also the Poincaré inequality states that there exists c depending only on G such that

$$\|w\|_{0,G}^2 \leq c \left\{ |w|_{1,G}^2 + \left(\int_G w \right)^2 \right\} \quad \forall w \in H^1(G) \quad (3.3.11)$$

$$\text{i.e.} \quad \|d^h\|_{0,\Omega}^2 \leq c \left\{ |d^h|_{1,\Omega}^2 + \left(\int_{\Omega} d^h \right)^2 \right\}$$

and from (3.3.7) $\int_{\Omega} d^h = \sum_{i=0}^2 \int_{\Omega_i^h} d_i^h = 0$, thus using (3.3.9) and (3.3.11):

$$\|d^h\|_{0,\Omega}^2 \leq c |d^h|_{1,\Omega} \leq \sum_{i=0}^2 a_i^h(d_i^h, d_i^h) = 0$$

which ensures q_i^h and r_i^h coincide as required and that a solution $\{r_i^h\}_{i=0}^2$ solving (3.3.7) is unique and thus also exists.

3.4 The H^1 Error Estimate

The error analysis requires the following preliminary lemmas.

Lemma 3.4.1

Noting (3.3.3a) and (3.3.6)

$$\sum_{i=1}^2 \left\| \sigma_i \frac{\partial(\tilde{r}_i + p_i)}{\partial \nu_i^h} - \sigma_o \frac{\partial \tilde{r}_o}{\partial \nu_i^h} \right\|_{0,\partial\Omega_i^h} \leq \begin{cases} ch^{\frac{k}{2}} \left(\sum_{i=0}^2 \|r_i\|_{2,\Omega_i} + \sum_{i=1}^2 \|p_i\|_{2,\Omega_o} \right) & (3.4.1a) \\ ch^{k-1} \left(\sum_{i=0}^2 \|r_i\|_{4,\Omega_i} + \sum_{i=1}^2 \|p_i\|_{4,\Omega_o} \right). & (3.4.1b) \end{cases}$$

Proof:

The purpose of this lemma is to give a bound on the discrepancy of $\sigma \partial \tilde{p} / \partial \nu_i^h$ across $\partial\Omega_i^h$ ($i = 1, 2$). This may be obtained by a simple adaptation of a similar interface result of Barrett and Elliott ([Barrett and Elliott 1985])

Lemma 2.1). A direct application of this result is not possible as $\tilde{p}$ is singular in Ω_i^h but by writing:

$$\frac{\partial \tilde{p}}{\partial \nu_i^h} = \frac{\partial}{\partial \nu_i^h} (\tilde{r}_i + p_i) \quad i = 1, 2$$

then noting (3.2.6) this discrepancy may be written as

$$\left\| \left(\sigma_o \frac{\partial \tilde{r}_o}{\partial \nu_i^h} - \sigma_i \frac{\partial p_i}{\partial \nu_i^h} \right) - \sigma_i \frac{\partial \tilde{r}_i}{\partial \nu_i^h} \right\|_{0, \partial \Omega_i^h}$$

to which the analysis of the quoted result may be applied directly to deliver the bounds (3.4.1a) and (3.4.1b). $\square$

Lemma 3.4.2

Noting the definition (3.3.5a), for $i = 0 \rightarrow 2$

$$\left| (\tilde{A}_i \tilde{r}_i, v)_{\Omega_i^h} \right| \leq ch^k \|r_i\|_{3, \Omega_i} \|v\|_{1, \Omega} \quad \forall v \in H^1(\Omega_i^h) \quad (3.4.2)$$

Proof:

From the definition of $\tilde{r}_i$;

$$(\tilde{A}_i \tilde{r}_i, v)_{\Omega_i^h} = (\tilde{A}_i \tilde{r}_i, v)_{\Omega_i^h \setminus \Omega_i} \quad i = 0 \rightarrow 2.$$

Now using (3.3.4a) and applying a trace inequality to the second norm on the r.h.s. gives:

$$\begin{aligned} \left| (\tilde{A}_i \tilde{r}_i, v)_{\Omega_i^h} \right| &\leq \left| (\tilde{A}_i \tilde{r}_i, v)_{\Omega_i^h \setminus \Omega_i} \right| \\ &\leq \left\| \tilde{A}_i \tilde{r}_i \right\|_{0, \Omega_i^h \setminus \Omega_i} \|v\|_{0, \Omega_i^h \setminus \Omega_i} \\ &\leq ch^k \left\| \tilde{A}_i \tilde{r}_i \right\|_{1, \Omega_i^h} \|v\|_{1, \Omega_i^h} \\ &\leq ch^k \|r_i\|_{3, \Omega_i} \|v\|_{1, \Omega_i^h} . \square \end{aligned}$$

The statement and proof of the main result now follows.

Theorem 3.4.3

$$\left(\sum_{i=0}^2 \|\tilde{r}_i - r_i^h\|_{1, \Omega_i^h}^2 \right)^{\frac{1}{2}} \leq ch^{k-1} \left[\sum_{i=0}^2 \|r_i\|_{k+2, \Omega_i}^2 + \sum_{i=1}^2 \|p_i\|_{4, \Omega_o}^2 \right]^{\frac{1}{2}} . \quad (3.4.3)$$

Proof:

Initially the error is split into two distinct quantities, the interpolation error and the discrepancy between the solution interpolant and the approximation. Thus:

$$\sum_{i=0}^2 \|\tilde{r}_i - r_i^h\|_{1,\Omega_i^h}^2 \leq 2 \left(\sum_{i=0}^2 \|\tilde{r}_i - \hat{r}_i^h\|_{1,\Omega_i^h}^2 + \|\hat{r}_i^h - r_i^h\|_{1,\Omega_i^h}^2 \right) \quad (3.4.4)$$

where $\hat{w}_i^h$ is a modification to the interpolant $\Pi_h \tilde{w}_i$ defined by

$$\hat{w}_i^h = \Pi_h \tilde{w}_i + \frac{1}{\underline{m}(\Omega)} \left(\mu - \sum_{i=0}^2 \int_{\Omega_i^h} \Pi_h \tilde{w}_i \right) \quad \forall w_i \in H^2(\Omega_i) \quad i = 0 \rightarrow 2. \quad (3.4.5)$$

This ensures that the $\{\hat{r}_i^h\}_{i=0}^2$ satisfy the integral condition of (3.3.7) so that on defining the second error component,

$$e_i^h = \hat{r}_i^h - r_i^h \quad \text{in } \Omega_i^h \quad i = 0 \rightarrow 2 \quad (3.4.6a)$$

then:

$$\sum_{i=0}^2 \int_{\Omega_i^h} e_i^h = 0. \quad (3.4.6b)$$

The first term of (3.4.4) may be treated by a further decomposition thus:

$$\sum_{i=0}^2 \|\tilde{r}_i - \hat{r}_i^h\|_{1,\Omega_i^h}^2 \leq 2 \left(\sum_{i=0}^2 \|\tilde{r}_i - \Pi_h \tilde{r}_i\|_{1,\Omega_i^h}^2 + \left(\mu - \sum_{i=0}^2 \int_{\Omega_i^h} \Pi_h \tilde{r}_i \right)^2 \right). \quad (3.4.7)$$

Now using (3.3.2) and (3.3.6):

$$\begin{aligned} \sum_{i=0}^2 \|\tilde{r}_i - \Pi_h \tilde{r}_i\|_{1,\Omega_i^h}^2 &\leq ch^{2(k-1)} \sum_{i=0}^2 \|\tilde{r}_i\|_{k,\Omega_i^h}^2 \\ &\leq ch^{2(k-1)} \sum_{i=0}^2 \|r_i\|_{k,\Omega_i}^2. \end{aligned} \quad (3.4.8)$$

Also noting the definition of μ from (3.2.8) the second term in (3.4.7) is bounded, again using (3.3.2) and (3.3.6) thus:

$$\left[\sum_{i=0}^2 \left(\int_{\Omega_i} r_i - \int_{\Omega_i^h} \Pi_h \tilde{r}_i \right) \right]^2 = \sum_{i=0}^2 \left[\int_{\Omega_i^h} (\tilde{r}_i - \Pi_h \tilde{r}_i) + \int_{\Omega_i \setminus \Omega_i^h} r_i - \int_{\Omega_i^h \setminus \Omega_i} \tilde{r}_i \right]^2$$

$$\begin{aligned}
&\leq c \sum_{i=0}^2 \left(h^{2k} \|r_i\|_{k,\Omega_i}^2 + m(\Omega_i \setminus \Omega_i^h)^2 \|r_i\|_{0,\infty,\Omega_i}^2 \right. \\
&\quad \left. + m(\Omega_i^h \setminus \Omega_i)^2 \|\tilde{r}_i\|_{0,\infty,\Omega_i^h}^2 \right) \\
&\leq ch^{2k} \sum_{i=0}^2 \|r_i\|_{k,\Omega_i}^2.
\end{aligned} \tag{3.4.9}$$

Thus combining (3.4.7), (3.4.8) and (3.4.9) gives the required result,

$$\sum_{i=0}^2 \|\tilde{r}_i - \tilde{r}_i^h\|_{1,\Omega_i^h}^2 \leq ch^{2(k-1)} \sum_{i=0}^2 \|r_i\|_{k,\Omega_i}^2. \tag{3.4.10}$$

To begin treatment of the second term in (3.4.4) a fundamental property of the error $\{e_i^h\}_{i=0}^2$ is now demonstrated, namely its continuity across $\partial\Omega_i^h$ ($i = 1, 2$). To show this α is defined thus:

$$\alpha = \frac{1}{\underline{m}(\Omega)} \left(\mu - \sum_{i=0}^2 \int_{\Omega_i^h} \Pi_h \tilde{r}_i \right)$$

and hence from (3.4.6) e_i^h may be written as:

$$e_i^h = \Pi_h \tilde{r}_i + \alpha - r_i^h \quad \text{in } \Omega_i^h \tag{3.4.11}$$

where from (3.3.8) on the boundaries $\partial\Omega_i^h$ ($i = 1, 2$)

$$\begin{aligned}
e_i^h - e_o^h &= \Pi_h \tilde{r}_i - \Pi_h \tilde{r}_o - (r_i^h - r_o^h) \\
&= \Pi_h \tilde{r}_i - \Pi_h \tilde{r}_o + \Pi_h p_i.
\end{aligned} \tag{3.4.12}$$

Now from the definition of interpolation on an isoparametrically fitted mesh, the values sampled to define the interpolating function will be obtained from the region being fitted or its boundary, but not outside this region. Thus $\Pi_h \tilde{r}_i$ and $\Pi_h \tilde{r}_o$ on $\partial\Omega_i^h$ will be derived from the values of $\tilde{r}_i$ and $\tilde{r}_o$ at the relevant nodal points on $\partial\Omega_i$, and on $\partial\Omega_i$ $\tilde{r}_i = r_i$. From this it can be seen that at nodal points of $\partial\Omega_i^h$, $\tilde{r}_i - \tilde{r}_o = r_i - r_o = -p_i$ and thus (from 3.4.12)) $e_i^h - e_o^h = 0$ which implies, more generally

$$e_i^h - e_o^h = 0 \quad \text{on } \partial\Omega_i^h \quad i = 1, 2. \tag{3.4.13}$$

Thus if e^h is now defined as follows:

$$e^h \equiv e_i^h \quad \text{in } \Omega_i^h \quad i = 0 \rightarrow 2 \quad (3.4.14)$$

then e^h is continuous across $\partial\Omega_i^h$ ($i = 1, 2$) and therefore $e^h \in S^h$.

The above result now allows a bound to be derived for the second term of (3.4.4) by use of Poincaré's inequality. Applying the inequality to e^h gives:

$$\begin{aligned} \sum_{i=0}^2 \|e^h\|_{1,\Omega_i^h}^2 &= \|e^h\|_{1,\Omega}^2 \leq c \left\{ |e^h|_{1,\Omega}^2 + \left(\int_{\Omega} e^h \right)^2 \right\} \\ &\leq c \sum_{i=0}^2 a_i^h(e^h, e^h) \end{aligned} \quad (3.4.15)$$

from (3.4.6) and ellipticity. Now:

$$\begin{aligned} \sum_{i=0}^2 a_i^h(e^h, e^h) &= \sum_{i=0}^2 \left(a_i^h(\hat{r}_i^h - \tilde{r}_i, e^h) + a_i^h(\tilde{r}_i, e^h) \right) \\ &\quad + \sum_{i=1}^2 \int_{\partial\Omega_i^h} \sigma_i \frac{\partial p_i}{\partial \nu_i^h} e^h \end{aligned} \quad (3.4.16)$$

using (3.3.7), and from (3.3.5) with (3.2.6)

$$\begin{aligned} \sum_{i=0}^2 a_i^h(\tilde{r}_i, e^h) + \sum_{i=1}^2 \int_{\partial\Omega_i^h} \sigma_i \frac{\partial p_i}{\partial \nu_i^h} e^h &= \sum_{i=1}^2 \int_{\partial\Omega_i^h} \left(\sigma_i \frac{\partial(\tilde{r}_i + p_i)}{\partial \nu_i^h} - \sigma_o \frac{\partial \tilde{r}_o}{\partial \nu_i^h} \right) e^h \\ &\quad + \sum_{i=0}^2 (\tilde{A}_i \tilde{r}_i, e^h)_{\Omega_i^h} \end{aligned} \quad (3.4.17)$$

which when combined with (3.4.16) yields:

$$\begin{aligned} \sum_{i=0}^2 a_i^h(e^h, e^h) &\leq \sum_{i=0}^2 \left| a_i^h(\hat{r}_i^h - \tilde{r}_i, e^h) \right| + \sum_{i=0}^2 \left| (\tilde{A}_i \tilde{r}_i, e^h)_{\Omega_i^h} \right| \\ &\quad + \sum_{i=1}^2 \left| \int_{\partial\Omega_i^h} \left(\sigma_i \frac{\partial(\tilde{r}_i + p_i)}{\partial \nu_i^h} - \sigma_o \frac{\partial \tilde{r}_o}{\partial \nu_i^h} \right) e^h \right|. \end{aligned} \quad (3.4.18)$$

These three derived quantities may now easily be bounded using earlier results. Firstly from the definition of $\hat{r}_i^h$:

$$\sum_{i=0}^2 \left| a_i^h(\hat{r}_i^h - \tilde{r}_i, e^h) \right| = \sum_{i=0}^2 \left| a_i^h(\Pi_h \tilde{r}_i - \tilde{r}_i, e^h) \right|$$

$$\leq ch^{k-1} \left(\sum_{i=0}^2 \|r_i\|_{k,\Omega_i}^2 \right)^{\frac{1}{2}} \|e^h\|_{1,\Omega} \quad (3.4.19)$$

using (3.3.5b) and (3.3.2). The bound on the second term follows directly from (3.5.2) i.e.

$$\begin{aligned} \sum_{i=0}^2 \left| (\tilde{A}_i \tilde{r}_i, e^h)_{\Omega_i^h} \right| &\leq ch^k \left(\sum_{i=0}^2 \|r_i\|_{3,\Omega_i} \right) \|e^h\|_{1,\Omega} \\ &\leq ch^k \left(\sum_{i=0}^2 \|r_i\|_{3,\Omega_i}^2 \right)^{\frac{1}{2}} \|e^h\|_{1,\Omega} \end{aligned} \quad (3.4.20)$$

and then from (3.4.1b) combined with the trace inequality,

$$\begin{aligned} \sum_{i=1}^2 \left| \int_{\partial\Omega_i^h} \left(\sigma_i \frac{\partial(\tilde{r}_i + p_i)}{\partial\nu_i^h} - \sigma_o \frac{\partial\tilde{r}_o}{\partial\nu_i^h} \right) e^h \right| \\ \leq \sum_{i=1}^2 \left\| \sigma_i \frac{\partial(\tilde{r}_i + p_i)}{\partial\nu_i^h} - \sigma_o \frac{\partial\tilde{r}_o}{\partial\nu_i^h} \right\|_{0,\partial\Omega_i^h} \|e^h\|_{0,\partial\Omega_i^h} \\ \leq ch^{k-1} \left(\sum_{i=0}^2 \|r_i\|_{4,\Omega_i} + \sum_{i=1}^2 \|p_i\|_{4,\Omega_o} \right) \|e^h\|_{1,\Omega} \\ \leq ch^{k-1} \left(\sum_{i=0}^2 \|r_i\|_{4,\Omega_i}^2 + \sum_{i=1}^2 \|p_i\|_{4,\Omega_o}^2 \right)^{\frac{1}{2}} \|e^h\|_{1,\Omega}. \end{aligned} \quad (3.4.21)$$

Thus combining these last four results with (3.4.15) gives:

$$\left(\sum_{i=0}^2 \|e^h\|_{1,\Omega_i^h}^2 \right)^{\frac{1}{2}} \leq ch^{k-1} \left(\sum_{i=0}^2 \|r_i\|_{k+2,\Omega_i}^2 + \sum_{i=1}^2 \|p_i\|_{4,\Omega_o}^2 \right)^{\frac{1}{2}} \quad (3.4.22)$$

which furnishes, with (3.4.10), the required bound (3.4.3). $\square$

3.5 The L_2 Error Estimate

Having now set up the main theory required for the error estimates in the previous section, the relevant bound in the L_2 norm can be obtained by an application of the Aubin-Nitsche trick.

Theorem 3.5.1

$$\left(\sum_{i=0}^2 \|\tilde{r}_i - r_i^h\|_{0,\Omega_i^h}^2 \right)^{\frac{1}{2}} \leq ch^k. \quad (3.5.1)$$

Proof:

Let e be the error with the following restrictions to the subdomains:

$$e \equiv e_i = \tilde{r}_i - r_i^h \quad \text{in } \Omega_i^h \quad i = 0 \rightarrow 2 \quad (3.5.2)$$

and set

$$\beta = \frac{1}{\underline{m}(\Omega)} \int_{\Omega} e. \quad (3.5.3)$$

Also let $z \in H^1(\Omega)$ such that $z \equiv z_i$ in Ω_i ($i = 0 \rightarrow 2$) satisfy

$$\begin{aligned} -\nabla \cdot (\sigma \nabla z) &= e - \beta \quad \text{in } \Omega \\ \text{and} \quad \frac{\partial z}{\partial \nu} &= 0 \quad \text{on } \partial\Omega \quad \text{with} \quad \int_{\Omega} z = 0 \end{aligned} \quad (3.5.4)$$

where ν is the unit outward normal on $\partial\Omega$; i.e. z solves an equation essentially equivalent to (3.1.1) with the singular right hand side replaced by a $e - \beta$. As $\int_{\Omega} (e - \beta) = 0$ this is a consistent problem for z and the uniqueness of z being ensured by the integral condition in (3.5.4). Also from elliptic regularity with $e - \beta \in L_2(\Omega)$, $z_i \in H^2(\Omega_i)$ ($i = 0 \rightarrow 2$) and in the following piecewise H^2 norm may be bounded thus:

$$\begin{aligned} \left(\sum_{i=0}^2 \|z_i\|_{2,\Omega_i}^2 \right)^{\frac{1}{2}} &\leq c \left(\sum_{i=0}^2 \|e - \beta\|_{0,\Omega_i}^2 \right)^{\frac{1}{2}} \\ &\leq c \left(\sum_{i=0}^2 \|\tilde{r}_i - r_i^h\|_{0,\Omega_i^h}^2 \right)^{\frac{1}{2}} \\ &\leq c \|e\|_{0,\Omega}. \end{aligned} \quad (3.5.5)$$

In terms of the above problem the L_2 error may now be written in the following form, noting the use of the extension operators E_i defined in (3.3.6):

$$\|e\|_{0,\Omega}^2 = (e, e)_{\Omega}$$

$$\begin{aligned}
&= \sum_{i=0}^2 \left[(e_i, e_i)_{\Omega_i^h} - (e_i, \tilde{A}_i \tilde{z}_i)_{\Omega_i^h} \right] + \sum_{i=0}^2 (e_i, \tilde{A}_i \tilde{z}_i)_{\Omega_i^h} \\
&= \sum_{i=0}^2 (e_i, e_i - A_i z_i)_{\Omega_i^h \cap \Omega_i} + \sum_{i=0}^2 (e_i, e_i - \tilde{A}_i \tilde{z}_i)_{\Omega_i^h \setminus \Omega_i} \\
&\quad + \sum_{i=0}^2 (e_i, \tilde{A}_i \tilde{z}_i)_{\Omega_i^h}. \tag{3.5.6}
\end{aligned}$$

Each of these terms, or quantities derived from them, will now be bounded by $ch^k \|e\|_{0,\Omega}$.

The first term in (3.5.6) may initially be bounded thus:

$$\begin{aligned}
\sum_{i=0}^2 \left| (e_i, e_i - A_i z_i)_{\Omega_i^h \cap \Omega_i} \right| &\leq c \sum_{i=0}^2 \|e_i\|_{0,\Omega_i^h} \left| \int_{\Omega_i^h \cap \Omega_i} r_i - r_i^h \right| \tag{3.5.7} \\
&\leq c \left(\sum_{i=0}^2 \left| \int_{\Omega_i \setminus \Omega_i^h} r_i - \int_{\Omega_i^h \setminus \Omega_i} r_i^h \right| \right) \|e\|_{0,\Omega}
\end{aligned}$$

via (3.5.3–4) and noting the integral conditions (3.2.9) and (3.3.7) on r_i and r_i^h respectively, and then:

$$\begin{aligned}
\sum_{i=0}^2 \left| \int_{\Omega_i \setminus \Omega_i^h} r_i - \int_{\Omega_i^h \setminus \Omega_i} r_i^h \right| &\leq \sum_{i=0}^2 \left\{ \left[m(\Omega_i \setminus \Omega_i^h) \right]^{\frac{1}{2}} \|r_i\|_{0,\Omega_i \setminus \Omega_i^h} + \right. \\
&\quad \left. \left[m(\Omega_i^h \setminus \Omega_i) \right]^{\frac{1}{2}} \|r_i^h\|_{0,\Omega_i^h \setminus \Omega_i} \right\}
\end{aligned}$$

which via (3.3.4) with a trace inequality

$$\leq ch^k \sum_{i=0}^2 \left(\|r_i\|_{1,\Omega_i} + \|r_i^h\|_{1,\Omega_i^h} \right) \tag{3.5.8}$$

and with (3.4.3) the result follows. The second term is then bounded thus:

$$\begin{aligned}
\sum_{i=0}^2 \left| (e_i, e_i - \tilde{A}_i \tilde{z}_i)_{\Omega_i^h \setminus \Omega_i} \right| &\leq c \sum_{i=0}^2 \|e_i\|_{0,\Omega_i^h \setminus \Omega_i} \|e_i\|_{0,\Omega_i^h} \\
&\leq ch^{\frac{k}{2}} \left(\sum_{i=0}^2 \|e_i\|_{1,\Omega_i^h} \right)^{\frac{1}{2}} \left(\sum_{i=0}^2 \|e_i\|_{0,\Omega_i^h}^2 \right)^{\frac{1}{2}} \\
&\leq ch^k \left(\sum_{i=0}^2 \|r_i\|_{k+2,\Omega_i}^2 + \sum_{i=1}^2 \|p_i\|_{4,\Omega_o}^2 \right)^{\frac{1}{2}} \|e\|_{0,\Omega} \tag{3.5.9}
\end{aligned}$$

using (3.3.4a) and a trace inequality, and then (3.4.3).

This now leaves the final sum of (3.5.6) which has to be further decomposed which using (3.3.5b) gives:

$$\begin{aligned}
\sum_{i=0}^2 (e_i, \tilde{A}_i \tilde{z}_i)_{\Omega_i^h} &= \sum_{i=0}^2 a_i^h(e_i, \tilde{z}_i) + \\
&\quad \sum_{i=1}^2 \int_{\partial\Omega_i^h} \sigma_o \frac{\partial \tilde{z}_o}{\partial \nu_i^h} (\tilde{r}_o - r_i^o) - \sigma_i \frac{\partial \tilde{z}_i}{\partial \nu_i^h} (\tilde{r}_i - r_i^h) \\
&= \sum_{i=0}^2 a_i^h(e_i, \tilde{z}_i - \Pi_h \tilde{z}_i) + \sum_{i=0}^2 a_i^h(e_i, \Pi_h \tilde{z}_i) \\
&\quad + \sum_{i=1}^2 \int_{\partial\Omega_i^h} \sigma_i \frac{\partial p_i}{\partial \nu_i^h} \Pi_h \tilde{z}_i \\
&\quad + \sum_{i=1}^2 \int_{\partial\Omega_i^h} \sigma_o \frac{\partial \tilde{z}_o}{\partial \nu_i^h} (\tilde{r}_o - r_o^h) - \sigma_i \frac{\partial \tilde{z}_i}{\partial \nu_i^h} (\tilde{r}_i - r_i^h) \\
&= \sum_{i=0}^2 a_i^h(e_i^h, \tilde{z}_i - \Pi_h \tilde{z}_i) \\
&\quad + \sum_{i=0}^2 a_i^h(\tilde{r}_i, \Pi_h \tilde{z}_i) - \sum_{i=0}^2 (\tilde{A}_i \tilde{r}_i, \Pi_h \tilde{z}_i)_{\Omega_i^h} \\
&\quad + \sum_{i=1}^2 \int_{\partial\Omega_i^h} \sigma_i \frac{\partial p_i}{\partial \nu_i^h} \Pi_h \tilde{z}_i \\
&\quad + \sum_{i=0}^2 (\tilde{A}_i \tilde{r}_i, \Pi_h \tilde{z}_i)_{\Omega_i^h} \\
&\quad + \sum_{i=1}^2 \int_{\partial\Omega_i^h} \left(\sigma_o \frac{\partial \tilde{z}_o}{\partial \nu_i^h} - \sigma_i \frac{\partial \tilde{z}_i}{\partial \nu_i^h} \right) (\tilde{r}_i - r_i^h) \\
&\quad + \sum_{i=1}^2 \int_{\partial\Omega_i^h} \sigma_o \frac{\partial \tilde{z}_o}{\partial \nu_i^h} [(\tilde{r}_o - r_o^h) - (\tilde{r}_i - r_i^h)] \\
&= \sum_{j=1}^5 I_j. \tag{3.5.10}
\end{aligned}$$

Using (3.3.2), (3.4.3) and (3.5.5) the first bound follows directly as:

$$\begin{aligned}
|I_1| &\leq \sum_{i=0}^2 \|e_i\|_{1,\Omega_i^h} \|\tilde{z}_i - \Pi_h \tilde{z}_i\|_{1,\Omega_i^h} \\
&\leq ch \sum_{i=0}^2 \|e_i\|_{1,\Omega_i^h} \|\tilde{z}_i\|_{2,\Omega_i^h} \\
&\leq ch^k \left(\sum_{i=0}^2 \|r_i\|_{k+2,\Omega_i}^2 + \sum_{i=1}^2 \|p_i\|_{4,\Omega_o}^2 \right)^{\frac{1}{2}} \|e\|_{0,\Omega}. \quad (3.5.11)
\end{aligned}$$

The second term I_2 must be split further and using (3.2.9b) then

$$\begin{aligned}
&\sum_{i=0}^2 a_i^h(\tilde{r}_i, \Pi_h \tilde{z}_i) - \sum_{i=0}^2 (\tilde{A}_i \tilde{r}_i, \Pi_h \tilde{z}_i)_{\Omega_i^h} + \sum_{i=1}^2 \int_{\partial\Omega_i^h} \sigma_i \frac{\partial p_i}{\partial \nu_i^h} \Pi_h \tilde{z}_i \\
&\leq \left| \sum_{i=0}^2 \left[a_i^h(\tilde{r}_i, \Pi_h \tilde{z}_i - \tilde{z}_i) - (\tilde{A}_i \tilde{r}_i, \Pi_h \tilde{z}_i - \tilde{z}_i)_{\Omega_i^h} \right] \right| \\
&\quad + \sum_{i=1}^2 \left| \int_{\partial\Omega_i^h} \sigma_i \frac{\partial p_i}{\partial \nu_i^h} (\Pi_h \tilde{z}_i - \tilde{z}_i) \right| \\
&\quad + \sum_{i=0}^2 \left| a_i^h(\tilde{r}_i, \tilde{z}_i) - a_i(r_i, z_i) \right| + \sum_{i=0}^2 \left| (\tilde{A}_i \tilde{r}_i, \tilde{z}_i)_{\Omega_i^h} \right| \\
&\quad + \sum_{i=1}^2 \left| \int_{\partial\Omega_i^h} \sigma_i \frac{\partial p_i}{\partial \nu_i^h} \tilde{z}_i - \int_{\partial\Omega_i} \sigma_i \frac{\partial p_i}{\partial \nu_i} z_i \right|
\end{aligned}$$

and with (3.3.5b):

$$\begin{aligned}
&\leq \sum_{i=1}^2 \left| \int_{\partial\Omega_i^h} \left(\sigma_o \frac{\partial \tilde{r}_o}{\partial \nu_i^h} - \sigma_i \frac{\partial (\tilde{r}_i + p_i)}{\partial \nu_i^h} \right) (\Pi_h \tilde{z}_i - \tilde{z}_i) \right| \\
&\quad + \sum_{i=0}^2 \left| a_i^h(\tilde{r}_i, \tilde{z}_i) - a_i(r_i, z_i) \right| + \sum_{i=0}^2 \left| (\tilde{A}_i \tilde{r}_i, \tilde{z}_i)_{\Omega_i^h} \right| \\
&\quad + \sum_{i=1}^2 \left| \int_{\partial\Omega_i^h} \sigma_i \frac{\partial p_i}{\partial \nu_i^h} \tilde{z}_i - \int_{\partial\Omega_i} \sigma_i \frac{\partial p_i}{\partial \nu_i} z_i \right| \\
&= \sum_{j=0}^4 |S_j|. \quad (3.5.12)
\end{aligned}$$

Now:

$$\begin{aligned}
|S_1| &\leq \sum_{i=1}^2 \left\| \sigma_o \frac{\partial \tilde{r}_o}{\partial \nu_i^h} - \sigma_i \frac{\partial(\tilde{r}_i + p_i)}{\partial \nu_i^h} \right\|_{0, \partial \Omega_i^h} \|\Pi_h \tilde{z}_i - \tilde{z}_i\|_{0, \partial \Omega_i^h} \\
&\leq ch^k \left(\sum_{i=0}^2 \|r_i\|_{k+2, \Omega_i}^2 + \sum_{i=1}^2 \|p_i\|_{4, \Omega_o}^2 \right)^{\frac{1}{2}} \left(\sum_{i=0}^2 \|\tilde{z}_i\|_{2, \Omega_i^h}^2 \right)^{\frac{1}{2}} \\
&\leq ch^k \left(\sum_{i=0}^2 \|r_i\|_{k+2, \Omega_i}^2 + \sum_{i=1}^2 \|p_i\|_{4, \Omega_o}^2 \right)^{\frac{1}{2}} \|e\|_{0, \Omega} \quad (3.5.13)
\end{aligned}$$

from Lemma 3.4.1, a trace inequality, (3.3.2) and (3.5.5). Also

$$\begin{aligned}
|S_2| &\leq \sum_{i=0}^2 \left| (\sigma_i \nabla \tilde{r}_i, \nabla \tilde{z}_i)_{\Omega_i^h \setminus \Omega_i} - (\sigma_i \nabla r_i, \nabla z_i)_{\Omega_i \setminus \Omega_i^h} \right| \\
&\leq c \sum_{i=0}^2 \left[\|\tilde{r}_i\|_{1, \Omega_i^h \setminus \Omega_i} \|\tilde{z}_i\|_{1, \Omega_i^h \setminus \Omega_i} + \|r_i\|_{1, \Omega_i \setminus \Omega_i^h} \|z_i\|_{1, \Omega_i \setminus \Omega_i^h} \right]
\end{aligned}$$

and then with (3.3.4) (setting $u = r_i, \tilde{r}_i, z_i, \tilde{z}_i$ and their first derivatives) and (3.5.5) with a trace inequality:

$$\begin{aligned}
&\leq ch^k \sum_{i=0}^2 \|r_i\|_{2, \Omega_i} \|z_i\|_{2, \Omega_i} \\
&\leq ch^k \left(\sum_{i=0}^2 \|r_i\|_{k+2, \Omega_i}^2 \right)^{\frac{1}{2}} \|e\|_{0, \Omega}. \quad (3.5.14)
\end{aligned}$$

From Lemma 3.4.2 S_3 is bounded directly on application of (3.5.5) i.e.

$$\begin{aligned}
\sum_{i=0}^2 \left| (\tilde{A}_i \tilde{r}_i, \tilde{z}_i)_{\Omega_i^h} \right| &\leq ch^k \sum_{i=0}^2 \|r_i\|_{3, \Omega_i} \|\tilde{z}_i\|_{1, \Omega_i^h} \\
&\leq ch^k \left(\sum_{i=0}^2 \|r_i\|_{k+2, \Omega_i}^2 \right)^{\frac{1}{2}} \|e\|_{0, \Omega}. \quad (3.5.15)
\end{aligned}$$

Finally to bound S_4 note that using the definition of the singular functions p_i :

$$\sum_{i=1}^2 \int_{\partial \Omega_i^h} \sigma_i \frac{\partial p_i}{\partial \nu_i^h} \tilde{z}_i - \int_{\partial \Omega_i} \sigma_i \frac{\partial p_i}{\partial \nu_i} z_i =$$

$$\begin{aligned}
&= \sum_{i=1}^2 \left[\int_{\Omega_i^h \setminus \Omega_i} \sigma_i \nabla^2 p_i \tilde{z}_i - \int_{\Omega_i \setminus \Omega_i^h} \sigma_i \nabla^2 p_i z_i \right. \\
&\quad \left. - \int_{\Omega_i \setminus \Omega_i^h} \sigma_i \nabla p_i \nabla z_i + \int_{\Omega_i^h \setminus \Omega_i} \sigma_i \nabla p_i \nabla \tilde{z}_i \right] \\
&\leq \sum_{i=1}^2 \left[\left| \int_{\Omega_i \setminus \Omega_i^h} \sigma_i \nabla p_i \nabla z_i \right| + \left| \int_{\Omega_i^h \setminus \Omega_i} \sigma_i \nabla p_i \nabla \tilde{z}_i \right| \right]
\end{aligned}$$

and thus using (3.4.4)

$$\begin{aligned}
|S_4| &\leq ch^k \sum_{i=1}^2 \left[(|\nabla p_i|_{1,\Omega_i \setminus \Omega_i^h} + \|\nabla p_i\|_{0,\partial\Omega_i}) \|z_i\|_{2,\Omega_i} \right. \\
&\quad \left. + (|\nabla p_i|_{1,\Omega_i^h \setminus \Omega_i} + \|\nabla p_i\|_{0,\partial\Omega_i^h}) \|\tilde{z}_i\|_{2,\Omega_i^h} \right]
\end{aligned}$$

which noting the regularity of p_i away from x_i and applying (3.5.5) allows:

$$|S_4| \leq ch^k \|e\|_{0,\Omega} \quad (3.5.16)$$

and thus completes the bound on $|I_2|$.

I_3 is bounded, as S_3 above, directly on application of Lemma 3.4.2 and (3.5.5) thus:

$$\begin{aligned}
|I_3| &\leq \sum_{i=0}^2 \left| (\tilde{A}_i \tilde{r}_i, \Pi_h \tilde{z}_i)_{\Omega_i^h} \right| \leq ch^k \sum_{i=0}^2 \|r_i\|_{3,\Omega_i} \|\Pi_h \tilde{z}_i\|_{1,\Omega_i^h} \\
&\leq ch^k \left(\sum_{i=0}^2 \|r_i\|_{k+2,\Omega_i}^2 \right)^{\frac{1}{2}} \|e\|_{0,\Omega}. \quad (3.5.17)
\end{aligned}$$

To deal with I_4 note that:

$$\begin{aligned}
|I_4| &\leq \sum_{i=1}^2 \left\| \sigma_o \frac{\partial \tilde{z}_o}{\partial \nu_i^h} - \sigma_i \frac{\partial \tilde{z}_i}{\partial \nu_i^h} \right\|_{0,\partial\Omega_i^h} \|\tilde{r}_i - r_i^h\|_{0,\partial\Omega_i^h} \\
&\leq ch^{\frac{k}{2}} \sum_{i=1}^2 (\|z_i\|_{2,\Omega_i} + \|z_o\|_{2,\Omega_o}) \|\tilde{r}_i - r_i^h\|_{1,\Omega_i^h}
\end{aligned}$$

where the first norm bound follows directly from Lemma 2.1 of [Barrett and Elliott 1985] (i.e. similar to Lemma 3.4.1), and then using

(3.4.3) and (3.5.5) gives:

$$|I_4| \leq ch^k \left(\sum_{i=0}^2 \|r_i\|_{k+2,\Omega_i}^2 + \sum_{i=1}^2 \|p_i\|_{4,\Omega_o}^2 \right)^{\frac{1}{2}} \|e\|_{0,\Omega}. \quad (3.5.18)$$

Finally to complete the bound on (3.5.10), using (3.3.8) gives

$$\begin{aligned} I_5 &= \sum_{i=1}^2 \int_{\partial\Omega_i^h} \sigma_o \frac{\partial \tilde{z}_o}{\partial \nu_i^h} (\tilde{r}_o - \tilde{r}_i - \Pi_h p_i) \\ &= \sum_{i=1}^2 \int_{\partial\Omega_i^h} \sigma_o \frac{\partial \tilde{z}_o}{\partial \nu_i^h} [(\tilde{r}_o - \Pi_h \tilde{r}_o) - (\tilde{r}_i - \Pi_h \tilde{r}_i)] \end{aligned}$$

and thus using (3.3.2), (3.5.5) and a trace inequality

$$\begin{aligned} |I_5| &\leq c \sum_{i=1}^2 \|\tilde{z}_o\|_{2,\Omega_o^h} \left(\|\tilde{r}_o - \Pi_h \tilde{r}_o\|_{0,\partial\Omega_i^h} + \|\tilde{r}_i - \Pi_h \tilde{r}_i\|_{0,\partial\Omega_i^h} \right) \\ &\leq ch^k \|\tilde{z}_o\|_{2,\Omega_o^h} \sum_{i=1}^2 \left(\|\tilde{r}_o\|_{k+1,\partial\Omega_i^h} + \|\tilde{r}_i\|_{k+1,\partial\Omega_i^h} \right) \\ &\leq ch^k \left(\sum_{i=0}^2 \|r_i\|_{k+2,\Omega_i}^2 \right)^{\frac{1}{2}} \|e\|_{0,\Omega} \end{aligned} \quad (3.5.19)$$

Thus combining (3.5.10–19) bounds the third term in (3.5.6), which with (3.5.7–9) supplies the required result (3.5.1). $\square$

3.6 Local Subtraction in the General Case

In the introduction it was explained that the test problem to which local subtraction was applied was a special case chosen to show the technique's full power. However in general σ will be variable and not necessarily discontinuous which potentially will substantially affect the method's performance and application. In the former case the effect is that regularity of r_i $\{i = 1, 2\}$ will be limited as described in section 2.3 and this will put an upper bound on the approximation power achievable. The effect on application of the latter though is advantageous as the natural restriction of aligning subtraction region boundaries with discontinuity lines will no longer be necessary

allowing a free choice in their positioning. The result of these observations will now be described.

There are two main effects of having σ non-constant within subdomains of Ω . The first of these differences is in the subtracted function's regularity i.e. compare that derived for p_r in section 2.3 and $\{r_i\}_{i=0}^2$ in section 3.2. In subtraction from a general variable σ problem the residual function will have a bound on its regularity i.e. will be in $H^{2-\epsilon}(\Omega)$ as opposed to when σ is constant when regularity is determined by domain smoothness. This limits the potential of a method to accurately mimic the solution's behaviour. The second difference can be seen by comparing (2.4.6) with (3.2.7) and shows that with variable σ a right hand side integral over Ω is present and this can be difficult to evaluate to sufficient accuracy.

The important result for deriving error bounds when using variable σ is that for $w \in H^{2-\epsilon}(\Omega)$ or equivalently $w \in W^{2,2-\epsilon}(\Omega)$ the interpolant error is given by:

$$|w - \Pi_h w|_{1,\Omega} \leq ch^{1-\epsilon}, \quad (3.6.1)$$

which may be obtained from Ciarlet [Ciarlet 1978] (see (5.2.14)). Thus when analysing the H^1 error obtained from a local subtraction approximation of r_i defined by (2.4.6) this limit on the interpolants approximation power will manifest itself in the associated error terms derived in proving an analogy to Theorem 3.4.3 for the variable σ case. The effect on the overall H^1 error will then be to limit approximation power to this interpolant error. The interpolant errors become the dominant terms when bounding $\|\tilde{r}_i - r_i^h\|_{1,\Omega_i^h}$ as all other terms are due to domain perturbation from isoparametric fitting. This occurs away from any singularities where solution regularity and so associated approximation power is not restricted by their influence. The H^1 error bound thus becomes:

$$\left(\sum_{i=0}^2 \|\tilde{r}_i - r_i^h\|_{1,\Omega_i^h}^2 \right)^{\frac{1}{2}} \leq ch^{1-\epsilon}. \quad (3.6.2)$$

To obtain an L_2 bound a similar application of the Aubin-Nitsche trick as in section 3.5 will give the same one power of h improvement in approximation power, again singularities present will not affect the domain perturbation

error terms, and thus the bound will be:

$$\left(\sum_{i=0}^2 \|\tilde{r}_i - r_i^h\|_{0,\Omega_i^h}^2 \right)^{\frac{1}{2}} \leq ch^{2-\epsilon}. \quad (3.6.3)$$

The second result of a generalisation of σ is that if continuous the domain of dependence of the local singularity now becomes effectively arbitrary. The only restriction results from the need to have subdomain boundaries an $O(1)$ distance from the singularities thus ensuring sufficient regularity of relevant functions on and in the immediate neighbourhood of these boundaries. This precludes the idea of reducing a subtraction region's area as the mesh is refined but is only a minimal constraint on the method. Apart from this if the subtraction domains are made as small as possible then this reduces the computational effort required in obtaining finite element approximations.

Chapter 4

Fracture Problems and Domain Induced Singularities

4.1 Introduction

Before investigating the fracture problem itself a related situation is looked at where the (constant) permeability around a well, as considered in Chapter 3, tends to infinity. This limiting process gives rise to a problem similar to that involving a fracture, the well-region becomes a flux source with infinite conductivity and thus a zero pressure gradient, i.e. the conditions that exist within a well or fracture source/sink. The inner region pressure is then an unknown constant and for consistency the total flux around the boundary of the inner region is equal to the well's injection rate. As will be shown below, along with relevant external boundary data these conditions give sufficient information to solve for the pressure. It is in fact the solution of a total flux problem as analysed by Barrett and Elliott [Barrett and Elliott 1987].

The chapter begins by formulating the infinite permeability well-region problem which is then used to motivate the modelling of a fracture by a total flux problem. A simple idealized problem is then directly obtained in section 4.2 and the approximation of this is discussed in section 4.3. The chapter continues in section 4.4 with a description of domain induced singularities

resultant from the fracture model and gives a singular expansion for the solution, thus introducing stress intensity factors (the coefficients in this expansion). Various approximation techniques are then discussed for this problem including the dual singular function method (DSFM). Finally in section 4.5 the DSFM is described in terms of its practical implementation and the basis of an alternative technique is presented.

4.2 The Pressure Problem with an Infinite Permeability Well Region

The purpose of this section is to show that this infinite permeability problem may be described in a total flux formulation. To do this a simplified version of (3.1.1) is considered with σ piecewise constant and arbitrarily large in the well region. This problem and the total flux formulation are now described and then the discrepancy in their solutions is bounded in terms of the well-region's vanishing conductivity (i.e. inverse of permeability). The required result then follows.

The vanishing conductivity problem formulated, with reference to Fig. 4.1, is a simplified version of (3.1.1) hence:

$$\begin{aligned} -\nabla \cdot (\sigma \nabla p^\epsilon) &= q\delta(x - x_n) = q\delta_n & \sigma &= \frac{1}{\epsilon} \text{ in } \Omega_n \\ & & \sigma &= 1 \text{ in } \Omega_o \\ p^\epsilon &= 0 & & \text{ on } \partial\Omega \end{aligned} \quad (4.2.1)$$

where $\bar{\Omega} = \bar{\Omega}_n \cup \bar{\Omega}_o$, $\bar{\Gamma}_n = \bar{\Omega}_n \cap \bar{\Omega}_o$ and $x_n \in \Omega_n$. The weak form of this problem is to find $p^\epsilon \in W_E^{1,1}(\Omega)$ ($p^\epsilon \equiv p_n^\epsilon$ on Ω_n and $p^\epsilon \equiv p_o^\epsilon$ on Ω_o) such that

$$\begin{aligned} (\sigma \nabla p^\epsilon, \nabla v)_\Omega &\equiv \left(\frac{1}{\epsilon} \nabla p_n^\epsilon, \nabla v \right)_{\Omega_n} + (\nabla p_o^\epsilon, \nabla v)_{\Omega_o} = (q\delta_n, v)_\Omega \\ &\forall v \in W_E^{1,\infty}(\Omega) \end{aligned} \quad (4.2.2)$$

where $W_E^{1,1}(\Omega) \equiv \{v | v \in W^{1,1}(\Omega), v = 0 \text{ on } \partial\Omega\}$ with an analogous definition for $W_E^{1,\infty}(\Omega)$. Alternatively each region comprising Ω can be treated separately and noting that Ω_o contains no source singularities this gives rise

to the following description: find $p_n^\epsilon \in W^{1,1}(\Omega_n)$, $p_o^\epsilon \in H_E^1(\Omega_o)$ such that $p_n^\epsilon = p_o^\epsilon$ on Γ_n and

$$\begin{aligned} \left(\frac{1}{\epsilon} \nabla p_n^\epsilon, \nabla v \right)_{\Omega_n} &= (q \delta_n, v)_{\Omega_n} + \left\langle \frac{1}{\epsilon} \frac{\partial p_n^\epsilon}{\partial \nu_n}, v \right\rangle_{\Gamma_n} & \forall v \in W^{1,\infty}(\Omega_n) \\ (\nabla p_o^\epsilon, \nabla v)_{\Omega_o} &= - \left\langle \frac{\partial p_o^\epsilon}{\partial \nu_n}, v \right\rangle_{\Gamma_n} & \forall v \in H_E^1(\Omega_o) \end{aligned} \quad (4.2.3)$$

where $H_E^1(\Omega_o) \equiv \{v | v \in H^1(\Omega_o), v = 0 \text{ on } \partial\Omega\}$ and ν_n is a unit outward normal to Ω_n . The total flux formulation, which is shown to be the limit of (4.2.1) as $\epsilon \rightarrow 0$ is:

$$\begin{aligned} -\nabla^2 p &= 0 & \text{in } \Omega_o \\ p &= \lambda & \text{(unknown constant) on } \Gamma_n \\ \int_{\Gamma_n} \frac{\partial p}{\partial \nu_n} &= -q & \text{and } p = 0 \text{ on } \partial\Omega. \end{aligned} \quad (4.2.4)$$

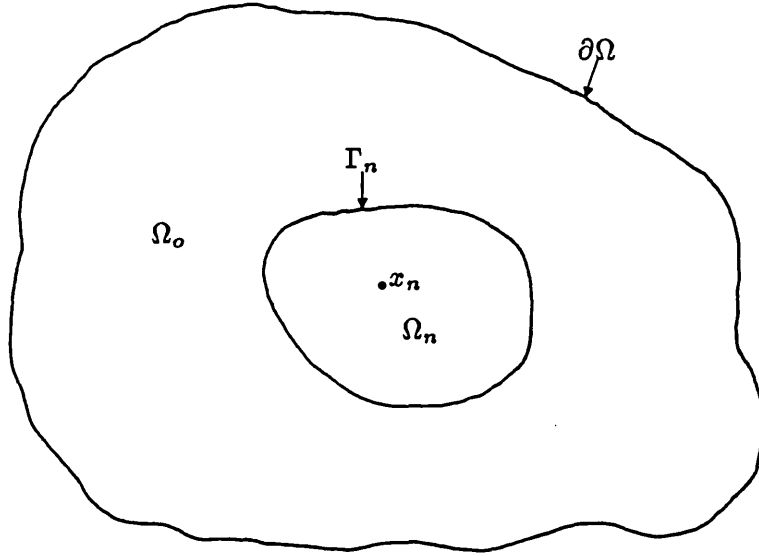


Figure 4.1: Reservoir section with vanishing conductivity region

To show the equivalence of the two problems in the limit, an ϵ bound on the error $\|p - p_o^\epsilon\|_{0,\Omega_o}$ will be demonstrated, this following from a bound on

$|p - p_o^\epsilon|_{1, \Omega_o}$. This seminorm bound can be obtained from (4.2.3) and (4.2.4) thus:

$$(\nabla p_o^\epsilon, \nabla v)_{\Omega_o} = - \left\langle \frac{\partial p_o^\epsilon}{\partial \nu_n}, v \right\rangle_{\Gamma_n} \quad \forall v \in H_E^1(\Omega_o)$$

from (4.2.3) and by performing a similar integration on (4.2.4),

$$(\nabla p, \nabla v)_{\Omega_o} = - \left\langle \frac{\partial p}{\partial \nu_n}, v \right\rangle_{\Gamma_n} \quad \forall v \in H_E^1(\Omega_o)$$

which on subtraction of the previous equation gives:

$$(\nabla(p - p_o^\epsilon), \nabla v)_{\Omega_o} = - \int_{\Gamma_n} \frac{\partial(p - p_o^\epsilon)}{\partial \nu_n} v \quad \forall v \in H_E^1(\Omega_o). \quad (4.2.5)$$

So if e is defined as $e = p - p_o^\epsilon$ and then v set equal to e in (4.2.5),

$$|e|_{1, \Omega_o}^2 = - \int_{\Gamma_n} \frac{\partial e}{\partial \nu_n} (p - p_o^\epsilon). \quad (4.2.6)$$

Now from (4.2.1) and (4.2.4)

$$\int_{\Gamma_n} \frac{\partial p_o^\epsilon}{\partial \nu_n} = -q = \int_{\Gamma_n} \frac{\partial p}{\partial \nu_n}$$

and as $p = \lambda$ on Γ_n then (4.2.6) may be rewritten:

$$|e|_{1, \Omega_o}^2 = - \int_{\Gamma_n} \frac{\partial e}{\partial \nu_n} (\mu - p_o^\epsilon) \quad \forall \mu \in R \quad (4.2.7)$$

and thus:

$$|e|_{1, \Omega_o}^2 \leq \left\| \frac{\partial e}{\partial \nu_n} \right\|_{-\frac{1}{2}, \Gamma_n} \|\mu - p_o^\epsilon\|_{\frac{1}{2}, \Gamma_n}. \quad (4.2.8)$$

Finally, again from (4.2.1) and (4.2.4), $-\nabla^2 e = 0$ in Ω_o and $e = 0$ on $\partial\Omega$ and then it follows from a trace inequality and Friedrich's inequality that

$$\left\| \frac{\partial e}{\partial \nu_n} \right\|_{-\frac{1}{2}, \Gamma_n} \leq c |e|_{1, \Omega_o} \quad (4.2.9)$$

(see Babuska and Aziz [Babuška and Aziz 1972]), and this combined with (4.2.8) then gives, noting that $p_n^\epsilon = p_o^\epsilon$ on Γ_n :

$$|e|_{1, \Omega_o} \leq c \|p_n^\epsilon - \mu\|_{\frac{1}{2}, \Gamma_n} \quad \forall \mu \in R. \quad (4.2.10)$$

Now the inner solution p_n^ϵ can be broken down into the sum of a known singular fundamental solution p_s^ϵ and a smooth function p_r^ϵ , which, following (4.2.3), satisfy:

$$\begin{aligned} \left(\frac{1}{\epsilon} \nabla p_s^\epsilon, \nabla v \right)_{\Omega_n} &= (q \delta_n, v)_{\Omega_n} + \left\langle \frac{1}{\epsilon} \frac{\partial p_s^\epsilon}{\partial \nu_n}, v \right\rangle_{\Gamma_n} \quad \forall v \in W^{1,\infty}(\Omega_n) \\ \left(\frac{1}{\epsilon} \nabla p_r^\epsilon, \nabla v \right)_{\Omega_n} &= \left\langle \frac{1}{\epsilon} \frac{\partial}{\partial \nu_n} (p_n^\epsilon - p_s^\epsilon), v \right\rangle_{\Gamma_n} \quad \forall v \in H^1(\Omega_n) \end{aligned} \quad (4.2.11)$$

where $p_s^\epsilon = (q/2\pi)\epsilon \log |x - x_n|^{-1}$ and $p_n^\epsilon = p_s^\epsilon + p_r^\epsilon$ in Ω_n , $\frac{\partial p_n^\epsilon}{\partial \nu_n} = \epsilon \frac{\partial p_s^\epsilon}{\partial \nu_n}$ on Γ_n . Using these functions (4.2.10) may now be written as:

$$|e|_{1,\Omega_o} \leq c \left(\epsilon \|\log |x - x_n|\|_{\frac{1}{2},\Gamma_n} + \|p_r^\epsilon - \mu\|_{\frac{1}{2},\Gamma_n} \right) \quad (4.2.12)$$

where the first normed term is bounded as $x_n \notin \Gamma_n$, thus:

$$|e|_{1,\Omega_o} \leq \epsilon K + c \|p_r^\epsilon - \mu\|_{\frac{1}{2},\Gamma_n} \quad \forall \mu \in R. \quad (4.2.13)$$

Now using a trace inequality $\|p_r^\epsilon - \mu\|_{\frac{1}{2},\Gamma_n} \leq c \|p_r^\epsilon - \mu\|_{1,\Omega_n}$ and if μ is chosen so that $\mu = \frac{1}{m(\Omega_n)} \int_{\Omega_n} p_r^\epsilon$ then $\int_{\Omega_n} (p_r^\epsilon - \mu) = 0$ and from the Poincaré inequality (3.3.11):

$$\|p_r^\epsilon - \mu\|_{1,\Omega_n} \leq c \left\{ |p_r^\epsilon|_{1,\Omega_n}^2 + \left(\int_{\Omega_n} (p_r^\epsilon - \mu)^2 \right)^{\frac{1}{2}} \right\} \leq c |p_r^\epsilon|_{1,\Omega_n}$$

and thus:

$$\|p_r^\epsilon - \mu\|_{\frac{1}{2},\Gamma_n} \leq c |p_r^\epsilon|_{1,\Omega_n}. \quad (4.2.14)$$

The seminorm in (4.2.14) can be bounded by letting $v = p_r^\epsilon - \mu$ in (4.2.11), thus

$$\begin{aligned} |p_r^\epsilon|_{1,\Omega_n}^2 &= \left\langle \frac{\partial}{\partial \nu_n} (p_n^\epsilon - p_s^\epsilon), p_r^\epsilon - \mu \right\rangle_{\Gamma_n} \\ &\leq \epsilon \left\| \frac{\partial p_s^\epsilon}{\partial \nu_n} \right\|_{-\frac{1}{2},\Gamma_n} \|p_r^\epsilon - \mu\|_{\frac{1}{2},\Gamma_n} + \epsilon \left\| \frac{\partial}{\partial \nu_n} (\log |x - x_n|) \right\|_{0,\Gamma_n} \|p_r^\epsilon - \mu\|_{0,\Gamma_n} \end{aligned}$$

and then via the trace and Poincaré inequalities:

$$|p_r^\epsilon|_{1,\Omega_n} \leq c \epsilon \left(\left\| \frac{\partial p_s^\epsilon}{\partial \nu_n} \right\|_{-\frac{1}{2},\Gamma_n} + \left\| \frac{\partial}{\partial \nu_n} (\log |x - x_n|) \right\|_{0,\Gamma_n} \right). \quad (4.2.15)$$

Now as (4.2.12), the second norm above is bounded as $x_n \notin \Gamma_n$ and as in (4.2.9)

$$\left\| \frac{\partial p_o^\epsilon}{\partial \nu_n} \right\|_{-\frac{1}{2}, \Gamma_n} \leq c |p_o^\epsilon|_{1, \Omega_o}.$$

Thus there exists a constant A such that

$$\begin{aligned} |p_r^\epsilon|_{1, \Omega_n} &\leq \epsilon (A + c |p_o^\epsilon|_{1, \Omega_o}) \\ &\leq \epsilon (A + c(|p|_{1, \Omega_o} + |e|_{1, \Omega_o})) \end{aligned} \quad (4.2.16)$$

and combining (4.2.13), (4.2.14) and (4.2.16) gives:

$$|e|_{1, \Omega_o} \leq \epsilon(k + c|e|_{1, \Omega_o}) \quad \text{where } k = K + A + c|p|_{1, \Omega_o}$$

and so

$$|e|_{1, \Omega_o} \leq C\epsilon \quad \text{for } \epsilon < \frac{1}{c}, \quad (4.2.17)$$

Finally noting $p = p_o^\epsilon$ on $\partial\Omega$ a similar bound follows directly for $\|p - p_o^\epsilon\|_{0, \Omega_o}$.

4.3 A Model Fracture Problem and Approximate Solutions

Physically a fracture can be considered as a thin region of infinite permeability, an empty hole which offers zero resistance to fluid flow. It will have a finite width but in comparison to its length and reservoir dimensions this will be negligible. The fracture therefore will appear in the model geometrically as a slit in the plane and in terms of fluid flow, as a line source. The strength of this source will be the well injectivity rate which is here equal to the fracture's total flux. Thus a fracture source/sink can be modelled as a special case of the above infinite permeability problem i.e. where the slit (well) is the internal region/boundary and its injectivity rate is the necessary prescribed total flux.

To avoid many complexities concerning the fracture problem an idealized model will be considered. The situation taken is where a single fracture is contained within a rectangular region Ω . The fracture is described by a straight slit, giving rise to an internal boundary Γ_n with reentrant corners

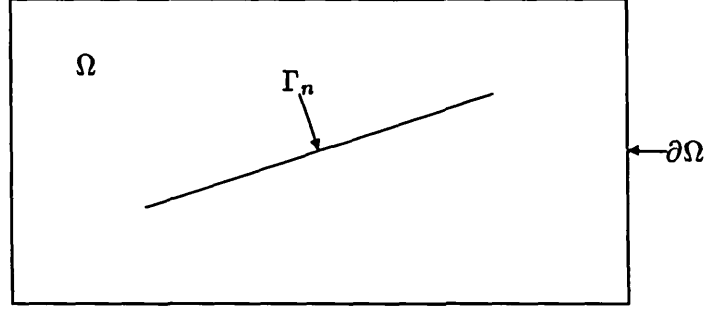


Figure 4.2: The model fracture problem domain

angle 2π at each end, injecting fluid at rate q . On the outer boundary of Ω , $\partial\Omega$, the solution u is known and thus in differential form the problem is (with reference to Fig. 4.2), given g and q find u such that:

$$\begin{aligned}
 -\nabla^2 u &= 0 & \text{in } \Omega \\
 u &= \lambda & \text{(unknown constant) on } \Gamma_n \\
 u &= g & \text{on } \partial\Omega \\
 \int_{\Gamma_n} \frac{\partial u}{\partial \nu_n} &= -q \left(= \int_{\partial\Omega} \frac{\partial u}{\partial \nu} \right).
 \end{aligned} \tag{4.3.1}$$

The associated weak form of this problem is then: find $u \in V_g(\lambda)$ such that

$$a(u, v) = \mu q \quad \forall v \in V_o(1), \forall \mu \in \mathbb{R} \tag{4.3.2}$$

where $V_g(\mu) \equiv \{w | w \in H^1(\Omega), w = g \text{ on } \partial\Omega \text{ and } w = \mu \text{ (constant) on } \Gamma_n\}$ and $a(u, v) \equiv (\nabla u, \nabla v)$. It is then convenient to decompose u into the form

$$u = u_o + \lambda u_1 \tag{4.3.3}$$

i.e. where $u_o \in V_g(0)$ and $u_1 \in V_o(1)$ are such that

$$\begin{aligned}
 a(u_o, v) &= 0 & \forall v \in V_o(0) \\
 a(u_1, v) &= 0 & \forall v \in V_o(0).
 \end{aligned} \tag{4.3.4}$$

From consistency λ must then satisfy the following equation

$$\int_{\Gamma_n} \frac{\partial u_o}{\partial \nu_n} + \lambda \int_{\Gamma_n} \frac{\partial u_1}{\partial \nu_n} = -q \quad (4.3.5a)$$

which also holds for any closed curve surrounding Γ_n , specifically $\partial\Omega$. Using (4.3.2) an equivalent formula for λ is (noting $u_1 \in V_o(1)$):

$$a(u_o, u_1) + \lambda a(u_1, u_1) = q, \quad (4.3.5b)$$

or if Γ_n in (4.3.5a) is replaced by $\partial\Omega$ it is also equivalently given by

$$a(u_o, 1 - u_1) + \lambda a(u_1, 1 - u_1) = -q \quad (4.3.5c)$$

as $1 - u_1 \in V_1(0)$. Thus λ may be recovered from u_o and u_1 as follows

$$\lambda = \frac{q - a(u_o, u_1)}{a(u_1, u_1)}. \quad (4.3.6)$$

Finally an approximation to u may be derived via (4.3.4) and (4.3.6) in a straightforward manner using standard finite elements. Using a piecewise polynomial subspace $S^h(\Omega)$ of $H^1(\Omega)$ the finite element analogue of $V_g(\mu)$, $V_g^h(\mu)$, is defined as follows $V_g^h(\mu) \equiv \{\chi | \chi \in S^h(\Omega), \chi = \Pi_h g \text{ on } \partial\Omega \text{ and } \chi = \mu \text{ on } \Gamma_n\}$, where Π_h is the interpolation operator associated with $S^h(\Omega)$. Also for $S^h(\Omega)$ to deliver well-defined approximations the associated decomposition of Ω must contain Γ_n as a mesh line (as would happen with any boundary) and with a node placed at each end. The approximations to u_o and u_1 then follow as $u_o^h \in V_g^h(0)$, $u_1^h \in V_o^h(1)$ s.t.

$$\begin{aligned} a(u_o^h, \chi) &= 0 & \forall \chi \in V_o^h(0) \\ a(u_1^h, \chi) &= 0 & \forall \chi \in V_o^h(0). \end{aligned} \quad (4.3.7)$$

Then defining λ^h to be the obvious approximation to λ from (4.3.6) i.e.

$$\lambda^h = \frac{q - a(u_o^h, u_1^h)}{a(u_1^h, u_1^h)} \quad (4.3.8)$$

an approximation to u follows directly thus:

$$u^h = u_o^h + \lambda^h u_1^h. \quad (4.3.9)$$

In fact (4.3.8) can be simplified further i.e. by noting (4.3.7) it collapses to:

$$\lambda^h = \frac{q - a(u_o^h, \eta_1^h)}{a(u_1^h, \eta_1^h)} \quad (4.3.10a)$$

where $\eta_1^h \in V_o^h(1)$ is s.t.

$$\eta_1^h = 0 \quad \text{at nodes in } \overline{\Omega} \setminus \Gamma_n.$$

Also via a similar process, using the second inner product definition for λ (4.3.5c) an equivalent approximation is given by:

$$\lambda^h = \frac{q + a(u_o^h, \eta_2^h)}{-a(u_1^h, \eta_2^h)} \quad (4.3.10b)$$

where $\eta_2^h \in V_1^h(0)$ and zero at all nodes in $\overline{\Omega} \setminus \partial\Omega$. This shows that the inner products in (4.3.10) are very easy to evaluate being just boundary skin integrals.

The reason for using a discrete form of (4.3.6) rather than (4.3.5) to derive λ^h is that, as shown by Barrett and Elliott [Barrett and Elliott 1987], the inner products give a greater power of approximation to the fluxes they represent than the equivalent boundary integrals. In fact in the case of a smooth domain assuming sufficient regularity of u_o and u_1 optimal approximations will be obtained for these fluxes and thus (from [Barrett and Elliott 1987]) for λ^h and u^h as well. However, there are singularities present in the fracture problem and these need to be investigated first.

4.4 Domain Induced Singularities and their Approximation

Detailed approximations of the fracture will now be left until Chapter 6 and in this section an important general feature of the analytic solution u and its effect on approximations will be discussed. The domain associated with the fracture model of section 4.3 contains two reentrant corners of angle 2π representing the fracture's ends and it is well known (see e.g. [Grisvard 1975])

that elliptic problems on non-convex polygonal domains give rise to solutions which have singularities induced by and centred on the reentrant corners. This will obviously give rise to problems when considering general approximations to such solutions and in fact not even finite elements with linear basis functions (i.e. the simplest case) converge at optimal rates. These problems and various methods to overcome them will be discussed at the end of this section but first some analytic preliminaries will be presented.

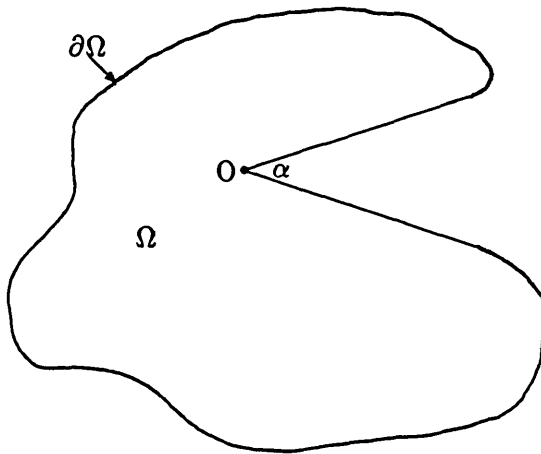


Figure 4.3: The general single reentrant corner domain

A more general situation than that directly relevant to (4.3.1) will be considered in view of approximation techniques later analysed i.e. the following Dirichlet problem:

$$\begin{aligned} -\nabla^2 u &= f & \text{in } \Omega \\ u &= g & \text{on } \partial\Omega \end{aligned} \quad (4.4.1)$$

where for ease of exposition Ω is an idealised domain smooth everywhere excepting one non-convex corner angle α , $\pi < \alpha \leq 2\pi$ centred on the origin O , see Fig. 4.3. From Grisvard [Grisvard 1975] u may be expanded in polar coordinates about the origin thus:

$$u = w + \sum_{j=1}^n k_j r^{j\beta} \sin(j\beta\theta) (\log r)^{m_j} \quad (4.4.2)$$

where:

$$\beta = \frac{\pi}{\alpha}; m_j = 1 \text{ when } j\beta \text{ integer, } 0 \text{ otherwise; } w \in H^{(n+1)\beta+1-\epsilon}(\Omega). \quad (4.4.3)$$

Here k_j (the j^{th} stress intensity factor or s.i.f.) is an unknown coefficient and the regularity of w assumes f and g are sufficiently smooth (i.e. for $w \in H^{p+2}(\Omega)$, $f \in H^p(\Omega)$ and $g \in H^{p+2}(\Omega)$) as will be done in the rest of this and the next chapter. (4.4.2–3) then demonstrate the singular behaviour in u , for example it will have the form $u = k_1 r^\beta \sin(\beta\theta) + v$, $\frac{1}{2} \leq \beta < 1$, with $v \in H^2(\Omega)$ assuming $f \in L_2(\Omega)$ and $g \in H^2(\Omega)$.

To give a complete description of the singular behaviour of u information is required on the coefficients k_j . For certain of these an explicit method of calculation (in terms of the problem's data) is available but for the remaining a direct method cannot be formulated. The former class of s.i.f.s are those associated with the log singularities and depend locally on the data f and g . These in fact may be obtained from the coefficients of a Taylor expansion of f and g about 0 (see [Grisvard 1975]). In the latter class however the k_j may be written implicitly in terms of integrals with a global dependence on f and g and involving the solution u . For ease of exposition the problem that will be dealt with below and in Chapter 5 is then a slight modification of (4.4.1) assuming that the log singularities have been explicitly calculated and the data f and g adjusted accordingly so that the singular expansion of u contains only terms involving $j\beta$ non-integer. Under this assumption from (4.4.2) and (4.4.3) u may be given by:

$$u = w + \sum_{i=1}^n k_i r^{i\beta} \sin(i\beta\theta), \quad (4.4.4)$$

where $w \in H^{(n+1)\beta+1-\epsilon}(\Omega)$. The above will be assumed throughout Chapters 4, 5 and 6.

Turning now to approximations of (4.4.1) the obvious starting point is to use the finite element method. The weak form of (4.4.1) is to find $u \in H_g^1(\Omega)$ such that

$$(\nabla u, \nabla v) = (f, v) \quad \forall v \in H_o^1(\Omega) \quad (4.4.5)$$

where $H_q^1(\Omega) \equiv \{w | w \in H^1(\Omega) \text{ and } w = q \text{ on } \partial\Omega\}$ and $(\cdot, \cdot)$ is the usual L_2 inner product over Ω . Now if $S^h(\Omega)$ is a space of piecewise polynomials defined over a partition of Ω into finite elements then the idealised finite element approximation (i.e. assuming no variational crimes, obviously unrealistic practically) of u is $u^h \in S_g^h(\Omega)$ so that:

$$(\nabla u^h, \nabla \chi) = (f, \chi) \quad \forall \chi \in S_o^h(\Omega) \subset H_o^1(\Omega) \quad (4.4.6)$$

where $S_q^h(\Omega) \equiv \{\psi | \psi \in S^h(\Omega) \text{ and } \psi = \Pi_h q \text{ on } \partial\Omega\}$ and Π_h is the interpolation operator associated with the subspace $S^h(\Omega)$. Due to the presence of the singularities in u just described this will be a poor approximation, H^1 and L_2 errors (as will be derived in section 5.2) having the following form:

$$\|u - u^h\|_{1,\Omega} \leq ch^{\beta-\epsilon} \quad (4.4.7a)$$

$$\|u - u^h\|_{0,\Omega} \leq ch^{2\beta-\epsilon} \quad (4.4.7b)$$

irrespective of the order of polynomials contained in $S^h(\Omega)$. Thus for finite elements to retain their full approximating power it is obviously necessary that some modification to the basic technique be used.

As with the source singularity problem the main strategy to overcome low approximation power is to exploit the a priori knowledge of the singular behaviour. Again the two standard approaches of mesh refinement and introducing relevant singular functions into the approximation may be used and form the basis of methods so far developed. Much work has been done on mesh refinements (see e.g. [Babuška 1970], [Babuška and Aziz 1972], [Schatz and Wahlbin 1979]) with the general idea being that using sufficient mesh grading in the vicinity of corners optimal L_2 , H^1 and L_∞ error estimates may be obtained while still requiring only $O(h^{-2})$ elements. However the grading conditions can be quite severe even for low order elements (see [Schatz and Wahlbin 1979]). This can lead to a great deal of effort being necessary to construct the required meshes.

The idea of explicitly using the singular functions as part of an approximation technique has also received much attention. The most straightforward approach is the SFM, as proposed in Babuška, Mejlík

and Vitasel [Babuška et al. 1961] and analysed by Fix, Gulati and Wakoff [Fix et al. 1973]. Here the finite element subspace is augmented by singular functions from the decomposition of u i.e. define

$$T^h(\Omega) = S^h(\Omega) \oplus_{i=1}^n \tilde{s}_i \quad (4.4.8)$$

with

$$s_i = r^{i\beta} \sin(i\beta\theta) \text{ and } \tilde{s}_i = \tau(r)s_i \quad (4.4.9)$$

where $\tau(r)$ is a smooth cutoff function having value 1 around the corner and 0 a distance $O(1)$ away from it (which minimises the loss of sparseness in the stiffness matrix when using these functions in the finite element inner products). An approximation $\bar{u}^h$ to u in (4.4.1) is then defined as $\bar{u}^h \in T^h(\Omega)$ such that $\bar{u}^h = \Pi_h g$ on $\partial\Omega$ and

$$(\nabla \bar{u}^h, \nabla \bar{\chi}) = (f, \bar{\chi}) \quad \forall \bar{\chi} \in T_o^h(\Omega) \equiv T^h(\Omega) \cap H_o^1(\Omega). \quad (4.4.10)$$

An alternative formulation for $\bar{u}^h$ may be given using the following projection operator $P_h : W^{2,q}(\Omega) \rightarrow S^h(\Omega)$ for $q \in (1, 2]$ such that $\forall v \in W^{2,q}(\Omega)$, $P_h v \in S^h(\Omega)$ satisfies

$$\begin{aligned} (\nabla(v - P_h v), \nabla \chi) &= 0 \quad \forall \chi \in S_o^h(\Omega) \\ P_h v &\equiv \Pi_h v \quad \text{on } \partial\Omega. \end{aligned} \quad (4.4.11)$$

$\bar{u}^h$ can then be written as

$$\bar{u}^h = \bar{w}^h + \sum_{i=1}^n k_i^h \tilde{s}_i \quad (4.4.12)$$

where the k_i^h are given by:

$$\sum_{i=1}^n k_i^h (\nabla(\tilde{s}_i - P_h \tilde{s}_i), \nabla \tilde{s}_j) = (f, \tilde{s}_j) - (\nabla P_h u, \nabla \tilde{s}_j) \quad j = 1 \rightarrow n \quad (4.4.13)$$

and $\bar{w}^h$ is an approximation to the regular part of u , i.e. w given in (4.4.4), defined by

$$\bar{w}^h = P_h u - \sum_{i=1}^n k_i^h P_h \tilde{s}_i. \quad (4.4.14)$$

The augmented subspace (4.4.8) is then a greatly improved approximation to the solution space and if n is chosen so that $w \in H^k(\Omega)$ with $k - 1$ order polynomials in $S^h(\Omega)$ then optimal H^1 and L_2 error approximations will be given by this method. However in respect of L_∞ or s.i.f. approximations the SFM does not perform as well (see [Dobrowolski 1981]). Optimal results are not obtained in these cases even if n can be chosen so that $w \in W^{k,\infty}(\Omega)$ and, significantly, s.i.f. estimates are not necessarily convergent. Thus for certain applications this method will also be unsuitable.

Recently a more sophisticated technique using similar ideas has been developed for approximating (4.4.1), the dual singular function method (DSFM) of Blum and Dobrowolski [Blum and Dobrowolski 1982]. This is a Petrov–Galerkin formulation where the trial space is augmented as in the SFM, providing a good approximation to the solution space, but the test space is augmented with dual singular functions s_{-i} given by:

$$s_{-i} = r^{-i\beta} \sin(i\beta\theta). \quad (4.4.15)$$

Blum and Dobrowolski in fact define their method using singular functions modified by a smooth cutoff function as in (4.4.9), i.e. $\tilde{s}_{\pm i}$, and augmented trial and test spaces T^h and U^h are then derived thus:

$$T^h(\Omega) = S^h(\Omega) \oplus_{i=1}^n \tilde{s}_i \quad U^h(\Omega) = S^h(\Omega) \oplus_{i=1}^n \tilde{s}_{-i} \quad (4.4.16)$$

The method is then defined formally in a manner similar to (4.4.9) with an appropriate modification to the bilinear form $(\nabla \cdot, \nabla \cdot)$ due to the terms $(\nabla \tilde{s}_i, \nabla \tilde{s}_{-j})$, see [Blum and Dobrowolski 1982] (where the method is described for the case $g = 0$).

For a certain class of problems this method provides optimal error estimates in L_2 , H^1 and L_∞ and also gives superconvergent approximations to the s.i.f.s, i.e. $\tilde{k}_i^h$ such that

$$|k_i - \tilde{k}_i^h| \leq ch^{2k-2} \quad (4.4.17)$$

if $k - 1$ order polynomials are used in $S^h(\Omega)$ and n in (4.4.11) is chosen so that $w \in W^{k,\infty}(\Omega)$. The restrictions on the class of problems able to be approximated by this method are that for s.i.f. estimates, which are

an integral part of the method, f is required to lie in a fractional order weighted Sobolev space whose order increases with n and (4.4.17) requires that no variational crimes be committed (as well as the usual restrictions on domain and data regularity). Practically this means for greater powers of approximation f will be required to have a zero of increasingly higher order (a fuller description of this will be given in the following section) and the full approximation power of (4.4.17) is lost in the presence of perturbation of domain errors. However the method when applicable is obviously very powerful and potentially gives very high accuracy s.i.f. estimates which are in general unobtainable via other methods.

4.5 A Practical Derivation of the DSFM

The DSFM will now be derived in a manner that leads naturally to a practical approximation scheme. Firstly the basic motivation for obtaining s.i.f. approximations is described which is then followed by the derivation of an analytic formula. This is then interpreted in two ways, one of which leads to Blum and Dobrowolski's computational form of the DSFM and the other an alternative formulation which forms the basis of the method analysed in Chapter 5.

The reason for seeking s.i.f. approximations is to exploit the knowledge of the singular expansion of u given by (4.4.4). With the projection operator introduced in (4.4.11) $P_h u$ may be decomposed as

$$P_h u = P_h w + \sum_i k_i P_h s_i, \quad (4.5.1)$$

(a full description being given in section 5.4). From this it is clear that the poor approximation power of $P_h u$ (see (4.4.7)) will be due to the singular function projections and with an approximation to k_i , k_i^h , this may be improved by defining u^h , noting (4.5.1), as

$$u^h = P_h u + \sum_i k_i^h (s_i - P_h s_i) = P_h w + \sum_i [(k_i - k_i^h) P_h s_i + k_i^h s_i]. \quad (4.5.2)$$

With $\{s_i\}_{i=1}^n$ known this is a straightforward calculation and its effect is to subtract off (up to the order of approximation in k_i^h) the poor behaviour of $P_h u$ (see 5.4 for a complete analysis).

Having given a method for exploiting s.i.f. estimates the problem of obtaining these estimates is addressed. This is done via an analytic expression for s.i.f.³ which is now derived. The expression is obtained from an invariant integral which allows the relevant coefficient to be isolated from the singular expansion for u (4.4.4) and written in terms of this integral. The quantities involved are u , $\bar{s}_i$ and $\bar{s}_{-i}$ where the latter two are respectively equal to s_i and s_{-i} (the singular and dual singular functions) in a region Ω_o surrounding the reentrant corner and contained within Ω . As these singular functions are harmonic in Ω_o , via Green's theorem the integral is re-interpreted in terms of u and the problem's data. If Ω is formally split into two subregions (Ω_1 and Ω_o) by Ω_o which is now taken to be the intersection of a disc centered on 0 with Ω , as shown, then the following line integral equation holds on Γ , the disc boundary contained in Ω (see Fig. 4.4):

$$\frac{1}{\pi} \int_{\Gamma} \left(\frac{\partial \bar{s}_i}{\partial r} \bar{s}_{-j} - \bar{s}_i \frac{\partial \bar{s}_{-j}}{\partial r} \right) = i \delta_{ij} \quad (4.5.3)$$

which is independent of the radius r that describes Γ .

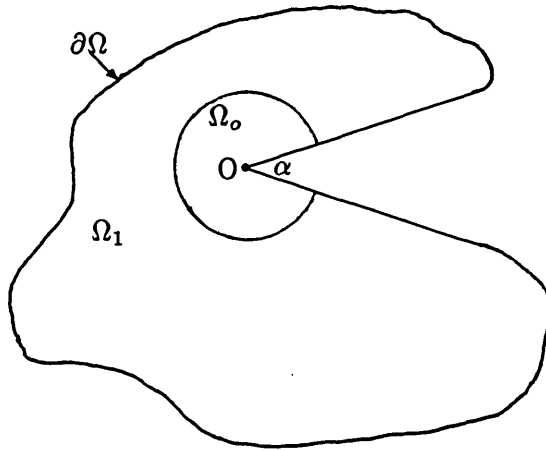


Figure 4.4: The decomposed reentrant corner domain

Now from (4.4.4) w , the smooth part of u , may be redefined as:

$$w = u - \sum_i k_i \bar{s}_i \quad (4.5.4)$$

and as the $\bar{s}_i$ are harmonic in Ω_o ,

$$-\nabla^2 u = -\nabla^2 w = f \quad \text{in } \Omega_o, \quad (4.5.5)$$

then from Green's theorem

$$\begin{aligned} \int_{\Omega_o} f \bar{s}_{-i} &= \int_{\partial\Omega_o} \left(w \frac{\partial \bar{s}_{-i}}{\partial \nu} - \frac{\partial w}{\partial \nu} \bar{s}_{-i} \right) \\ \int_{\Omega_1} f \bar{s}_{-i} &= \int_{\partial\Omega_1} \left(u \frac{\partial \bar{s}_{-i}}{\partial \nu} - \frac{\partial u}{\partial \nu} \bar{s}_{-i} \right) - \int_{\Omega_1} u \nabla^2 \bar{s}_{-i} \end{aligned} \quad (4.5.6)$$

where ν is the outward normal to the boundary defining each integral's domain.

From (4.5.3) and (4.5.4) it is clear that k_i may be written:

$$k_i = \frac{1}{i\pi} \int_{\Gamma} \left(\frac{\partial(u-w)}{\partial r} \bar{s}_i - (u-w) \frac{\partial \bar{s}_{-i}}{\partial r} \right) \quad (4.5.7)$$

and this line integral is contained within the right hand side of (4.5.6) i.e.

$$\begin{aligned} \int_{\Omega_o} f \bar{s}_{-i} &= \int_{\Gamma} \left(w \frac{\partial \bar{s}_{-i}}{\partial r} - \frac{\partial w}{\partial r} \bar{s}_{-i} \right) + \int_{\partial\Omega_o \setminus \Gamma} \left(w \frac{\partial \bar{s}_{-i}}{\partial \nu} - \frac{\partial w}{\partial \nu} \bar{s}_{-i} \right) \\ \int_{\Omega_1} f \bar{s}_{-i} &= - \int_{\Gamma} \left(u \frac{\partial \bar{s}_{-i}}{\partial r} - \frac{\partial u}{\partial r} \bar{s}_{-i} \right) + \int_{\partial\Omega_1 \setminus \Gamma} \left(u \frac{\partial \bar{s}_{-i}}{\partial \nu} - \frac{\partial u}{\partial \nu} \bar{s}_{-i} \right) \\ &\quad - \int_{\Omega_1} u \nabla^2 \bar{s}_{-i}. \end{aligned} \quad (4.5.8)$$

At this stage it is noted that on $\partial\Omega_o \setminus \Gamma$ $\bar{s}_{\pm i} = 0$, thus $u = w$, and $\nabla^2 \bar{s}_{-i} = 0$ in Ω_o and so using (4.5.7) and (4.5.8) k_i may be written hence:

$$\begin{aligned} k_i &= \frac{1}{i\pi} \left[\int_{\Omega} f \bar{s}_{-i} - \int_{\partial\Omega_o \setminus \Gamma} u \frac{\partial \bar{s}_{-i}}{\partial \nu} - \int_{\partial\Omega_1 \setminus \Gamma} u \frac{\partial \bar{s}_{-i}}{\partial \nu} + \int_{\partial\Omega_1 \setminus \Gamma} \frac{\partial u}{\partial \nu} \bar{s}_{-i} \right. \\ &\quad \left. + \int_{\partial\Omega_o \setminus \Gamma} \frac{\partial u}{\partial \nu} \bar{s}_{-i} + \int_{\Omega} u \nabla^2 \bar{s}_{-i} \right]. \end{aligned} \quad (4.5.9)$$

Now obviously $(\partial\Omega_o \setminus \Gamma) \cup (\partial\Omega_1 \setminus \Gamma) \equiv \partial\Omega$ (on which $u = g$) so (4.5.9) may be written simply as:

$$k_i = \frac{1}{i\pi} \left[\int_{\Omega} f \bar{s}_{-i} - \int_{\partial\Omega} g \frac{\partial \bar{s}_{-i}}{\partial \nu} + \int_{\partial\Omega} \frac{\partial u}{\partial \nu} \bar{s}_{-i} + \int_{\Omega} u \nabla^2 \bar{s}_{-i} \right] \quad (4.5.10)$$

(given these integrals exist which will be addressed later in the section).

The feature of (4.5.10) that is obviously important in terms of s.i.f. estimation is that the expression contains u i.e. it is not an explicit representation. However it may easily be exploited by using an approximation to u in (4.5.10) which will obviously give an estimate of the s.i.f. This can be seen as the fundamental step in both the DSFM and the alternative formulation presented in Chapter 5 and leads to potentially optimal approximations. These methods follow from a distinction of two separate s.i.f. formulations from (4.5.10) which will now be described along with a derivation of the approximation process defined by the DSFM.

Different s.i.f. formulations are deriveable by explicitly defining $\bar{s}_{\pm i}$ outside Ω_o with the two obvious choices being as follows:

$$\begin{aligned} \text{(i)} \quad \bar{s}_{\pm i} &= s_{\pm i} & \text{in } \Omega \\ \text{(ii)} \quad \bar{s}_{\pm i} &= \tau(r) s_{\pm i} = \tilde{s}_{\pm i} & \text{in } \Omega \end{aligned} \quad (4.5.11)$$

i.e. $\bar{s}_{\pm i}$ is taken identical to $s_{\pm i}$ throughout Ω , or modified by the smooth cutoff function as described in (4.4.9) for the SFM (here $\tau(r)$ being taken as 1 within Ω_o and only varying in Ω_1). These then lead to the following s.i.f. formulae, the first used in the alternative DSFM and the second being Blum and Dobrowolski's formulation (noting they analyse the case $g = 0$) i.e.

$$k_i = \frac{1}{i\pi} \left[(f, s_{-i}) - \left\langle g, \frac{\partial s_{-i}}{\partial \nu} \right\rangle + \left\langle \frac{\partial u}{\partial \nu}, s_{-i} \right\rangle \right] \equiv F_i(f, g, u) \quad (4.5.12)$$

or:

$$k_i = \frac{1}{i\pi} \left[(f, \tilde{s}_{-i}) + (u, \nabla^2 \tilde{s}_{-i}) \right], \quad (4.5.13)$$

where $(\cdot, \cdot)$ and $\langle \cdot, \cdot \rangle$ are the usual L_2 inner products over Ω and $\partial\Omega$ respectively and the functional form in (4.5.12) is defined for later use.

Now the first step is to seek an improvement u^h over the basic finite element approximation $P_h u$ which may be defined by (4.5.2) using the following s.i.f. estimate k_i^h given by:

$$k_i^h = \frac{1}{i\pi} \left[(f, \tilde{s}_{-i}) + (P_h u, \nabla^2 \tilde{s}_{-i}) \right]. \quad (4.5.14)$$

Then immediately an iterative process is suggested by replacing $P_h u$ in (4.5.14) by the newly formed u^h which will potentially lead to a sequence of successively more accurate approximations to u and s.i.f. estimates. This process Blum and Dobrowolski investigate [Blum and Dobrowolski 1982] and they show it converges to the solution $\hat{u}^h$ of their Petrov–Galerkin formulation referred to in 4.4. Analogous to the SFM form (4.4.12–14) they also give a computational form for $\hat{u}^h$ thus

$$\hat{u}^h = P_h u + \sum_{i=1}^n \hat{k}_i^h (\tilde{s}_i - P_h \tilde{s}_i) \quad (4.5.15)$$

with n sufficiently large for $P_h w$ (defined in (4.5.1)) to be an optimal approximation to w (i.e. so w is contained in $W^{k,\infty}(\Omega)$, see (4.4.4), if $k - 1$ order polynomials are used in the approximation) and where the $\hat{k}_i^h$ are defined implicitly by:

$$\hat{k}_i^h = \frac{1}{i\pi} \left[(f, \tilde{s}_{-i}) + (\hat{u}^h, \nabla^2 \tilde{s}_{-i}) \right]. \quad (4.5.16)$$

This then uses the notion of combining an approximate solution enhancement, as (4.5.2), with a s.i.f. estimation formula, as in (4.5.14), to give an implicit definition of $\hat{u}^h$ and $\{\hat{k}_i^h\}_{i=1}^n$ via (4.5.15–16). The latter are then obtained from the following matrix equation (derived from combining these two equations):

$$\sum_{j=1}^n a_{i,j} \hat{k}_j^h = (f, \tilde{s}_{-i}) + (P_h u, \nabla^2 \tilde{s}_{-i}) \quad i = 1 \rightarrow n \quad (4.5.17a)$$

where

$$a_{i,i} = i\pi + (P_h \tilde{s}_i - \tilde{s}_i, \nabla^2 \tilde{s}_{-i}) \quad (4.5.17b)$$

$$a_{i,j} = (P_h \tilde{s}_j, \nabla^2 \tilde{s}_{-i}) \quad i \neq j \quad (4.5.17c)$$

noting $(\tilde{s}_j, \nabla^2 \tilde{s}_{-i}) = 0$.

Having now defined the DSFM various practical details associated with its applicability and use will be noted. These will be used at the end of Chapter 5 (in 5.5) to give a comparison between this method and the alternative formulation derived in that chapter. They concern firstly the general method's applicability and then secondly problems arising from implementation in general situations.

Central to the method is existence of s.i.f. estimates which require (4.5.13) (or (4.5.10) for the more general case of g non-zero) to be well defined for appropriate values of i . This is in fact a harsh restriction when i becomes large as for (f, s_{-i}) and $\langle g, \partial s_{-i} / \partial \nu \rangle$ to exist f must have a zero of order $i\beta - 2 + \epsilon$ and g a zero order $i\beta + \epsilon$ at the corner (in addition to constraints necessary for the regularity of w). This will restrict the number of s.i.f.² able to be approximated and thus the smoothness of w hence the overall approximation power achievable for $\hat{u}^h$ and $\hat{k}_i^h$. The situation can however be improved in certain cases by a simple pre-processing of the problem to modify f into an appropriate form, details of this will be given in the discussion of the DSFM applicability in section 5.5.

Now dealing with practical problems relating to implementation of the DSFM, these arise from variational crimes and the use of a cutoff function $\tau(r)$ in defining $s_{\pm i}$. The effect of general boundary fitting on the method has no great consequence in terms of approximation to u , with standard isoparametric elements the same convergence rates will be observed. However it will destroy error estimates in negative norms required for the superconvergent s.i.f. estimates $\hat{k}_i^h$, though they will still be $O(h^k)$. Secondly numerical integration as will be used for a practical implementation will with standard methods give inaccurate results when integrating s_{-i} around the corner for large i due to the strong singularity that will be present. A certain amount of extra work would then have to be put into providing sufficiently accurate integrals in this case. Finally the smooth cutoff function is a computational inconvenience. It is an undefined quantity and leads, in comparison to the definition using $s_{\pm i}$, to a complex formulation involving its derivatives and an integral involving the approximation of u over Ω rather than over just the elements adjacent to its boundary $\partial\Omega$.

Chapter 5

A Dual Singular Function Method Without Cutoff Functions

5.1 Introduction

In this chapter using the s.i.f. representation (4.5.12) an alternative DSFM to that described in section 4.5 is developed. The process begins by deriving an abstract error bound for s.i.f. estimates obtained via approximations to the solution u . The bound is then used to derive an asymptotic convergence rate when using the simplest (or “basic”) s.i.f. approximations obtained with a finite element approximation to u .

Once it has been established that s.i.f.^s may be approximated with a certain degree of accuracy a singular decomposition of the finite element solution is demonstrated. From this a straightforward method of using s.i.f. estimates to enhance the finite element solution is then described and an abstract bound is constructed for this new approximation. The bound is given in terms of s.i.f. approximation errors which in turn are bounded by finite element solution errors previously derived and so a natural iteration process to obtain optimal approximations is immediately apparent. A straightforward and practical formulation of this iteration idea is then derived which involves simple matrix operations on the basic s.i.f. approximations and subsequently

in turn replaced by a single matrix inversion. This final stage is then the alternative DSFM formulation and optimal asymptotic convergence rates are established for both this and the iterative method. These steps follow the route taken by Blum and Dobrowolski [Blum and Dobrowolski 1982] in their exposition of the DSFM.

With the derivation and analysis of the alternative formulation complete the chapter closes with a comparison between this and the DSFM via a discussion of aspects of its implementation. Also it is shown that with linear finite elements the method has general applicability and optimal approximations are very easily obtained. Demonstration of these theoretical results via numerical experiments are then left until Chapter 6 where they are presented in conjunction with an extension of the method.

5.2 Stress Intensity Factor Approximation and Basic Error Analysis

The process of accurately approximating solutions of reentrant corner problems is now begun by deriving an approximation to the s.i.f. formulation (4.5.12). The most obvious approach is to find u^h approximating u in (4.4.1) and then substitute it into the third integral of (4.5.12). However this is not generally satisfactory as the resultant normal derivative is a poor approximation of the true flux $\partial u / \partial \nu$. This problem is though easily and practically overcome for a whole class of approximations as will now be shown. Then a bound on the resultant error in terms of $\|u - u^h\|_{0,\Omega}$ and other higher order terms is derived.

To give an improved flux approximation the general ideas of total flux estimation, as given by Barrett and Elliott [Barrett and Elliott 1987], are followed. Here the first stage is to define a ψ_i as follows:

$$\psi_i = s_{-i} \quad \text{on } \partial\Omega \quad \psi_i \in H^2(\Omega). \quad (5.2.1)$$

The smoothness condition is applied to allow definition of a sufficiently accurate interpolant for later use. With the arms of the corner being straight lines (see Fig. 4.3) s_{-i} will be zero on $\partial\Omega$ in a neighbourhood of the origin

and so ψ_i may be defined as follows,

$$\psi_i = a(r) \sin i\beta\theta$$

with $a(r) = r^{-i\beta}$ away from O and smoothly decaying to zero in its neighbourhood. Now using ψ_i the exact flux integral may be written via Green's theorem thus:

$$\int_{\partial\Omega} \frac{\partial u}{\partial \nu} s_{-i} = (\nabla u, \nabla \psi_i) - (f, \psi_i) \quad (5.2.2)$$

which can be approximated directly as follows:

$$\int_{\partial\Omega} \frac{\partial u}{\partial \nu} s_{-i} \approx (\nabla u^h, \nabla \Pi_h \psi_i) - (f, \Pi_h \psi_i) \quad (5.2.3)$$

where Π_h is an interpolation operator. A s.i.f. approximation follows by combining (4.5.12) and (5.2.3.) to yield

$$k_i^h = \frac{1}{i\pi} \left[\int_{\Omega} f s_{-i} - \int_{\partial\Omega} g \frac{\partial s_{-i}}{\partial \nu} + (\nabla u^h, \nabla \Pi_h \psi_i) - (f, \Pi_h \psi_i) \right]. \quad (5.2.4)$$

So far the function ψ_i used to define (5.2.3) has only been given on $\partial\Omega$, with a certain ambiguity associated with its behaviour in Ω . This may be overcome by finding an appropriate $a(r)$ as suggested though it complicates matters unnecessarily and would make the inner products messy to evaluate. However for $u^h \in S_g^h(\Omega)$ being a Galerkin finite element approximation and $\Pi_h \psi_i \in S^h(\Omega)$ interpolating s_{-i} on $\partial\Omega$, (5.2.3) reduces to a "boundary skin" integral. The finite element approximation is defined (see section 4.4) as the function $u^h \in S_g^h(\Omega)$ such that

$$(\nabla u^h, \nabla \chi) = (f, \chi) \quad \forall \chi \in S_o^h(\Omega). \quad (5.2.5)$$

Then decomposing S^h as $S^h \equiv S_o^h \oplus \partial S^h$ where ∂S^h is the subspace of S^h containing functions zero at nodes internal to Ω it follows that

$$\begin{aligned} \Pi_h \psi_i &= \chi_i + \phi_i & \text{where } \chi_i &\in S_o^h(\Omega) \text{ and} \\ \phi_i & (= \Pi_h \psi_i \equiv \Pi_h s_{-i} \text{ on } \partial\Omega) &\in \partial S^h. \end{aligned} \quad (5.2.6)$$

Thus using (5.2.5), (5.2.3) becomes:

$$\int_{\partial\Omega} \frac{\partial u}{\partial \nu} s_{-i} \approx (\nabla u^h, \nabla \phi_i) - (f, \phi_i) \quad (5.2.7)$$

which is of course now practical and simple to evaluate since the right hand side of (5.2.7) only involves an integration over elements adjacent to $\partial\Omega$.

It should be noted here that by referring (5.2.5) back to (4.4.6) it has been implicitly assumed that variational crimes are not present. For ease of exposition this will continue for the remainder of this chapter and in Chapter 6. Allowance for perturbation of domain and numerical integration errors could be made but it would render most of the analysis very involved akin to that of Chapter 3. No deterioration in error behaviour would result (excepting that already indicated in relation to the superconvergence result (4.4.17)).

Finally before defining the s.i.f. approximation it is important to note that a flux approximation of the form (5.2.7) is not specific to simple finite element solutions. Any approximation u^h defined so that the ambiguous part of $\Pi_h \psi_i$ may be eliminated from (5.2.3) can be used to construct an estimate of the flux as in (5.2.7). This observation will be used later for solution enhancement but now the complete approximate s.i.f. formulation, along with a discrete version of F_i , is given as follows:

$$\begin{aligned} k_i^h &= F_i^h(f, g, u^h) \\ &= \frac{1}{i\pi} \left[\int_{\Omega} f s_{-i} - \int_{\partial\Omega} g \frac{\partial s_{-i}}{\partial \nu} + (\nabla u^h, \nabla \phi_i) - (f, \phi_i) \right] \end{aligned} \quad (5.2.8)$$

with ϕ_i as defined in (5.2.6) and valid for any u^h as described. To now derive an error in this approximation the initial general flux estimate formulation (5.2.3) is used to give a bound for arbitrary u^h in abstract form from which this and other more specific results will follow. Thus:

Theorem 5.2.1

With k_i and k_i^h respectively defined by (4.5.12) and (5.2.4) the error in k_i^h is bounded as follows:

$$\begin{aligned} |k_i - k_i^h| &\leq c \left\{ \|u - u^h\|_{0,\Omega} + h|u - u^h|_{1,\Omega} \right. \\ &\quad \left. + \|s_{-i} - \Pi_h s_{-i}\|_{0,\infty,\partial\Omega} + \|u - u^h\|_{0,\partial\Omega} \right\} \end{aligned} \quad (5.2.9)$$

with $u^h \in H^1(\Omega)$ an approximation to u and Π_h an associated interpolation operator.

Proof

From the definitions of k_i and k_i^h

$$\begin{aligned} |k_i - k_i^h| &= |F_i(f, g, u) - F_i^h(f, g, u^h)| \\ &= \frac{1}{i\pi} \left| \int_{\partial\Omega} \frac{\partial u}{\partial \nu} s_{-i} - (\nabla u^h, \nabla \Pi_h \psi_i) + (f, \Pi_h \psi_i) \right| \end{aligned} \quad (5.2.10)$$

and via Green's theorem applied to $\int_{\partial\Omega} \partial u / \partial \nu \Pi_h s_{-i}$ noting (5.2.1)

$$\begin{aligned} &= \frac{1}{i\pi} \left| \int_{\partial\Omega} \frac{\partial u}{\partial \nu} (s_{-i} - \Pi_h s_{-i}) + (\nabla(u - u^h), \nabla \Pi_h \psi_i) \right| \\ &= \frac{1}{i\pi} \left| \int_{\partial\Omega} \frac{\partial u}{\partial \nu} (s_{-i} - \Pi_h s_{-i}) - (\nabla(u - u^h), \nabla(\psi_i - \Pi_h \psi_i)) \right. \\ &\quad \left. + (\nabla(u - u^h), \nabla \psi_i) \right|. \end{aligned} \quad (5.2.11)$$

Now using the regularity of ψ given in (5.2.1) the last term in (5.2.11) may be rewritten as

$$\int_{\partial\Omega} \left(\frac{\partial \psi_i}{\partial \nu} (u - u^h) \right) - (\nabla^2 \psi_i, u - u^h)$$

and thus:

$$\begin{aligned} |k_i - k_i^h| &\leq \frac{1}{i\pi} \left[\left| \left\langle \frac{\partial u}{\partial \nu}, s_{-i} - \Pi_h s_{-i} \right\rangle \right| + |(\nabla(u - u^h), \nabla(\psi_i - \Pi_h \psi_i))| \right. \\ &\quad \left. + |(\nabla^2 \psi_i, u - u^h)| + \left| \left\langle \frac{\partial u}{\partial \nu}, (u - u^h) \right\rangle \right| \right]. \end{aligned}$$

The first term may be bounded thus:

$$\left| \left\langle \frac{\partial u}{\partial \nu}, s_{-i} - \Pi_h s_{-i} \right\rangle \right| \leq \left\| \frac{\partial u}{\partial \nu} \right\|_{0,1,\partial\Omega} \|s_{-i} - \Pi_h s_{-i}\|_{0,\infty,\partial\Omega}$$

and noting that $u \in W^{2,1}(\Omega)$ the right hand side is bounded, via a trace inequality, by $\|u\|_{2,1,\Omega}$. The other terms may be bounded directly noting that $\psi \in H^2(\Omega)$ and the theorem follows. $\square$

Interpreting this general result in terms of finite element approximations is straightforward but before this is done various errors associated with the method, as will be required for (5.2.9), must be obtained. Firstly elements of general order k (≥ 2) will be considered with interpolation errors for general $w \in H^m(\Omega)$, $2 \leq m \leq k$, being as follows:

$$\|w - \Pi_h w\|_{r,\Omega} \leq ch^{m-r} \|w\|_{m,\Omega} \quad r = 0, 1 \quad (5.2.12)$$

and for $w \in W^{m,\infty}(\Omega)$, $2 \leq m \leq k$:

$$\|w - \Pi_h w\|_{0,\infty,\Omega} \leq ch^m. \quad (5.2.13)$$

The error behaviour of an approximation using k^{th} order elements will be considered optimal in the relevant norms if it satisfies (5.2.12) or (5.2.13) with $m = k$. Also boundary interpolant errors are required and for piecewise smooth boundaries (as present in this problem) the following norm is defined:

$$\|w\|_{m,\partial\Omega} = \left\{ \sum_{i=1}^p \|w\|_{m,\partial\Omega_i}^2 \right\}^{\frac{1}{2}}$$

where $\overline{\partial\Omega} = \cup_i \overline{\partial\Omega_i}$ and the $\partial\Omega_i$ are smooth. Interpolation errors of the form (5.2.12–13) will then hold over $\partial\Omega$, using this piecewise norm, for $w \in H^m(\partial\Omega)$, $w \in W^{m,\infty}(\partial\Omega)$, $2 \leq m \leq k$, respectively.

Also error bounds are required for the f.e. solution u^h approximating u which solves the reentrant corner problem (4.4.1). From [Ciarlet 1978] the H^1 interpolation error for $w \in W^{2,p}(\Omega)$, $1 \leq p < 2$, is given by:

$$|w - \Pi_h w|_{1,\Omega} \leq ch^{2-\frac{2}{p}} |w|_{2,p,\Omega}. \quad (5.2.14)$$

Now from (4.4.4) the leading singular behaviour in u results from s_1 , and has the form r^β , thus $u \in W^{2,p}(\Omega)$, $p = \frac{2}{2-\beta} - \epsilon$, for any $\epsilon > 0$ so:

$$|u - \Pi_h u|_{1,\Omega} \leq ch^{\beta-\epsilon}. \quad (5.2.15)$$

Then as

$$(\nabla(u - u^h), \nabla\chi) = 0 \quad \forall \chi \in S_o^h(\Omega)$$

from (4.4.1) and (4.4.6) the error may be written as follows:

$$\begin{aligned} |u - u^h|_{1,\Omega}^2 &= (\nabla(u - u^h), \nabla(u - u^h)) = (\nabla(u - u^h), \nabla(u - \Pi_h u)) \\ &\leq |u - u^h|_{1,\Omega} |u - \Pi_h u|_{1,\Omega} \end{aligned}$$

and using (5.2.15)

$$|u - u^h|_{1,\Omega} \leq ch^{\beta-\epsilon} \quad (5.2.16)$$

a result which also of course holds for s_1 replacing u . This then gives an H^1 error bound and from this, using the standard Aubin-Nitsche argument, an L_2 bound may be derived as follows. Let

$$\begin{aligned} -\nabla^2 z &= u - u^h & \text{in } \Omega \\ z &= 0 & \text{on } \partial\Omega \end{aligned}$$

then from the irregularity of Ω , $z \in W^{2, \frac{2}{2-\beta}-\epsilon}$ and $\|z\|_{2, \frac{2}{2-\beta}-\epsilon, \Omega} \leq c\|e\|_{0, \Omega}$ and so

$$\begin{aligned} \|u - u^h\|_{0, \Omega}^2 &= (u - u^h, -\nabla^2 z) = (\nabla(u - u^h), \nabla z) - \int_{\partial\Omega} (g - \Pi_h g) \frac{\partial z}{\partial \nu} \\ &\leq (\nabla(u - u^h), \nabla(z - \Pi_h z)) + \|g - \Pi_h g\|_{0, \infty, \partial\Omega} \left\| \frac{\partial z}{\partial \nu} \right\|_{0, 1, \partial\Omega} \\ &\leq |u - u^h|_{1, \Omega} |z - \Pi_h z|_{1, \Omega} + ch^2 |z|_{1, 1, \Omega} \end{aligned}$$

from the regularity of g . Thus using (5.2.14) and (5.2.16)

$$\|u - u^h\|_{0, \Omega} \leq ch^{2\beta-\epsilon}. \quad (5.2.17)$$

The L_∞ error is more complicated to derive and follows from [Schatz and Wahlbin 1979] noting that $u \in H^{\beta+1-\epsilon}(\Omega)$:

$$\|u - u^h\|_{0, \infty, \Omega} \leq ch^{\beta-\epsilon} \quad (5.2.18)$$

a result that again also obviously holds for s_1 replacing u . Finally a result is required on the L_∞ error behaviour of errors in finite element projections. From [Schatz 1980] with $w \in W^{k, \infty}(\Omega)$ and P_h the projection operator, defined by (4.4.11), associated with an k^{th} ($k \geq 2$) order finite element subspace $S^h(\Omega)$ then:

$$\|w - P_h w\|_{0, \infty, \Omega} \leq ch^k |\ln h|^r \quad (5.2.19)$$

where $r = 1$ if $k = 2$ and zero otherwise. This implies that finite element projections are optimal in L_∞ unless $k = 2$ when asymptotically they fail by the minimal factor h^ϵ (ϵ any small positive number).

With all the relevant preliminary error results now obtained the second main task of this section may now be accomplished, namely bounding the

derived s.i.f. approximations' error when using a practical procedure to find k_i^h . This is done via the following theorem:

Theorem 5.2.2.

If a standard finite element solution u^h approximating u from (4.4.1) is used to define k_i^h via (5.2.8) then assuming $g \in H^{k+2}(\Omega)$,

$$|k_i - k_i^h| \leq ch^{2\beta-\epsilon}. \quad (5.2.20)$$

Proof

Note that from (4.4.1) $u = g$ on $\partial\Omega$ thus:

$$\|u - u^h\|_{0,\partial\Omega} = \|g - \Pi_h g\|_{0,\partial\Omega} \leq ch^k$$

via the boundary interpolation errors derived above and the smoothness of g (see 4.4) and with s_{-i} smooth on $\partial\Omega$ a similar bound holds for $\|s_{-i} - \Pi_h s_{-i}\|_{0,\infty,\partial\Omega}$. Thus from (5.2.9) and using (5.2.16) and (5.2.17)

$$|k_i - k_i^h| \leq c\{h^{2\beta-\epsilon} + h^{1+\beta-\epsilon} + h^k\}.$$

The result now follows immediately noting $\frac{1}{2} \leq \beta < 1$ and $k \geq 2$. $\square$

5.3 Enhanced Finite Element Approximations

Having now shown that s.i.f. approximation is possible the method outlined in section 4.5 to use this information for increasing the accuracy of u^h is now presented. First the low accuracy components of u^h are isolated and then a method for solution enhancement is given for a general s.i.f. estimate. Simple error bounds for improved approximations in terms of the s.i.f. approximation error are then obtained and practical consequences of these follow immediately. Finally to close the section a general method for increasing s.i.f. approximation accuracy is given.

Via the projection operator P_h introduced in (4.4.11) and by writing

$$P_h v \equiv v^h \quad (5.3.1)$$

(4.5.1) gives a singular decomposition of u^h thus:

$$u^h = w^h + \sum_{i=1}^n k_i s_i^h. \quad (5.3.2)$$

Here the s_i^h (necessary for later approximation enhancement) are explicitly given by

$$\begin{aligned} (\nabla s_i^h, \nabla \chi) &= 0 & \forall \chi \in S_o^h(\Omega) \\ s_i^h &= \Pi_h s_i & \text{on } \partial\Omega \end{aligned} \quad (5.3.3)$$

as the s_i are harmonic in Ω (and known).

Now using the decomposition of u^h in (5.3.2) the suboptimal components it contains may be isolated. From (5.2.19), when using k^{th} order elements $w \in W^{k,\infty}(\Omega)$ is required for w^h to be optimal in L^∞ (or minimally suboptimal when $k = 2$). This will be satisfied if n is such that

$$n\beta \geq k - \beta.$$

(along with the global regularity requirements of $f \in W^{k-2,\infty}(\Omega)$, $g \in W^{k,\infty}(\Omega)$) i.e. so that the next singular function is at least of order $r^{(n+1)\beta}$ (and thus in $W^{k,\infty}(\Omega)$). This implies that taking n such that

$$n \geq \frac{k - \beta}{\beta} \quad (5.3.4)$$

the summation in (5.3.2) will contain all suboptimal behaviour of u^h . Then finally combining (4.4.4) and (5.3.2) allows an associated error decomposition

$$\|u - u^h\| \leq \|w - w^h\| + \sum_{i=1}^n |k_i| \|s_i - s_i^h\| \quad (5.3.5)$$

where the role of s_1^h in limiting the convergence of u^h becomes obvious. This above analysis of course does nothing to help enhance u^h 's approximation power though does isolate the area of difficulty. It is by using this information along with the s.i.f. estimates now available that improvements are made.

For use in later analysis a general s.i.f. estimate $\bar{k}_i^h$ is now introduced and is used to improve the basic finite element solution u^h within the following method. First calculate the s_i^h and then use $\bar{k}_i^h s_i^h$ to partially subtract the $k_i s_i^h$ in (5.3.2) (thus leaving a residual of $(k_i - \bar{k}_i^h) s_i^h$) and replace it by $\bar{k}_i^h s_i$. The first part is thus a simple finite element calculation with which a new

approximation $\bar{u}^h$ is then defined as follows:

$$\bar{u}^h = u^h + \sum_{i=1}^n \bar{k}_i^h (s_i - s_i^h). \quad (5.3.6)$$

The relevant error is then reduced by a factor $|k_i - \bar{k}_i^h|$ as will be demonstrated in the following theorem:

Theorem 5.3.1

If u^h is a finite element approximation to u defined in (4.4.1) (with n s.t. $n \geq (k - \beta)/\beta$) and $\bar{k}_i^h$ is an estimate of the s.i.f. k_i defined in (4.5.12) then $\bar{u}^h$ as given in (5.3.6) is an enhanced approximation satisfying the following bounds:

$$\begin{aligned} \|u - \bar{u}^h\|_{0,\Omega} &\leq c \left(h^{2\beta-\epsilon} \sum_{i=1}^n |k_i - \bar{k}_i^h| + h^{k-r\epsilon} \right) \\ \|u - \bar{u}^h\|_{1,\Omega} &\leq c \left(h^{\beta-\epsilon} \sum_{i=1}^n |k_i - \bar{k}_i^h| + h^{k-1} \right) \\ \|u - \bar{u}^h\|_{0,\infty,\Omega} &\leq c \left(h^{\beta-\epsilon} \sum_{i=1}^n |k_i - \bar{k}_i^h| + h^{k-r\epsilon} \right) \end{aligned} \quad (5.3.7)$$

where $r = 1$ for $k = 2$ and zero otherwise.

Proof

With $\|\cdot\|$ representing a general norm:

$$\|u - \bar{u}^h\| = \|u - u^h + \sum_{i=1}^n \bar{k}_i^h (s_i - s_i^h)\|$$

and from the singular expansions of u and u^h :

$$\begin{aligned} \|u - \bar{u}^h\| &= \|w - P_h w + \sum_{i=1}^n k_i (s_i - s_i^h) - \bar{k}_i^h (s_i - s_i^h)\| \\ &\leq \|w - P_h w\| + \sum_{i=1}^n |k_i - \bar{k}_i^h| \|s_i - s_i^h\|. \end{aligned} \quad (5.3.8)$$

Now with s_1 being the strongest singularity, the summation's error will obviously be bounded by $c\|s_1 - s_1^h\|$ and thus from (5.3.8)

$$\|u - \bar{u}^h\| \leq c \left\{ \|w - P_h w\| + \|s_1 - s_1^h\| \sum_{i=1}^n |k_i - \bar{k}_i^h| \right\}. \quad (5.3.9)$$

(5.3.7) then follows directly from (5.2.16), (5.2.17) and (5.2.18). The right hand side quantity in each bracket represents the convergence of $P_h w$ i.e. always optimal in H^1 and a minimal violation in the other norms (here, due to domain irregularity L_2 bounds are derived from the L_∞ error result (5.2.19)). $\square$

The above then gives a straightforward procedure to utilise s.i.f. approximations for enhancement of u^h . Using k_i^h as defined in theorem (5.2.2) in (5.3.6) gives, from the basic s.i.f. error result (5.2.20), an increase of $h^{2\beta-\epsilon}$ (at most, depending on the size of k) in error bounds over the respective estimates in (5.2.16), (5.2.17) and (5.2.19). With this new approximation it then follows from (5.2.8) that there is the possibility of producing improved s.i.f. estimates as well. This, as it happens, is an entirely practical idea and immediately suggests an iteration procedure for further enhancements to the approximations of u and $\{k_i\}_{i=1}^n$. In fact optimal error bounds may be derived as a result and details of these with relevant methods are presented in the following section.

5.4 Optimal Approximations

Before developing the final theory associated with approximation of these reentrant corner problems a brief resumé of the main results developed so far is useful as they form a basis for this section's work. The first of these is that using an approximation to u satisfying (5.2.5) a s.i.f. estimate may be defined by (5.2.8) with the resultant error given by (5.2.9), which when using standard finite elements converts to $h^{2\beta-\epsilon}$. These s.i.f. estimates may then be used to define an enhanced approximation to u via (5.3.6) with the accuracy in any norm increased by a factor of these estimate's errors (unless greater than optimal) as given in (5.3.7). Finally it was then suggested that an enhanced f.e. solution of the form (5.3.6) could be used to produce improved s.i.f. estimates with an iteration procedure then obvious to obtain further increases in accuracy. From this idea a practical procedure for optimal approximations is developed immediately below and this is then modified to produce an alternative DSFM which closes the section.

The first step towards optimal approximations is to use an enhanced finite element solution of the form (5.3.6) to produce an s.i.f. estimate. As described in section 5.2 for this to be practical $\bar{u}^h$ must satisfy (5.2.5) i.e.

$$(\nabla \bar{u}^h, \nabla \chi) = (f, \chi) \quad \forall \chi \in S_o^h(\Omega).$$

This is obvious as from (5.3.6) the equation may be written thus:

$$(\nabla u^h, \nabla \chi) + \sum_{i=1}^n \bar{k}_i^h (\nabla(s_i - s_i^h), \nabla \chi) = (f, \chi) \quad \forall \chi \in S_o^h(\Omega) \quad (5.4.1)$$

with the second left hand side term zero by definition of $S^h(\Omega)$. Thus a s.i.f. estimate may be given by:

$$\bar{k}_i^h = F_i^h(f, g, \bar{u}^h) \quad (5.4.2)$$

with the error therefore bounded by the right hand side of (5.2.9) i.e. having $\bar{u}^h$ replacing u^h and thus, by Theorem 5.3.1, an improvement over the basic estimate k_i^h .

Having determined that a technique for s.i.f. estimate enhancement is possible, the iteration procedure mentioned is now developed and details of its implementation are then given. Firstly using the ideas of 5.3 an enhancement to the basic f.e. solution u_1^h is defined thus:

$$u_1^h = u^h + \sum_{i=1}^n k_i^h (s_i - s_i^h) \quad (5.4.3)$$

where $k_i^h = F_i^h(f, g, u^h)$. From the above k_i^1 may now be defined as:

$$k_i^1 = F_i^h(f, g, u_1^h). \quad (5.4.4)$$

k_i^1 is then a better estimate of k_i than k_i^h . The iteration procedure now follows by defining for a general l , k_i^l and u_l^h thus:

$$u_l^h = u^h + \sum_{i=1}^n k_i^{l-1} (s_i - s_i^h) \quad (5.4.5a)$$

$$k_i^l = F_i^h(f, g, u_l^h). \quad (5.4.5b)$$

Now by expanding the right hand side of (5.4.5b) using (5.2.8) and (5.4.5a) k_i^l may be written thus:

$$\begin{aligned}
k_i^l &= F_i^h(f, g, u^h) + \sum_{j=1}^n k_j^{l-1} F_i^h(0, 0, s_j - s_j^h) \\
&= k_i^h + \sum_{j=1}^n k_j^{l-1} (\nabla(s_j - s_j^h), \nabla \phi_i) \\
&= k_i^h + \sum_{j=1}^n d_{ij} k_j^{l-1}
\end{aligned} \tag{5.4.6}$$

where

$$d_{ij} = (\nabla(s_j - s_j^h), \nabla \phi_i)$$

and via the harmonicity of s_j , in more convenient form:

$$= \int_{\partial\Omega} \frac{\partial s_j}{\partial \nu} \Pi_h s_{-i} - (\nabla s_j^h, \nabla \phi_i). \tag{5.4.7}$$

This means that if $\underline{k}^h$ and $\underline{k}_i^h$ are defined as:

$$\underline{k}^h = (k_1^h, \dots, k_n^h) \text{ and } \underline{k}_i^h = (k_1^h, \dots, k_n^h) \tag{5.4.8}$$

then from (5.4.6), with $D = \{d_{ij}\}$,

$$\underline{k}^l = \underline{k}^h + D \underline{k}^{l-1}. \tag{5.4.9}$$

So it follows that a general enhanced s.i.f. estimate $\underline{k}^p$ may be defined simply in terms of the “small correction factor” matrix D and the basic estimate $\underline{k}^h$ noting $\underline{k}^1 = \underline{k}^h + D \underline{k}^h$ thus:

$$\underline{k}^p = \left(\sum_{i=0}^{p-1} D^i \right) \underline{k}^h \tag{5.4.10}$$

or the more computationally convenient iterative form:

$$\underline{k}^1 = \underline{k}^h + D \underline{k}^h \quad \underline{k}^p = \underline{k}^h + D \underline{k}^{p-1}, \quad p \geq 2. \tag{5.4.11}$$

This thus reduces the iteration procedure to a series of very simple operations (and relevant u_{p+1}^h may be defined at any stage but are unnecessary).

Using (5.4.10), (5.3.6) and other results the following theorem defines completely this iteration process and provides bounds for the errors in s.i.f. estimates and approximations to u thus defined.

Theorem 5.4.1

If u^h is the standard finite element approximation to u defined by (4.4.6) and $\underline{k}^h$ is the initial s.i.f. vector computed with u^h in (5.2.8) then using D given by (5.4.7) $\underline{k}^p$ where:

$$\underline{k}^p = \left(\sum_{i=0}^p D^i \right) \underline{k}^h \quad (5.4.12)$$

satisfies the error bound:

$$|\underline{k} - \underline{k}^p| \leq ch^{\min((2(p+1)\beta-\epsilon, k-r\epsilon))} \quad (5.4.13)$$

where $|\underline{k} - \underline{k}^p| = \sum_{i=1}^n |k_i - k_i^p|$, assuming $n \geq (k - \beta)/\beta$. Also if u_{p+1}^h ($p \geq 0$) is defined by

$$u_{p+1}^h = u^h + \sum_{i=1}^n k_i^p (s_i - s_i^h) \quad (5.4.14)$$

then with r as in Theorem 5.3.1:

$$\begin{aligned} \|u - u_{p+1}^h\|_{0,\Omega} &\leq ch^{\min((2p+4)\beta-\epsilon, k-r\epsilon)} \\ \|u - u_{p+1}^h\|_{1,\Omega} &\leq ch^{\min((2p+3)\beta-\epsilon, k-1)} \\ \|u - u_{p+1}^h\|_{0,\infty,\Omega} &\leq ch^{\min((2p+3)\beta-\epsilon, k-r\epsilon)}. \end{aligned} \quad (5.4.15)$$

Proof

Using (5.2.20), (5.4.13) is satisfied for $p = 0$ which then means via (5.3.7) so is (5.4.15). Now assume (5.4.13) and (5.4.15) true for $p = l$ then from Theorem 5.2.1:

$$\begin{aligned} |k_i - k_i^{l+1}| &\leq c \left\{ \|u - u_{l+1}^h\|_{0,\Omega} + h \|u - u_{l+1}^h\|_{1,\Omega} \right. \\ &\quad \left. + \|s_{-1} - \Pi_h s_{-1}\|_{0,\infty,\partial\Omega} + \|u - u_{l+1}^h\|_{0,\partial\Omega} \right\} \\ &\leq c \left\{ h^{\min((2l+4)\beta-\epsilon, k)} + h \cdot h^{\min((2l+3)\beta-\epsilon, k-1)} + ch^k \right\} \\ &\leq ch^{\min(2(l+2)\beta-\epsilon, k-\epsilon)} \end{aligned}$$

as $\frac{1}{2} \leq \beta < 1$. This proves (5.4.14) is satisfied for $p = l + 1$ which then via (5.3.7) shows (5.4.15) is also true, thus by induction the theorem follows. $\square$

Corollary

Optimal approximations to u and $\{k_i\}_{i=1}^n$ (i.e. $O(h^{k-r\epsilon})$ to the latter) will be achieved in p iterations (giving $\{k_i^p\}_{i=1}^n$ and u_{p+1}^h) where

$$p \geq \frac{k - 2\beta + (1 - r)\epsilon}{2\beta}. \quad (5.4.16)$$

This statement is obvious in view of (5.4.13) and an immediate consequence is that using piecewise linear elements optimal approximations are delivered by simply one iteration (noting $1 > \beta \geq \frac{1}{2}$). $\square$

Finally having demonstrated $\bar{u}^h$ as defined in (5.3.6) satisfies (5.2.5) and then produced the practical iteration step (5.4.9), optimal approximations can be obtained directly by means of this last equation by using an implicit definition for the optimal s.i.f.² $\hat{k}_i^h$. This process in fact defines the DSFM without cutoff functions and resultant optimal error bounds are now demonstrated.

Theorem 5.4.2 - The DSFM Without Cutoff Functions

Let $\underline{\hat{k}}^h = (\hat{k}_1^h, \dots, \hat{k}_n^h)$ be defined by:

$$\underline{\hat{k}}^h = \underline{k}^h + D\underline{\hat{k}}^h \quad \text{i.e.} \quad \underline{\hat{k}}^h = (I - D)^{-1}\underline{k}^h \quad (5.4.17)$$

where D is the correction factor matrix defined in (5.4.7) and

$$\hat{u}^h = u^h + \sum_{i=1}^n \hat{k}_i^h (s_i - s_i^h). \quad (5.4.18)$$

Then assuming $n \geq (k - \beta)/\beta$ and with r as previously defined:

$$|\underline{k}_i - \underline{\hat{k}}_i^h| \leq ch^{k-r\epsilon} \quad (5.4.19)$$

and:

$$\begin{aligned} \|u - \hat{u}^h\|_{0,\Omega} &\leq ch^{k-r\epsilon} \\ \|u - \hat{u}^h\|_{0,\Omega} &\leq ch^{k-1} \\ \|u - \hat{u}^h\|_{0,\infty,\Omega} &\leq ch^{k-r\epsilon}. \end{aligned} \quad (5.4.20)$$

Proof

From Theorems 5.2.1 and 5.3.1, and finite element interpolation results:

$$\begin{aligned} |k_i - \hat{k}_i^h| &\leq c \left\{ \|u - \hat{u}^h\|_{0,\Omega} + h|u - \hat{u}^h|_{1,\Omega} + h^{k-r\epsilon} \right\} \\ &\leq c \left\{ h^{k-r\epsilon} + (h^{2\beta-\epsilon} + h^{\beta+1-\epsilon}) \sum_{i=1}^n |k_i - \hat{k}_i^h| \right\}. \end{aligned}$$

Thus summing over i gives:

$$|\underline{k}_i - \hat{\underline{k}}_i^h|(1 - Kh^{2\beta-\epsilon}) \leq ch^{k-r\epsilon}$$

and so for h sufficiently small (5.4.19) is satisfied. As previously the other bounds follow from this result and Theorem 5.3.1. $\square$

This error bound is derived in the same manner as Blum and Dobrowolski's bound for their DSFM, the only difference in the methods being their d.s.f. definitions.

5.5 General Applicability of a DSFM

Having defined the DSFM of Blum and Dobrowolski in 4.5 and presented an alternative version in this chapter, the applicability of these methods will now be discussed. First the general class of problems on which the methods may be used is investigated. From this, constraints on the use of high order elements result and their implications are then considered. Various practical details relating to implementation are then discussed and finally advantages of the new method are described.

The fundamental factor which influences both methods is data regularity. The theory limiting approximation power by the global regularity of f and g still applies but of relevance here are the extra constraints resulting from the data integrals required for s.i.f. estimation (as outlined at the end of Chapter 4). The problem results from the increasing degree of singularity in the d.s.f. or its derivative against which f or g (respectively) have to be integrated. For arbitrary order elements this requires the data to have high order zeros to obtain sufficient s.i.f. estimates necessary to leave the smooth residual w in an appropriate Sobolev space for optimal approximation. (Also

if s.i.f. estimates are independently required then these zeros in the data are again obviously a necessity). The exact constraints on the data will now be investigated.

The relation that governs the number of singular functions n , and thus the highest order d.s.f., required for optimal approximations using a general finite element subspace of order k is (5.3.4) i.e.

$$n \geq \frac{k - \beta}{\beta}. \quad (5.5.1)$$

n will then be chosen as the minimum value satisfying this i.e. so that

$$(n + 1)\beta \geq k > n\beta \quad (5.5.2)$$

and thus the maximum order of d.s.f. is given by,

$$s_{-n} \approx r^{n\beta} \approx r^{-k+\epsilon} \quad \forall \epsilon > 0. \quad (5.5.3)$$

From this, constraints are necessary on f and g at the corner to ensure existence of the integrals involving these data functions used in the s.i.f. estimates (see (4.5.10)). Now f is integrated against the d.s.f. itself across the whole domain and thus must have a zero of order p at the corner such that

$$p - k + 1 + \epsilon \geq -1 + \epsilon \quad \Rightarrow \quad p \geq k - 2. \quad (5.5.4)$$

g however is integrated against the d.s.f. normal derivative along only the boundary in which case requires a zero of order q at the corner such that

$$q - k - 1 + \epsilon \geq -1 + \epsilon \quad \Rightarrow \quad q \geq k. \quad (5.5.5)$$

(5.5.4) and (5.5.5) then give the constraints above those of global regularity necessary for sufficient s.i.f.^s to be approximated to allow optimal approximations.

In relation to a practical implementation if linear elements are used then $k = 2$ and (5.5.4–5) imply there will be no restriction on f (other than $f \in L_\infty(\Omega)$) but g will need a zero of order two at the corner (along with the requirement $g \in W^{2,\infty}(\Omega)$). For quadratic elements these constraints

will be raised by one i.e. f will require a first order and g a third order zero at the corner, which then becomes rather a severe restriction on the method's applicability. However for linears a simple pre-processing may be applied to the problem leaving it in a form where relevant data integrals will be well-defined. For the method to be generally applicable all that is required is to isolate the constant and linear dependence of g at the corner which, from its regularity, may be given using cartesian and polar coordinates

$$g = a + bx + cy + O(r^2). \quad (5.5.6)$$

If the problem is then redefined to solve for:

$$\bar{u} = u - a - bx - cy \quad (5.5.7)$$

the s.i.f. boundary integral will exist and thus s.i.f. estimates will be readily obtainable. For quadratics the r^2 term in (5.5.6) will have to be removed in order for the boundary integral to exist i.e. $\bar{u}$ (in (5.5.7)) will then have to contain second order terms. However the redefined problem will then in general have a constant term on the right hand side (from the Laplacian operating on the second order terms) which will then not be integrable against $r^{-3+\epsilon}$ as required for the domain integral to exist. This problem will recur for higher order approximations and indicates that for other than linear finite elements very special data is required for optimal approximations using the DSFM.

From other points of view also linear elements are a desirable choice for use in the DSFM due to their simplicity and, as above, their requirement of only low order dual singular functions. The latter will reduce the effort in data integral evaluation which would become large for high order singularities. Some form of numerical integration would be used in a practical implementation and in the presence of highly singular functions this will require a good deal of computational effort. Also the number of singular function projections necessary for approximations to u will be minimized and they will also be able to be evaluated quickly.

Finally advantages of the new DSFM formulation presented in this Chapter will be discussed. Its basic appeal is in the use of unmodified singular

functions which allow a simple and explicit definition of the s.i.f. estimates. The smooth cutoff functions are an inconvenience in terms of their general definition and because they result in increased computational effort. This is demonstrated by considering the requirements on a cutoff function in the simplest case of a piecewise linear approximation. When calculating the basic s.i.f. estimate the integral $\int_{\Omega} u^h \nabla^2 \tilde{s}_{-i}$ is required and in a practical implementation this will be done via numerical integration. For the integral to be $O(h^2)$ accurate $\nabla^2 \tilde{s}_{-i} \in H^2(\Omega)$ is required and with $\tilde{s}_{-i} = \tau(r)s_{-i}$ ($\tau(r) = 1$ $r < a$ and $\tau(r) = 0$ $r > b$, $0 < a < b < 1$) then

$$\nabla^2 \tilde{s}_{-i} = [r^{-i\beta} \tau'' + r^{-(i\beta+1)}(1 - 2i\beta)\tau'] \sin i\beta\theta, \quad (5.5.8)$$

so τ must satisfy:

$$\tau \in C^3(0,1)$$

and

$$\begin{aligned} \tau(a) = 1, \quad \tau'(a) = \tau''(a) = \tau'''(a) = 0 \\ \tau(b) = \tau'(b) = \tau''(b) = \tau'''(b) = 0. \end{aligned} \quad (5.5.9)$$

Thus τ must be a seventh order polynomial. Also due to the domain of definition of $\tau(r)$ being an $O(1)$ interval then the integral will be over $O(1/h^2)$ elements rather than the $O(1/h)$ contained in the equivalent unmodified singular function boundary skin integral $(\nabla u^h, \nabla \phi_i)$.

The definition of a DSFM without cutoff functions is also a natural extension of the work of Babuska and Miller [Babuška and Miller 1984]. They give recipes for post-processing finite element solutions to obtain s.i.f. estimates in a wide range of elliptic problems with reentrant corner domains i.e. involving Neumann and mixed as well as Dirichlet boundary data. For these problems only when Dirichlet data is present other than on the reentrant corner arms do they require cutoff functions for their s.i.f. estimation. Thus the work presented here removes this restriction and allows the general use of unmodified singular functions for s.i.f. estimation.

Finally the superconvergent s.i.f. estimates (i.e. $O(h^{2k-2})$) that result from using $\tilde{s}_{\pm i}$ can be reproduced within the new formulation by using an L_2

projection of s_{-i} on the boundary rather than the interpolant $\Pi_h s_{-i}$. This would then allow negative norm estimates to be introduced in the convergence proofs as in Blum and Dobrowolski. However for practical purposes these questions are academic as if any boundary fitting (e.g. via isoparametric elements) is required then negative norm estimates are destroyed and thus both methods will give the same $O(h^k)$ convergence. Thus this new formulation can be seen as a useful practical extension of the DSFM, especially in the case of $f = 0$ or with an implementation using linear elements which has a simple and general applicability

Chapter 6

Application of the DSFM to Two Test Problems

6.1 Introduction

This chapter describes implementation of the practical DSFM for the approximation of two model problems. First a simple test problem is constructed on a slit square domain so that there is one reentrant corner with the magnitude of the angle being 2π . The second problem concerns an idealized fractured well. This gives rise to a domain which contains two reentrant corners with both angles having the magnitude 2π .

In solving these problems several of the details concerned with the applicability of the DSFM, raised in Chapters 4 and 5, are encountered. With reentrant corners of angle 2π , from (4.4.3), $\beta = \frac{1}{2}$ thus $i\beta$ is integer for i even and the log terms of (4.4.2) will potentially be present. For s.i.f. approximations numerical integration will have to be performed on the data integrals in (5.2.8). In addition the presence of more than one reentrant corner has not yet been discussed. For this a modification of the technique described in Chapter 5 is required allowing s.i.f. integrals to be calculated over sub-domains of the problem's domain of definition and then using a simple cutoff function technique to recover optimal approximations. These concerns will be dealt with as they arise while presenting the problems in following sections.

6.2 A Simple Test Problem

To demonstrate the DSFM the following test problem is approximated, find u such that

$$\nabla^2 u = 0 \quad \text{in } \Omega \quad (6.2.1a)$$

$$u = s_1 + s_3 + s_5 \text{ on } \partial\Omega \quad (6.2.1b)$$

where $s_i = r^{\frac{i}{2}} \sin \frac{i\theta}{2}$ and Ω is a rectangular slit domain (size 1×0.6) as shown in Fig. 6.1. This gives rise to a domain with one reentrant corner of angle 2π (being chosen as the origin) and as the s_i are harmonic in Ω then the exact solution to (6.2.1) is obviously

$$u = s_1 + s_3 + s_5. \quad (6.2.2)$$

The approximation is obtained using piecewise linear basis functions.

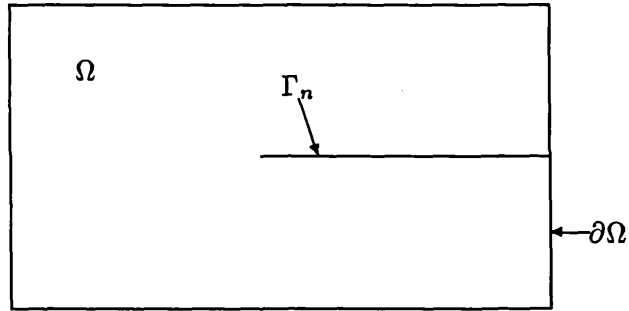


Figure 6.1: The simple slit domain

Now (6.2.1) has a solution of the form (4.4.4) and is amenable to approximation by the DSFM. Firstly the number of singular functions required in (6.2.2) for optimal approximations must be determined. The requirement is that w may be optimally approximated by its finite element projection so

that from (5.3.4), with approximation order $k = 2$ and $\beta = \frac{1}{2}$, n singular functions are necessary where:

$$n \geq \frac{2 - \frac{1}{2}}{\frac{1}{2}} = 3. \quad (6.2.3)$$

The second singular function corresponds to the case $i\beta$ being an integer and is assumed to have been eliminated a priori by massaging the data. Clearly by inspection this is not necessary for (6.2.1) either from the explicit solution or from the data and so the required singular decomposition is:

$$u = w + k_1 s_1 + k_3 s_3, \quad (6.2.4)$$

with $w \in W^{2,\infty}(\Omega)$. The theory of Chapter 5 then states that using the DSFM to obtain enhanced s.i.f estimates $\hat{k}_1^h, \hat{k}_3^h$, these will be $O(h^2)$ and via projections of s_1 and s_3 an optimal approximation $\hat{u}^h$ to u may be constructed. The approximation process is thus as follows. Calculate the finite element approximation of u i.e., from (4.4.6) and (6.2.1), $u^h \in S_g^h(\Omega)$, $g = s_1 + s_3 + s_5$, such that:

$$(\nabla u^h, \nabla \chi) = 0 \quad \forall \chi \in S_o^h(\Omega) \quad (6.2.5)$$

and the projections of s_1 and s_3 , s_1^h and s_3^h , as defined in (5.3.3). Then use u^h to construct basic s.i.f estimates k_i^h , $i = 1, 3$, where, using (5.2.8),

$$k_i^h = \frac{1}{i\pi} \left\{ (\nabla u^h, \nabla \phi_i) - \left\langle g, \frac{\partial s_{-i}}{\partial \nu} \right\rangle_h \right\} \quad (6.2.6)$$

where ϕ_i is defined by (5.2.6) and $\langle v, w \rangle_h$ a trapezoidal approximation to the exact integral $\int_{\partial\Omega} vw$. As $g = 0$ on the slit and is smooth elsewhere the numerical integral in (6.2.6) will be an $O(h^2)$ approximation thus not affecting the $O(h^{2\beta-\epsilon})$ bound on $|k_i - k_i^h|$ furnished by (5.2.20).

Using these elementary approximations the DSFM techniques may now be applied to obtain optimal approximations. Firstly for the s.i.f. estimates $\hat{k}_i^h$ define elements of the correction matrix D , d_{ij} such that:

$$d_{ij} = \left\langle \frac{\partial s_i}{\partial \nu}, \Pi_h s_{-i} \right\rangle_h - (\nabla s_j^h, \nabla \phi_i) \quad i, j = 1, 3 \quad (6.2.7)$$

	10	20	40	80	Obs.	Theory
$ k_1 - k_1^h $	35.10^{-3}	18.10^{-3}	9.10^{-3}	$4.5.10^{-3}$	ch	$ch^{1-\epsilon}$
$ k_1 - \hat{k}_1^h $	573.10^{-6}	144.10^{-6}	36.10^{-6}	9.10^{-6}	ch^2	$ch^{2-\epsilon}$
$ k_3 - k_3^h $	49.10^{-4}	36.10^{-4}	21.10^{-4}	11.10^{-4}	$\rightarrow ch$	$ch^{1-\epsilon}$
$ k_3 - \hat{k}_3^h $	192.10^{-5}	47.10^{-5}	12.10^{-5}	3.10^{-5}	ch^2	$ch^{2-\epsilon}$
$\ u - u^h\ _\infty$	0.053	0.040	0.030	0.021	$\rightarrow ch^{\frac{1}{2}}$	$ch^{\frac{1}{2}-\epsilon}$
$\ u - \hat{u}^h\ _\infty$	285.10^{-6}	67.10^{-6}	16.10^{-6}	4.10^{-6}	ch^2	$ch^{2-\epsilon}$

Table 6.1: Numerical Results for the Simple Crack Problem

and then calculate $\hat{k}_i^h$, $i = 1, 3$, such that

$$\begin{pmatrix} \hat{k}_1^h \\ \hat{k}_3^h \end{pmatrix} = (I - D)^{-1} \begin{pmatrix} k_1^h \\ k_3^h \end{pmatrix}. \quad (6.2.8)$$

Finally using these optimal s.i.f. estimates an optimal approximation $\hat{u}^h$ follows from:

$$\hat{u}^h = u^h + \hat{k}_1^h(s_1 - s_1^h) + \hat{k}_3^h(s_3 - s_3^h). \quad (6.2.9)$$

To demonstrate the performance of this method, the above procedure was implemented to solve for $\hat{k}_i^h$, $\hat{u}^h$ on a series of grids. Starting with a coarse grid of ten elements per side three refinements were performed with the finest partition having 80 elements per side. The expected convergences (see (5.4.19) and (5.4.20)) were obtained, the L_∞ error being measured only at nodes, as is shown in Table 6.1.

6.3 Approximation of the Model Fracture Problem

The model fracture problem has already been presented in section 4.3 with (4.3.3–6) being the defining equations. Also the basic solution procedure will be that outlined in (4.3.7–9) i.e. find approximations to u_0 and u_1 in (4.3.4) then using these to obtain an estimate of λ via (4.3.8) and finally use

(4.3.9) to give an approximation to the initial unknown u defined in (4.3.2). However from the description of the physical problem it is clear that its domain of definition Ω would lead to poor approximations if (4.3.7–9) were implemented. From the results of Chapter 4 the internal boundary Γ_n with its reentrant corners of angle 2π will lead to singularities in u_0 and u_1 and thus u_0^h and u_1^h given in (4.3.7) would exhibit the low convergence rates demonstrated in Chapter 5. It is thus proposed to use the DSFM to provide better approximations for these quantities and its implementation for this problem will now be described.

Firstly the second equation in (4.3.4) is chosen as the example on which the following theory is developed. This will be rewritten using the general notation of Chapter 4 thus; find $u_1 \in V_o(1)$ such that:

$$a(u_1, v) = 0 \quad \forall v \in V_o(0) \quad (6.3.1)$$

which is the weak form of the associated differential equation for u_1 i.e.:

$$\begin{aligned} -\nabla^2 u_1 &= 0 && \text{in } \Omega \\ u_1 &= 0 && \text{on } \partial\Omega \\ u_1 &= 1 && \text{on } \Gamma_n \end{aligned}$$

($\partial\Omega$ being the outer boundary of Ω and Γ_n the internal slit). Now it will be noticed that the boundary data is non-zero on the slit and thus at the reentrant corners and so, from section 5.5, the problem needs pre-processing if it is to have a form to which the DSFM may be applied. This is simply done by defining a new variable $u_2 = 1 - u_1$ whence (6.3.1) becomes: find $u_2 \in V_1(0)$ such that

$$a(u_2, v) = 0 \quad \forall v \in V_o(0). \quad (6.3.2)$$

Having defined the problem to be approximated the singular nature of the solution will be presented. With x_1 and x_2 the slit end points then, from the theory of Chapter 4, u_2 may be written in the vicinity of the reentrant corners thus:

$$u_2 = w_1 + \sum_{i=1}^n k_i s_i, \quad u_2 = w_2 + \sum_{i=1}^n l_i t_i. \quad (6.3.3)$$

For $n = 3$ w_j ($j = 1, 2$) will be contained in $W^{2,\infty}$ in the relevant regions (see (5.3.4)), with the s_i, t_i singularities of the form $r^{i/2} \sin(i\theta/2)$ centred on x_1, x_2 respectively. (The $r^2 \ln r$ singularities in general present at the rectangle's corners will not be found in this case as $u_2 = 1 - u_1$ and u_1 will not exhibit these singularities because the associated data is zero at and near these corners.) This then ensures sufficient regularity of the w_j for them to be optimally approximated using piecewise linear basis functions.

Before proceeding with an approximation of (6.3.2) using the DSFM it must be noted that the problem does not fall into the class for which this method was developed. Ω is no longer a simply connected region, the reentrant corners are present on an internal boundary and they number more than one. The main problems induced in this situation are that the s.i.f. integral formulation (4.5.12) will no longer hold (though (4.5.13) will of course) and a global singular decomposition analogous to (4.4.4) may not be defined. The singular/dual singular functions are no longer harmonic in Ω thus they cannot directly give rise to a singular decomposition. Also the integral identities in (4.5.6) will not be valid and an interpretation of the invariant s.i.f. integral (4.5.7) in terms of integrals over Ω and $\partial\Omega$ unobtainable. The former problem may be overcome by using cutoff functions with the singular functions and the latter by reformulating the s.i.f. formula over a region "local" to the associated singular point where singularity harmonicity is ensured. The local region may in fact be defined arbitrarily so long as it contains the corner and singularities maintain their harmonic behaviour. This allows a very compact formulation for s.i.f. estimation and with a global singular decomposition the main bulk of the theory of Chapter 5 is applicable.

To derive the described local s.i.f. formulation the first singular point x_1 will be considered and thus, with reference to (6.3.3), the first set of s.i.f.s $\{k_i\}_{i=1}^n$. A subdomain Ω_1 of Ω will be considered containing x_1 and an $O(1)$ neighbourhood but excluding x_2 , see Fig. 6.2. Using the theory of Chapter 4 an s.i.f. representation for k_i may be given (from (4.5.12)) by:

$$k_i = \frac{1}{i\pi} \left[\int_{\Omega_1} f s_{-i} - \int_{\partial\Omega_1} u \frac{\partial s_{-i}}{\partial \nu_i} + \int_{\partial\Omega_1} \frac{\partial u}{\partial \nu_1} s_{-i} \right] \quad (6.3.4)$$

with ν_1 the outward normal to Ω_1 and s_{-i} being the (same) d.s.f. associated with s_i . The only difference between this and (4.5.12) is that $\partial\Omega_1$ is no longer the boundary on which u is specified (though u will be specified on part of $\partial\Omega_1$, i.e. at least on that part coinciding with the reentrant corner). This however does not compromise the use of (6.3.4) to estimate s.i.f.^s as will now be demonstrated.

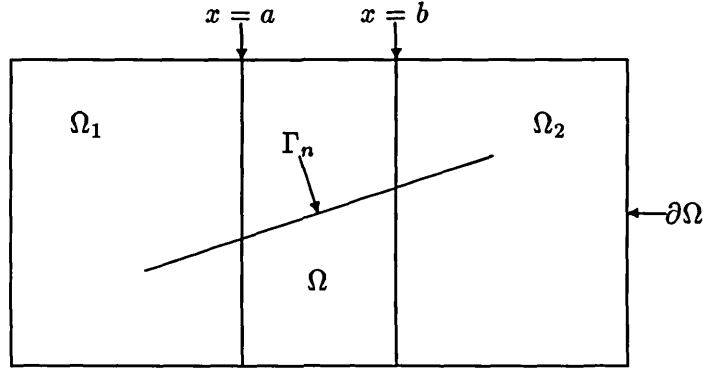


Figure 6.2: The idealised fracture domain and computational subdomains

The first stage of s.i.f. estimation is to use an approximation to u in (6.3.4). Following Chapter 5, with a finite element approximation $u^h \in S_g^h(\Omega)$ such that

$$(\nabla u^h, \nabla \chi) = (f, \chi) \quad \forall \chi \in S_o^h(\Omega)$$

(see (4.4.6)), define an s.i.f. estimate

$$k_i^h = \frac{1}{i\pi} \left[\int_{\Omega_1} f s_{-i} - \int_{\partial\Omega_1} u^h \frac{\partial s_{-i}}{\partial \nu_1} + (\nabla u^h, \nabla \phi_i)_{\Omega_1} - (f, \phi_i)_{\Omega_1} \right] \quad (6.3.5)$$

with $(\cdot, \cdot)_{\Omega_1}$ the L_2 inner product over Ω_1 and ϕ_i defined over Ω_1 (taken as a union of elements) as follows:

$$\begin{aligned} \phi_i &\in \partial S^h(\Omega_1) \text{ (where } S^h(\Omega_1) \equiv S_o^h(\Omega_1) \oplus \partial S^h(\Omega_1)) \\ \text{and } \phi_i &= \Pi_h s_{-i} \quad \text{on } \partial\Omega_1. \end{aligned} \quad (6.3.6)$$

This is obviously the analogue to (5.2.8) and an equivalent error formulation to that given in (5.2.9) is directly deriveable as follows. With ψ_i defined analogously to (5.2.1) thus:

$$\psi_i \equiv s_{-i} \text{ on } \partial\Omega_1 \text{ such that } \psi_i \in H^2(\Omega_1) \quad (6.3.7)$$

then from (6.3.4) and (6.3.5)

$$\begin{aligned} |k_i - k_i^h| &= \frac{1}{i\pi} \left| \int_{\partial\Omega_1} (u^h - u) \frac{\partial s_{-i}}{\partial \nu_1} + \int_{\partial\Omega_1} \frac{\partial u}{\partial \nu_1} s_{-i} - (\nabla u^h, \nabla \Pi_h \psi_i)_{\Omega_1} \right. \\ &\quad \left. + (f, \Pi_h \psi_i)_{\Omega_1} \right| \\ &= \frac{1}{i\pi} \left| \int_{\partial\Omega_1} (u^h - u) \frac{\partial s_{-i}}{\partial \nu_1} + \int_{\partial\Omega_1} \frac{\partial u}{\partial \nu_1} (s_{-i} - \Pi_h s_{-i}) \right. \\ &\quad \left. + (\nabla(u - u^h), \nabla \Pi_h \psi_i)_{\Omega_1} \right| \\ &= \frac{1}{i\pi} \left| \int_{\partial\Omega_1} \frac{\partial u}{\partial \nu_1} (s_{-i} - \Pi_h s_{-i}) - (\nabla(u - u^h), \nabla(\psi_i - \Pi_h \psi_i))_{\Omega_1} \right. \\ &\quad \left. + (u - u^h, -\nabla^2 \psi_i)_{\Omega_1} \right| \\ &\leq c \left\{ \|s_{-i} - \Pi_h s_{-i}\|_{0,\infty,\partial\Omega_1} + h|u - u^h|_{1,\Omega_1} + \|u - u^h\|_{0,\Omega_1} \right\} \quad (6.3.8) \end{aligned}$$

via the reasoning of the proof to Theorem 5.2.1. Theorem 5.2.2 may then be applied to these s.i.f. estimates indicating that using the basic approximations this local method provides the same error behaviour as its global predecessor.

Having now defined basic s.i.f. estimates the next step is to define a solution enhancement procedure. The same method as developed in Chapter 5 is inapplicable here as the required global singular decomposition using unmodified singular functions of the form (4.4.4) is invalid in this context. With the s_i and t_i having discontinuous gradients within Ω (repectively along the fracture line's continuation to the right and left) a global subtraction of the singularity sums defined in (6.3.3) would leave the residual $w \notin H^2(\Omega)$ and thus unable to be optimally approximated. For this reason a singular decomposition will require the use of cutoff functions though of a simpler form than the radial $\tau(r)$ used to define $\tilde{s}_{\pm i}$ in (4.5.11). Thus, with reference

to Fig. 6.3.1, returning to the problem (6.3.2) let u_2 be written as:

$$u_2 = w + \sum_{i=1}^3 k_i [\tau(x)s_i + l_i(1 - \tau(x))t_i] = w + \sum_{i=1}^3 k_i [\bar{s}_i + l_i \bar{t}_i] \quad (6.3.9)$$

where $w \in W^{2,\infty}(\Omega)$ and $\tau(x)$ is a one-dimensional cutoff function such that

$$\tau(x) = \begin{cases} 1 & x < a \\ 0 & x > b \end{cases} \quad (6.3.10)$$

where $x = a$ and $x = b$ are mesh lines. This then isolates the singular behaviour in u_2 and as will become apparent when considering solution enhancement will not lead to a large increase in computational effort. Firstly the order of polynomial required for $\tau(x)$ is less than for the radial cutoff function and secondly the s.i.f. estimate formula, with its use of unmodified dual singular functions and arbitrary local domain of definition, will require no more effort to evaluate, only in the singularity subtraction will an increase be required.

With a global singular decomposition for u_2 now defined the ideas of sections 5.3 and 5.4 may be applied to allow optimal approximations to be derived. Firstly a projection of the decomposition (6.3.9) will give a singular decomposition for u_2^h analogous to that in (5.3.2). The solution enhancement process may then proceed along the lines of 5.3 and 5.4 after calculating the projections of $\bar{s}_i$ and $\bar{t}_i$, $\bar{s}_i^h$ and $\bar{t}_i^h$ ($i = 1, 3$). It is this process which gives rise to extra computational effort as the singular functions are no longer harmonic in Ω and so a right hand side must be calculated to obtain their projections e.g. (see (4.4.11)) the right hand side of the defining equation for $\bar{s}_i^h$:

$$(\nabla \bar{s}_i^h, \nabla \chi) = (\nabla \bar{s}_i, \nabla \chi) \equiv (-\nabla^2 \bar{s}_i, \chi) \quad \forall \chi \in S_o^h(\Omega).$$

This also demonstrates the reduced smoothness requirement on $\tau(x)$ over that derived in (5.5.9) for the radial cutoff function. Since $\tau(x)$ is aligned with the mesh a quadrature approximation with sufficient accuracy only requires $-\nabla^2 \bar{s}_i \in H^2(\Delta)$ for each triangle Δ . Thus the requirement on τ is simply

$$\tau \in C^1(0, 1) \text{ with } \tau(a) = 1, \tau'(a) = \tau(b) = \tau'(b) = 0$$

which is satisfied by having τ merely cubic in x .

Once having obtained $\bar{s}_i^h, \bar{t}_i^h$ the basic solution enhancement using s.i.f. estimates, i.e. analogous to (5.3.6), follows with an analogy to Theorem 5.3.1 directly obtainable. Then using the new arbitrary s.i.f. estimate (6.3.5) the process of obtaining optimal solutions given in 5.4 may be paralleled. An approximation $\hat{u}_1^h$ can be defined as:

$$\hat{u}_2^h = u_2^h + \sum_{i=1}^3 \hat{k}_i^h(\bar{s}_i - \bar{s}_i^h) + \hat{l}_i^h(\bar{t}_i - \bar{t}_i^h) \quad (6.3.11)$$

where the $\hat{k}_i^h, \hat{l}_i^h$ are derived from a formula analogous to (5.4.17) with typical correction factor matrix entries being:

$$\begin{aligned} (\nabla(\bar{s}_j - \bar{s}_j^h), \nabla\phi_i)_{\Omega_1} &= \int_{\partial\Omega_1} \frac{\partial\bar{s}_j}{\partial\nu_i} \Pi_h s_{-i} - (\nabla\bar{s}_j^h, \nabla\phi_i)_{\Omega_1} \\ (\nabla(\bar{t}_j - \bar{t}_j^h), \nabla\phi_i)_{\Omega_1} &= -(\nabla\bar{t}_j^h, \nabla\phi_i)_{\Omega_1}. \end{aligned}$$

These examples assume that Ω_1 is contained in the region where $\tau(x) = 1$ and is identical with the expression derived in (5.4.7) (except being defined over a subregion of Ω). Error bounds of the form given in Theorem 5.4.2 may then be developed for the resulting approximations as (6.3.8) is equivalent to the abstract bound furnished by Theorem 5.2.1 and as just discussed an analogy to Theorem 5.3.1 for the approximation (6.3.11) is available.

The application of the DSFM to the model fracture problem is based on the local s.i.f. estimator, (6.3.5). A numerical experiment was thus performed to demonstrate the convergence of s.i.f. estimates using this formulation. The problem chosen was (6.3.2) and as no analytic solution is available in this case the results presented in Table 6.2 give the actual estimates though it is easily seen that they are behaving as expected. The experiment was performed with Ω_1 a region an eighth of the size of Ω and with $\partial\Omega_1$ not coinciding at all with $\partial\Omega$. The discretization of Ω was similar to that used in 6.2.

Finally to return to the problem this section was designed to tackle the question of approximations to u from (4.3.1-2), i.e. the total flux problem solution, must be addressed. u_o , as defined in (4.3.7), may obviously be

	10	20	40	80	Obs.	Theory
k_1^h	2.3731	2.2859	2.2459	2.2266	$\sim ch$	$ch^{1-\epsilon}$
k_2^h	0.1513	0.2720	0.3180	0.3380	$> ch$	$ch^{1-\epsilon}$

Table 6.2: Basic Stress Intensity Factor Estimates for the Fracture Problem

approximated in a similar maner to $1 - u_1$ and will give rise to an optimal $\hat{u}_o^h$. So all that remains to do then is to construct the estimate $\hat{\lambda}^h$. For this (4.3.10b) will be used with $\hat{u}_o^h$ and $1 - \hat{u}_2^h$. This will reduce the computational effort in evaluating the integrals over the alternative (4.3.10a) as the $\bar{s}_i$ and $\bar{t}_i$ within $\hat{u}_o^h$ and $\hat{u}_2^h$ are smooth in the outer boundary skin where the integration is performed. Thus the approximation $\hat{\lambda}^h$ is given by

$$\hat{\lambda}^h = \frac{q + a(\hat{u}_o^h, \eta_2^h)}{a(\hat{u}_2^h, \eta_2^h)},$$

and from the theory of 4.3 this will be optimal and an optimal approximation to u , $\hat{u}^h$ will follow as:

$$\hat{u}^h = \hat{u}_o^h + \hat{\lambda}^h \hat{u}_1^h. \quad (6.3.12)$$

Part III

A Chemical Flood Model

Chapter 7

A Modified Roe's Scheme for a Parabolically Degenerate System of Hyperbolic Conservation Laws

7.1 Introduction

This chapter departs from the work presented until now in two respects; namely a complete idealised (i.e. one-dimensional) model of a chemical flood in a reservoir is considered and the resultant system contains no pressure equation. The model gives rise to a system of hyperbolic p.d.e.s whose derivation will be briefly described in the following section. The system is then analysed and an analytic solution constructed following solutions derived for similar situations in Pope [Pope 1980] and Isaacson [Isaacson 1985]. Following on from this a numerical method is derived from Roe's anti-diffusive flux scheme for conservation laws [Roe 1981a] and then finally this is implemented with the results compared against both the analytic solution and an approximation using basic upwinding.

7.2 The Physical Problem and Modelling Equations

Chemical flooding is an enhanced oil recovery (e.o.r.) technique designed to increase production efficiency by making the oil more mobile (i.e. σ in (1.2.1) is increased). Physically the effect is to reduce the minimum oil saturation level (below which water is not able to displace oil due to surface tension effects) which then becomes a function of the introduced chemical's concentration. The equations modelling this process are now derived for a general reservoir and then from these the system to be analysed is obtained.

Recalling the general outline of a reservoir model given in 1.2 there are two basic constituents, a pressure equation and mass conservation or continuity equation(s). Continuity of the aqueous phase gives:

$$-\nabla \cdot (\rho_w \mathbf{v}_w) + q_w = \frac{\partial}{\partial t}(\phi \rho_w S_w) \quad (7.2.1)$$

where ρ_w is the water density, $\mathbf{v}_w$ its velocity, q_w the injection/production rate, ϕ rock porosity, and S_w the phase saturation (an analogous equation for the oil phase may obviously be derived). A chemical continuity equation follows in exactly the same manner as:

$$-\nabla \cdot (c \rho_w \mathbf{v}_w) + c q_w = \frac{\partial}{\partial t}(\phi c \rho_w S_w) \quad (7.2.2)$$

where c is the concentration and the chemical is assumed to remain in the water phase. Now returning to the pressure it is noted that an equation of the form (1.2.1) (i.e. Darcy's law) is required for both oil and water. The mobility σ is influenced by the presence of a second phase and so separate 'phase mobilities' are defined for each phase. Also the pressures in the phases are generally unequal due to capillary pressure with the fluid velocities being given by:

$$\begin{aligned} \mathbf{v}_w &= -\sigma_w(\nabla p_w - \rho_w \nabla z) \\ \mathbf{v}_o &= -\sigma_o(\nabla p_o - \rho_o \nabla z) \end{aligned} \quad (7.2.3)$$

where subscript o denotes oil associated quantities and p_o and p_w are related by the capillary pressure p_c thus:

$$p_c = p_o - p_w. \quad (7.2.4)$$

The preceeding four equations give in full generality a description of the chemical flood. However before formulating a mathematical model from these equations certain simplifying assumptions will be made. The densities and porosities are assumed to be constant and the sources to be zero thus from conservation:

$$\begin{aligned} -\nabla \cdot \mathbf{v}_w &= \phi \frac{\partial S_w}{\partial t} \\ -\nabla \cdot \mathbf{v}_o &= \phi \frac{\partial S_o}{\partial t} \end{aligned} \quad (7.2.5)$$

and

$$-\nabla \cdot (c\mathbf{v}_w) = \phi \frac{\partial}{\partial t}(cS_w). \quad (7.2.6)$$

Also it will be assumed that capillary pressure is sufficiently small so that its gradient is negligible and there are no depth variations, thus (7.2.3) becomes:

$$\begin{aligned} \mathbf{v}_w &= -\sigma_w \nabla p \\ \mathbf{v}_o &= -\sigma_o \nabla p. \end{aligned} \quad (7.2.7)$$

It is now convenient to consider the sum of each individual phase velocity $\mathbf{v}_t$

$$\mathbf{v}_t = \mathbf{v}_w + \mathbf{v}_o \quad (7.2.8)$$

and then from (7.2.5) (noting $S_o + S_w = 1$) the standard incompressibility condition results:

$$\nabla \cdot \mathbf{v}_t = 0.$$

With this and rewriting $\mathbf{v}_w$ in terms of $\mathbf{v}_t$ from (7.2.7–8), (7.2.5–6) become

$$\begin{aligned} \frac{\partial S_w}{\partial t} + \frac{\mathbf{v}_t}{\phi} \cdot \nabla f_w &= 0 \\ \frac{\partial(cS_w)}{\partial t} + \frac{\mathbf{v}_t}{\phi} \cdot \nabla(cf_w) &= 0 \end{aligned} \quad (7.2.9)$$

where $f_w = \frac{\sigma_w}{\sigma_w + \sigma_o}$. This system with associated boundary conditions will give a full description of the behaviour of the fluid and chemical concentrations.

From (7.2.9) the actual chemical flood model which will be analysed in the rest of this chapter may now be derived. A one-dimensional situation is

considered and after normalising (7.2.9) becomes:

$$\begin{aligned}\frac{\partial u}{\partial t} + \frac{\partial f}{\partial x} &= 0 \\ \frac{\partial(cu)}{\partial t} + \frac{\partial(cf)}{\partial x} &= 0\end{aligned}\tag{7.2.10}$$

where u is aqueous phase saturation and c the chemical concentration. The phase mobility (or flux) function f will have the form:

$$f(u) = \frac{(u - u_1)^2}{(u - u_1)^2 + d(u_2(c) - u)^2}\tag{7.2.11}$$

where $u_2(c)$ is the maximum water saturation (i.e. one minus minimum oil saturation), with c taking values between 0 and 1. u_1 is the minimum water saturation and d a parameter dependent on the fluids' relative viscosities and permeabilities. The function $u_2(c)$ is taken to be linear with a maximum value of 1 (i.e. complete absence of oil) and minimum value u_o . Initial and boundary conditions for the system will be derived in the next section from an assumption of a semi-infinite (linear) reservoir containing a single well injecting at the origin.

7.3 Analysis of the Model Equations and Solution of the Riemann Problem

The system (7.2.10) may be rewritten in conservation law form by setting $v = uc$ i.e.

$$\frac{\partial u}{\partial t} + \frac{\partial f}{\partial x} = 0\tag{7.3.1a}$$

$$\frac{\partial v}{\partial t} + \frac{\partial}{\partial x} \left(\frac{v}{u} f \right) = 0\tag{7.3.1b}$$

which in the equivalent vector form is:

$$\mathbf{u}_t + \mathbf{F}_x = 0 \text{ where } \mathbf{u} = \begin{pmatrix} u \\ v \end{pmatrix}, \mathbf{F} = \begin{pmatrix} f \\ \frac{v}{u} f \end{pmatrix}.\tag{7.3.2}$$

In general solutions to (7.3.2) will contain discontinuities and for this situation the weak form and its associated solution must be considered (see [Lax 1957] for a full treatment).

With the problem defined on a semi-infinite interval two conditions are needed to complete its definition. An initial state is needed for the reservoir which typically takes the following form

$$\mathbf{u}_r = \begin{pmatrix} u_r \\ v_r \end{pmatrix} = \begin{pmatrix} u_I \\ 0 \end{pmatrix} \quad (7.3.3)$$

i.e. a given constant aqueous saturation and zero chemical. Also a left hand boundary condition is required which will take the form:

$$\mathbf{u}(0, t) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (7.3.4)$$

i.e. an injection of pure aqueous phase with maximum chemical concentration (noting $u_2(1) = 1$). This is also equivalent to having the left half line as a constant vector state i.e.

$$\mathbf{u}_l = \begin{pmatrix} u_l \\ v_l \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (7.3.5)$$

and so the problem may be treated as a pure initial value problem. This is an example of a Riemann problem (see [Lax 1957]) with the single condition:

$$\mathbf{u}(x, 0) = \begin{cases} \begin{pmatrix} u_I \\ 0 \end{pmatrix} & x > 0 \\ \begin{pmatrix} 1 \\ 1 \end{pmatrix} & x \leq 0. \end{cases} \quad (7.3.6)$$

With a complete problem definition now given analysis of the equations is now begun via the quasi-linear form of (7.3.2) i.e.

$$\mathbf{u}_t + A(\mathbf{u})\mathbf{u}_x = 0 \text{ where } A(\mathbf{u}) = \begin{pmatrix} \frac{\partial f_i}{\partial u_j} \end{pmatrix} \quad (7.3.7)$$

with $\mathbf{F} = \{f_i\}^T$ and $\mathbf{u} = \{u_i\}^T$. The system is hyperbolic if the Jacobian matrix $A(\mathbf{u})$ has real and distinct eigenvalues for all values of $\mathbf{u}$. Two degeneracies from this state are non-strict hyperbolicity and parabolic degeneracy. The former occurs when a linearly degenerate field exists, this being where the gradient of an eigenvalue λ_i is perpendicular to its eigenvector e_i i.e.

$$(\nabla \lambda_i) \cdot e_i = 0. \quad (7.3.8)$$

Parabolic degeneracy is a more serious deviation and occurs when coincident eigenvalues lead to the Jacobian matrix being non-diagonalisable.

An important feature of these non-linear systems is that in general continuous solutions will not exist and the weak form of the problem must be considered in order to admit the necessary discontinuous solutions. These have to satisfy the integral form of the conservation law from which can be deduced the Rankine–Hugoniot jump conditions (see [Lax 1957]) that any jump must satisfy:

$$S[\mathbf{u}] = [\mathbf{F}] \quad (7.3.9)$$

where S is the speed of the discontinuity and $[x]$ denotes the associated (discontinuous) change in quantity x . The class of weak solutions satisfying (7.3.9) is infinite in number and the uniqueness property of smooth solutions is lost. However only one of these weak solutions is physically relevant and may be identified via the entropy conditions (which will be introduced later).

Fundamental to an analytic investigation of the system are the eigenvalues and eigenvectors of the Jacobian matrix A of the quasilinear form (7.3.7). The eigenvectors define directions along which components of a solution $\mathbf{u}$ can travel and the associated eigenvalue is the speed with which each component moves. Each eigenvector describes a vector field in the solution space $(u_1, \dots, u_n)$ of $\mathbf{u}$ i.e. where trajectories, or integral curves, of this field are paths followed by a point moving in the direction of the eigenvector. The eigenvalues in turn define trajectories in the variable space, i.e. the (x, t) plane, known as characteristics, which are curves satisfying:

$$\frac{\partial x}{\partial t} = \lambda_i \quad (7.3.10)$$

where λ_i is a function of $\mathbf{u}$ which is itself, of course, a function of x and t . Thus from a point in the variable space where a solution vector $\mathbf{u}$ is defined n characteristics will issue each with a gradient that is the inverse of the eigenvalue associated with the particular characteristic.

The eigenvalues and eigenvectors can be further used to indicate the behaviour of linearly degenerate fields and to define special quantities known as Riemann invariants. Linear degeneracy occurs when the eigenvalue gradient is normal to its eigenvector i.e. the eigenvector is parallel to surfaces

$\lambda_i = C$ which then contain the integral curves. Therefore a solution derived from this field will have constant speed. A k -Riemann invariant $v_k = v_k(\mathbf{u})$ is a scalar quantity such that:

$$\mathbf{r}_k \cdot \nabla v_k = 0 \quad (7.3.11)$$

for all values of $\mathbf{u}$, where $\mathbf{r}_k$ is the k^{th} (right) eigenvector. The Riemann invariants are constant in the region of variable space (the (x, t) plane) adjacent to a region where the solution $\mathbf{u}$ is constant (see [Lax 1957]) and here the solution is known as a simple wave. Also in a two-component system lines upon which the Riemann invariants are constant will be parallel to the corresponding eigenvector i.e. they will be integral curves in the solution space (and thus will represent solutions).

Finally the characteristics are used to obtain the entropy conditions from which the physically relevant solution can be identified. The requirement is simply that a solution must have evolved along characteristics from the initial state. In the discontinuous solution case this says that from any shock line in the variable plane a characteristic corresponding to each eigenvalue may be drawn back from this shock to the initial line with the second characteristic of one of the fields (i.e. the shock field) also developing from the initial line. This follows directly from the requirement of the shock field eigenvalues to satisfy

$$\lambda_k(\mathbf{u}_l) > S > \lambda_k(\mathbf{u}_r)$$

across the shock and then hyperbolicity ensuring that (with ordering of the eigenvalues)

$$\begin{aligned} \lambda_j(\mathbf{u}_l) &< \lambda_k(\mathbf{u}_l) & j < k \\ \lambda_j(\mathbf{u}_r) &> \lambda_k(\mathbf{u}_r) & j > k. \end{aligned}$$

For a shock in a genuinely non-linear field the entropy condition then states that in an n -component system $n + 1$ must impinge on the shock line in the (x, t) plane. In a linearly degenerate field as the eigenvalue associated with the field is continuous across the shock, i.e. $\lambda(\mathbf{u}_l) = \lambda(\mathbf{u}_r)$, and its value is the shock speed then the two associated characteristics are obviously coincident

with the shock. In this case the obvious modification is then to require that of the remaining characteristics $n - 1$ must impinge on the shock line.

Returning to the problem (7.3.5) the quasilinear form is:

$$\mathbf{u}_t + \begin{pmatrix} f_u & f_v \\ \frac{v}{u}f_u - \frac{v}{u^2}f & \frac{1}{u}f - \frac{v}{u}f_v \end{pmatrix} \mathbf{u}_x = 0 \quad (7.3.12)$$

from which the Jacobian matrix

$$A(\mathbf{u}) = \begin{pmatrix} f_u & f_v \\ \frac{v}{u}f_u - \frac{v}{u^2}f & \frac{1}{u}f - \frac{v}{u}f_v \end{pmatrix}. \quad (7.3.13)$$

The eigenvalues are then the solutions of the characteristic equation:

$$\lambda^2 - \frac{1}{u}(f + uf_u + vf_v) + \frac{f}{u}(f_u + \frac{v}{u}f_v) = 0 \quad (7.3.14)$$

i.e.

$$\lambda_1 = \frac{f}{u} \text{ and } \lambda_2 = f_u + \frac{v}{u}f_v, \quad (7.3.15)$$

and from these can be deduced the respective eigenvectors:

$$\mathbf{e}_1 = \begin{pmatrix} f_v \\ \frac{f}{u} - f_u \end{pmatrix} \text{ and } \mathbf{e}_2 = \begin{pmatrix} u \\ v \end{pmatrix}. \quad (7.3.16)$$

The non-strict hyperbolicity of the system follows as

$$\mathbf{e}_1 = u \begin{pmatrix} \frac{f_v}{u} \\ \frac{f}{u^2} - \frac{f_u}{u} \end{pmatrix} = u \begin{pmatrix} \frac{\partial}{\partial v} \left(\frac{f}{u} \right) \\ -\frac{\partial}{\partial u} \left(\frac{f}{u} \right) \end{pmatrix} \quad (7.3.17)$$

and as $\lambda_1 = f/u$ then the field is linearly degenerate i.e.

$$\nabla \lambda_1 \cdot \mathbf{e}_1 = 0. \quad (7.3.18)$$

To show the parabolic degeneracy properties of the actual flux function are required to demonstrate that eigenvalue equality can occur in this system. This will be done while constructing a solution at the end of this section.

To define the admissible shock waves of the system the Rankine-Hugoniot jump condition must be applied. From (7.3.9) the shock speed S satisfies:

$$S(u_l - u_r) = f(\mathbf{u}_l) - f(\mathbf{u}_r) \quad (7.3.19a)$$

$$S(v_l - v_r) = \frac{v_l}{u_l} f(u_l) - \frac{v_r}{u_r} f(u_r) \quad (7.3.19b)$$

and when eliminated the shock conditions follow:

$$\frac{f_l}{u_l} = \frac{f_r}{u_r} \text{ or } \frac{v_r}{u_r} = \frac{v_l}{u_l}. \quad (7.3.20)$$

Substituting (7.3.20) into (7.3.19) then gives the respective shock speeds as

$$S = \frac{f_l}{u_l} = \frac{f_r}{u_r} \text{ and } S = \frac{f_l - f_r}{u_l - u_r}. \quad (7.3.21)$$

The first of these conditions and speeds confirms the first characteristic field degeneracy i.e. the shocks (known as contact discontinuities) will move with speed λ_1 (which is thus continuous across the shock as required).

Finally the Riemann invariants, the quantities v_i ($i = 1, 2$) satisfying (7.3.11), will be derived. In general $n - 1$ independent invariants will be deriveable for each characteristic field but in the two-component system case this simplifies to a single invariant. The first of these was constructed when showing the degeneracy of the first field as (7.3.18) is equivalent to (7.3.11) with $v = \lambda_1 = f/u$ so

$$v_1 = \frac{f}{u}. \quad (7.3.22a)$$

For the second field invariant note that:

$$\mathbf{e}_2 = \begin{pmatrix} u \\ v \end{pmatrix} = u^2 \begin{pmatrix} \frac{1}{u} \\ \frac{v}{u^2} \end{pmatrix} = u^2 \begin{pmatrix} \frac{\partial}{\partial v} \left(\frac{v}{u} \right) \\ -\frac{\partial}{\partial u} \left(\frac{v}{u} \right) \end{pmatrix}$$

from which it follows directly that:

$$v_2 = \frac{v}{u}. \quad (7.3.22b)$$

The invariants then define the class of solutions known as simple waves i.e. the families of curves defined by $v_1 = k_1$ and $v_2 = k_2$. These then define the integral curves of the first and second eigenvectors respectively (in a higher order system there would be more than one Riemann invariant corresponding to each field with each invariant defining a surface) and thus the general class of solutions for the problem. The resulting waves will take the form of either smooth solutions (rarefaction waves) or shocks.

Finally to identify the physically relevant solution an interpretation of the entropy conditions for this system is required. Noting that the first field is linearly degenerate for shocks to be admissible one characteristic must impinge on (or enter) the shock line. With the second field genuinely non-linear then three characteristics must impinge on the shock line.

Before developing a framework for and constructing a solution it is convenient to reinterpret the quantities derived for this process in terms of the initial system variables u and c (see (7.2.10)). Firstly a new 'flux' function g is defined as

$$g(u, c) = f(u, v) = f(u, uc) \quad (7.3.23)$$

using which the eigenvalues are given from (7.3.15) as follows:

$$\lambda_1 = \frac{f}{u} = \frac{g}{u} \text{ and } \lambda_2 = f_u + \frac{v}{u} f_v = f_u + c f_{(uc)}.$$

The second eigenvalue is then able to be rewritten thus:

$$\lambda_2 = f_u + \frac{v}{u} f_v = f_u + c f_{(uc)} = f_u + \frac{\partial(uc)}{\partial u} f_{(uc)} = \frac{\partial}{\partial u} (f(u, uc)) = g_u$$

and therefore:

$$\lambda_1 = \frac{g}{u} \text{ and } \lambda_2 = g_u. \quad (7.3.24)$$

This is one valuable simplification i.e. the second eigenvalue is simply the gradient of the flux function when graphed against u at constant c . A second simplification is seen when considering the Riemann invariants i.e. with $v/u = c$ these reduce to:

$$v_1 = \frac{g}{u} \text{ and } v_2 = c. \quad (7.3.25)$$

The graph of g against u at different values of c shows the usefulness of this approach. First of all the curves themselves represent the second family of Riemann invariants and the straight lines issuing from the origin with gradients k , i.e. so that $g/u = k$, represent both the first family of Riemann invariants and the first family of eigenvalues. As has been noted already the second eigenvalue is simply the gradient of these flux curves and from this observation the parabolic degeneracy can be easily recognised. It occurs

when the straight lines representing the first eigenvalues are tangent to these flux curves and so the eigenvalues are equal. This then defines a line on the graph of g against u (see Fig. 7.1), designated T the transition curve, and is where the inequality between the two eigenvalues changes direction.

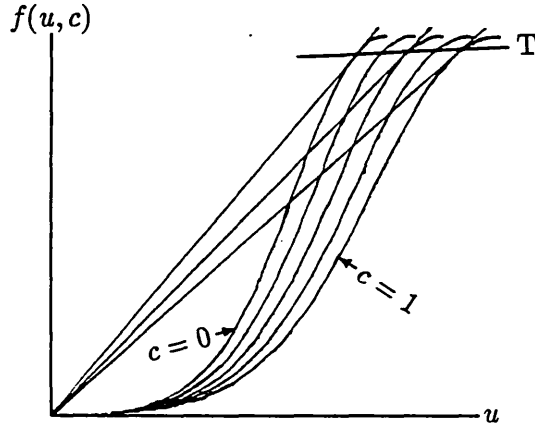


Figure 7.1: Riemann invariants in the (u, f) plane

Now to construct a solution to a given initial value problem the integral curves defined by v_1 and v_2 constant are required and these will be graphed in the transformed solution space (u, c) as shown in Fig. 7.2. Here again the transition curve T is given which is the boundary between the two regions where $v_2 = k$ is monotone in c i.e. where a given value of c can be matched with a unique value of u on a particular invariant curve. The simple wave solutions lie along the lines of Fig. 7.2 and thus a solution can be constructed by taking segments of these curves to form a continuous line in the (u, c) plane which represents a series of simple waves.

To determine which particular combination of simple waves gives the physical solution to the problem the entropy conditions are called upon which can also be formulated in terms of the flux function and integral curve diagrams of Figs. 7.1 and 7.2. These will pick out the admissible shock waves which can then be used to identify the physical solution. For

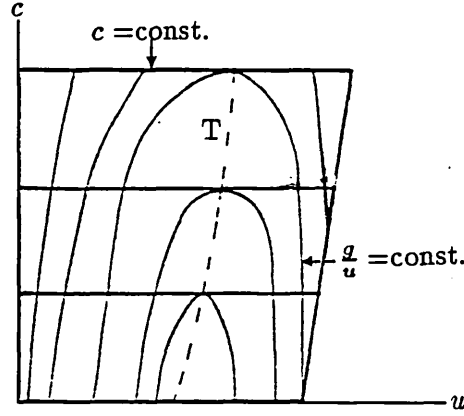


Figure 7.2: Riemann invariants in the (u, c) plane

degenerate field shocks two of the characteristics already coincide with the shock line so only the characteristic pair from the second (non degenerate) field is required to determine admissibility. The characteristic gradient is given by the inverse of the eigenvalue g_u , and for precisely one of the two characteristics to coincide with the shock both of their gradients need either to be less or greater than that of the degenerate field eigenvalue g/u . This will not be satisfied if the second field eigenvalues are derived from states either side of the transition curve (see Fig. 7.1). Thus the states must be contained within one of the two regions in the (u, c) plane defined by T . If one state should lie on T then its second eigenvalue will obviously coincide with the shock line thus the other state should not impinge on the shock line for admissibility. Thus if the off- T state is u_l then the gradient of $\lambda_2(u_l)$ must be less than that of the shock line, $\lambda_1(u_r) = \lambda_1(u_l) = \lambda_2(u_r)$ i.e. u_l must lie to the right of T in the solution space. Conversely if the off T state is u_r then it must lie to the left of T in the (u, c) plane.

For second field shocks consideration has to be given to all four characteristics and the entropy condition is most conveniently interpreted in terms of the flux function as graphed in Fig. 7.1. A shock in this field will join

two points on one of the (constant concentration) flux curves with its speed being the gradient of the connecting line (or flux curve shock line – Fig. 7.1), see (7.3.21). From this it can be seen that the first field characteristics will have slopes both greater or less than the flux curve shock line (i.e. the three lines join in a triangle where the two characteristic lines issue from the origin and the shock line joins them so it must have a slope greater than or less than both). Thus only one of these characteristics will impinge on the shock line and so for shock admissibility both of the second field characteristics must impinge on the shock line in variable space. For this to happen the flux shock line must lie below the flux curve if $u_l < u_r$ and above if $u_l > u_r$. This then guarantees that the characteristic slopes are respectively greater and less than the shock line slope which is necessary and that they are situated correctly in order to impinge on the shock line in the variable space. The special case in which the two points are joined by a single first field characteristic line is the inadmissible case given in the contact discontinuity situation described above.

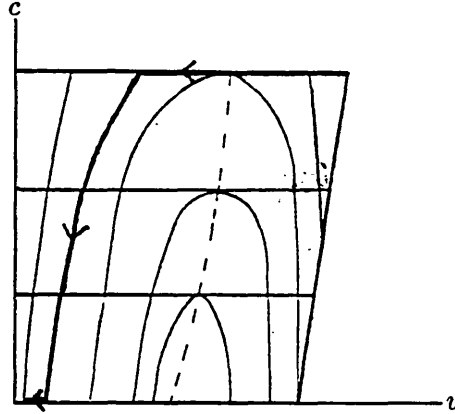


Figure 7.3: Solution path in the (u, c) plane

This application of the theory gives all the information for the construction of a solution to the Riemann problem. A physically interesting problem

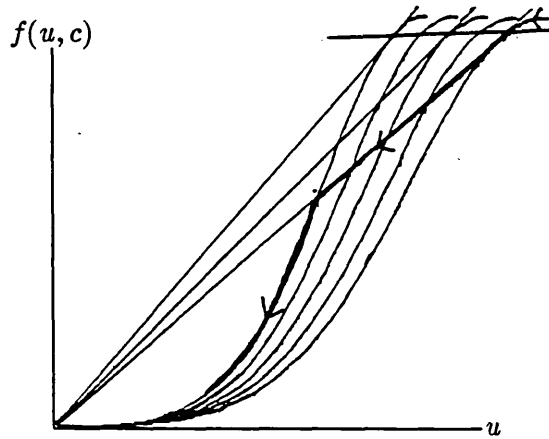


Figure 7.4: Solution path in the (u, f) plane

is one where the state u_l is given maximum values and u_r has close to minimum values. With reference to Figs. 7.3 and 7.4 the general idea is to trace a path between the states u_l and u_r along the lines defining simple waves, and then the resulting group of wave sections thus defined will form the basis of the solution. In practical terms this solution path for the given situation is arrived at in the following way (see Fig. 7.3). u_l is initially joined to an intermediate state u_1 by a 2-wave (here a rarefaction wave) which then joins with a second intermediate state u_2 along a 1-wave (a contact discontinuity) and then joins to the right hand state u_r along a second 2-wave (this time a shock wave). As can be seen this combination satisfies the necessary entropy conditions and so is the physical solution. By following the resultant path on the flux curve diagram, Fig. 7.4 the solution is seen to be physically reasonable i.e. while moving from left to right each subsequent portion of the solution moves out with increasing speed. These speeds are of course not continuous and thus to complete the solution between these simple waves the resultant spaces are then filled by constant states. Examples of solutions are given in Figs. 7.6–10 where they are graphed with numerical solutions to the problems.

7.4 Numerical Methods for the Chemical Flood System

The numerical methods used to approximate the system will be based on upwinding i.e. obtaining a numerical representation of the equation which incorporates information about the direction of motion of the variables. The value of this approach can be seen by considering the first order discretisations of the flux derivative f_x at (x_i, t)

$$\begin{aligned} f_x(x_i, t) &= \frac{f(x_i, t) - f(x_{i-1}, t)}{x_i - x_{i-1}} + O(h) \\ &= \frac{f(x_i, t) - f(x_{i+1}, t)}{x_i - x_{i+1}} + O(h). \end{aligned} \quad (7.4.1)$$

Now though these have the same local accuracy if $f_u(x_i, t) > 0$ (i.e. flow direction positive) then the first of these will be more accurate as here information in the system will be carried from left to right. When approximating (7.3.2) in scalar form in this situation the upwinding discretization will give an approximation thus:

$$u(x_i, t_{j+1}) = u(x_i, t_j) - \frac{t_{j+1} - t_j}{x_i - x_{i-1}} (f(x_i, t_j) - f(x_{i-1}, t_j)) \quad (7.4.2)$$

i.e. mimicing the physical behaviour by determining the flux in the cell $[x_{i-1}, x_i]$ to first order and adding it to the grid point value towards which it is moving.

In the case of systems such a simple approach is not possible as flow is not just positive or negative. Flow (at a point) then takes place along vectors in variable space, the Jacobian matrix eigenvectors. However by considering the quasi-linear form of the scalar equation, a suitable discretization of the flux term is, with $u_i = u(x_i, t)$,

$$f_u u_x \simeq \frac{f_i - f_{i-1}}{u_i - u_{i-1}} \frac{u_i - u_{i-1}}{x_i - x_{i-1}} \quad (7.4.3)$$

i.e. precisely the flux term in (7.4.2) and a combination of discretizations of the "one-dimensional Jacobian matrix" f_u i.e. $\frac{f_i - f_{i-1}}{u_i - u_{i-1}}$, and u_x . The flux function difference

$$\Delta f_{i-\frac{1}{2}} = f_i - f_{i-1}$$

can then be written as

$$\Delta f_{i-\frac{1}{2}} = \frac{f_i - f_{i-1}}{u_i - u_{i-1}} \Delta u_{i-\frac{1}{2}} \quad (7.4.4)$$

and is naturally extended to systems thus

$$\Delta \mathbf{F}_{i-\frac{1}{2}} = A^h(\mathbf{u}_i, \mathbf{u}_{i-1}) \Delta \mathbf{u}_{i-\frac{1}{2}} \quad (7.4.5)$$

where A^h is a discretization of the Jacobian matrix $A(u)$. In the extension (7.4.5) the flow directions are then identified as the eigenvectors of A^h and from hyperbolicity these will form a basis for the solution space thus the vector $\Delta \mathbf{u}_{i-\frac{1}{2}}$ may be expanded in terms of the eigenvectors. From this a decomposition of the flux difference in terms of the eigenvectors follows directly and an interpretation of this decomposition using the idea of upwinding gives rise to Roe's flux-splitting scheme for systems [Roe 1981b].

The process operates as follows. Using the eigenvalues $\mathbf{e}_1^{i-\frac{1}{2}}, \dots, \mathbf{e}_n^{i-\frac{1}{2}}$ and associated eigenvectors $\lambda_1^{i-\frac{1}{2}}, \dots, \lambda_n^{i-\frac{1}{2}}$ of $A^h(\mathbf{u}_i, \mathbf{u}_{i-1})$ write

$$\Delta \mathbf{u}_{i-\frac{1}{2}} = \sum_{k=1}^n a_k^{i-\frac{1}{2}} \mathbf{e}_k^{i-\frac{1}{2}} \quad (7.4.6)$$

and in view of (7.4.5) the flux difference then follows as:

$$\Delta \mathbf{F}_{i-\frac{1}{2}} = \sum_{k=1}^n a_k^{i-\frac{1}{2}} \lambda_k^{i-\frac{1}{2}} \mathbf{e}_k^{i-\frac{1}{2}}. \quad (7.4.7)$$

Now to upwind these flux difference components as with scalar upwinding note is taken of the sign of the eigenvalues (i.e. direction of motion of the quantity) in order to assign the relevant component to the grid point value towards which it is moving. The resulting upwinding difference equation for (7.3.2) is thus written as:

$$\mathbf{u}_i^{j+1} = \mathbf{u}_i^j - \frac{\Delta t}{h} \left(\sum_{(\lambda_k^{i-\frac{1}{2}} > 0)} a_k^{i-\frac{1}{2}} \lambda_k^{i-\frac{1}{2}} \mathbf{e}_k^{i-\frac{1}{2}} + \sum_{(\lambda_k^{i+\frac{1}{2}} < 0)} a_k^{i+\frac{1}{2}} \lambda_k^{i+\frac{1}{2}} \mathbf{e}_k^{i+\frac{1}{2}} \right) \quad (7.4.8)$$

using (7.4.6) and a similar decomposition for $\Delta \mathbf{u}_{i+\frac{1}{2}}$, and where Δt and h are the (uniform) temporal and spatial partition spacings.

Having now described the procedure for obtaining an upwind approximation to (7.3.2) using a Roe decomposition a generalisation of the Roe scalar anti-diffusive flux scheme to systems is now described. Its main aim is to improve the inherent first order accuracy of upwinding by reducing the second order error term which is diffusive and accounts for the smearing of shocks. However, eliminating the second order error term, as done in the Lax-Wendroff and Warming and Beam second order schemes, leads to unrealistic oscillations in the solution near discontinuities and non-entropy satisfying solutions being obtained. Roe has devised a monotone scheme [Roe 1981a] i.e. one that has the effect of introducing no new local maxima or minima (thus outlawing oscillations), which eliminates all the second order error where the solution is smooth by adding in the anti-diffusive flux term but where it varies rapidly this correction is limited, suppressing oscillations.

For the scalar equation the second order term has the general form

$$C_{i-\frac{1}{2}} = -\frac{1}{2} (1 - |\nu_{i-\frac{1}{2}}|) \nu_{i-\frac{1}{2}} \Delta u_{i-\frac{1}{2}}$$

where $\nu_{k-\frac{1}{2}}$ is the discretized flux derivative (or one-dimensional eigenvalue). In order to maintain monotonicity [Roe 1981b] the correction terms actually used are chosen from the five around the relevant grid point and on the basis of the sign of $\nu_{i\pm\frac{1}{2}}$ and the size and sign of the correction terms involved. Thus the difference equation correction term is:

$$-\text{sgn}(\nu_{i-\frac{1}{2}})b_{i-\frac{1}{2}} + \text{sgn}(\nu_{i+\frac{1}{2}})b_{i+\frac{1}{2}} \quad (7.4.9)$$

where:

$$b_{i-\frac{1}{2}} = \begin{cases} B(C_{i+\frac{1}{2}}, C_{i-\frac{1}{2}}) & \text{if } \lambda_{i+\frac{1}{2}} < 0 \\ B(C_{i+\frac{3}{2}}, C_{i+\frac{1}{2}}) & \text{if } \lambda_{i+\frac{1}{2}} > 0 \end{cases} \quad (7.4.10)$$

and:

$$B(x, y) = \begin{cases} x & \text{if } x < y \text{ and } xy > 0 \\ y & \text{if } x > y \text{ and } xy > 0 \\ 0 & \text{if } xy < 0 \end{cases} \quad (7.4.11)$$

There are other forms of limiters, this one, known as the Minmod limiter from the definition of the function B , being a good example and easy to generalize for systems.

The generalization is not, however, obvious due to the loss of fixed directions of motion. In the scalar case the presence of a single variable means that information is constrained to move along the line, i.e. an invariant direction, whereas the vector directions defined in the system case are dependent on the discretization. Thus while the same corrections C can be defined (corresponding to each eigenvalue) the definition of the function B has to be extended in order to cope with the corrections being multi-componented. The problem results from the two vectors to be compared in the B function being in general non-parallel vectors, thus a size and relevant sign comparison is insufficient.

There would appear to be three courses of action to construct a reasonable vector B function $\mathbf{B}$. Each has been numerically tested with little difference in performance and no sign of oscillations near discontinuities, so further investigations numerically and analytically will have to be pursued in order to resolve the problem. The different methods are each derived from applying the Minmod principle to different aspects of the vectors involved. Firstly the choice can be made by just examining one component of the vector, then there is the idea of forming a composite vector by applying the B function to each component pair in turn and finally the choice can be made on the length of the vector, with the zero result being returned when any of the components change sign. The first and third methods will return one of the vectors in the function or zero if the relevant components have opposite signs and thus will send information along one of the derived wave directions when the result is non-zero. The second method constructs an approximation to the direction of the anti-diffusive flux as some sort of Minmod average of the vectors and thus introduces a new direction.

Now returning to the problem the second order correction (to be added to the upwind approximation (7.4.8)) will have a term corresponding to each eigenvalue derived from the cells adjacent to the relevant gridpoint, thus:

$$\begin{aligned} & -\text{sgn}(\lambda_1^{i-\frac{1}{2}})b_1^{i-\frac{1}{2}} - \text{sgn}(\lambda_2^{i-\frac{1}{2}})b_2^{i-\frac{1}{2}} \\ & \quad + \text{sgn}(\lambda_1^{i+\frac{1}{2}})b_1^{i+\frac{1}{2}} + \text{sgn}(\lambda_2^{i+\frac{1}{2}})b_2^{i+\frac{1}{2}} \end{aligned} \quad (7.4.12)$$

where the $b_k^{i-\frac{1}{2}}$ ($k = 1, 2$) are defined by the $\mathbf{B}$ function analogously to those

in (7.4.10) and the general second order term $C_i^{i-\frac{1}{2}}$ is

$$C_k^{i-\frac{1}{2}} = -\frac{1}{2} \left(1 - |\lambda_k^{i-\frac{1}{2}}| \right) \lambda_k^{i-\frac{1}{2}} a_k^{i-\frac{1}{2}} \mathbf{e}_k^{i-\frac{1}{2}}.$$

$\mathbf{B}$ may then be given by:

$$\mathbf{B}(\mathbf{x}, \mathbf{y}) = (\alpha_1, \alpha_2) \text{ where } \alpha_1 = B(x_1, y_1), \alpha_2 = B(x_2, y_2) \quad (7.4.13)$$

and $\mathbf{x} = (x_1, x_2)^T$, $\mathbf{y} = (y_1, y_2)^T$. This is the Minmod averaging function, the second of the three described above. Examples of the effect of this function are given in Fig. 7.5.

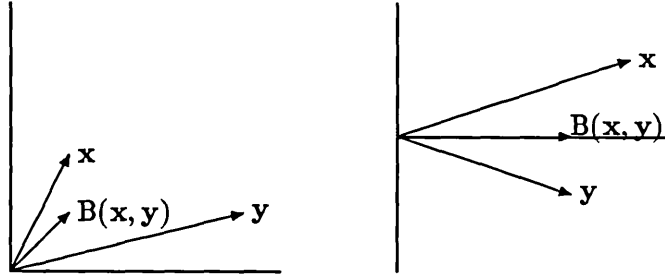


Figure 7.5: The Minmod average of two sets of vectors

Finally the numerical scheme must take in to account the parabolic degeneracy of the system. It is clear that with the discretized Jacobian matrix designed to behave like its continuous counterpart it obviously may also degenerate at certain points. Thus at these points the discretized matrix will be non-diagonalisable, the (discrete) eigenvectors will not span the solution space and splitting the flux difference along the wave directions will be impossible.

A way to overcome this problem in the case of a two component system becomes clear when recognising that $\Delta \mathbf{u}_{i-\frac{1}{2}}$ is an approximation to the direction along which $\mathbf{u}$ in $[x_{i-1}, x_i]$ is moving. Also applying the upwinding

idea if $\lambda_1^{i-\frac{1}{2}} (= \lambda_2^{i-\frac{1}{2}}) > 0$ then $\Delta \mathbf{u}_{i-\frac{1}{2}}$ will most accurately represent the direction of $\mathbf{u}$ at $\mathbf{u}_i$ and at $\mathbf{u}_{i-1}$ if $\lambda_1^{i-\frac{1}{2}} > 0$. Thus the eigenvector associated with this eigenvalue, i.e. the wave direction given by A^h , will directly approximate $\Delta \mathbf{u}_{i-\frac{1}{2}}$. Therefore the obvious step in this case is to use $\Delta \mathbf{u}_{i-\frac{1}{2}}$ and (7.4.5) to define $\Delta \mathbf{F}_{i-\frac{1}{2}}$ with the upwinding process assigning the whole flux to one endpoint depending on the sign of $\lambda_1^{i-\frac{1}{2}}$.

It is clear that all that is happening here that with the degeneracy the (two component) system becomes locally scalar with a consequential single direction for the information to travel along. This means that the simple use of $\Delta \mathbf{u}_{i-\frac{1}{2}}$ to define the first order fluxes in this case can be extended to define the second order anti-diffusive fluxes (directly from the scalar case) as:

$$\mathbf{C}_{i-\frac{1}{2}} = -\frac{1}{2} \left(1 - |\lambda_1^{i-\frac{1}{2}}| \right) \lambda_1^{i-\frac{1}{2}} \mathbf{u}_{i-\frac{1}{2}}.$$

The obvious way to then use $\mathbf{C}_{i-\frac{1}{2}}$ in the anti-diffusive flux transfer is to assign half of it to compare with each of the distinct fluxes derived from the adjacent cells in the $\mathbf{B}$ function in (7.4.13), assuming the cells to be non-degenerate. If an adjacent cell is also degenerate then the $\mathbf{B}$ function can then be applied directly to compare the two single anti-diffusive fluxes from each cell.

7.5 Numerical Experiments and Implementation of Methods

Numerical experiments were designed to compare the efficiency of Roe's scheme with the basic upwinding scheme and the exact solution. This required a series of computations run with different grid spacing as well as the programming of the exact solution to give graphical comparisons. There was no detailed analysis of the errors, the general trends as indicated by the graphical results were sufficient for the requirements of the investigation.

The first step in setting up the difference scheme is to discretize the Jacobian matrix given in (7.3.13). The matrix A^h must satisfy (7.4.5) and

also that

$$A^h(\mathbf{u}_{i-1}, \mathbf{u}_i) \rightarrow A(\mathbf{u}_{i-\frac{1}{2}}) \text{ as } h \rightarrow 0. \quad (7.5.1)$$

A way to enforce the first of these conditions is not necessarily obvious though is facilitated by decomposing A thus:

$$A(\mathbf{u}) = \begin{pmatrix} a & b \\ c(a - df) & df + bc \end{pmatrix}, \text{ with } a = f_u, b = f_v, c = \frac{v}{u}, d = \frac{1}{u} \quad (7.5.2)$$

with the eigenvalues and eigenvectors thus being:

$$\lambda_1 = df, \lambda_2 = a + bc, \mathbf{e}_1 = \begin{pmatrix} b \\ df - a \end{pmatrix}, \mathbf{e}_2 = \frac{1}{d} \begin{pmatrix} 1 \\ c \end{pmatrix}. \quad (7.5.3)$$

Now with discretisations a^h, b^h, c^h, d^h and f^h of the elements of (7.5.2)

$$a^h = \frac{f(u_{i-1}, v_{i-1}) - f(u_i, v_{i-1})}{u_{i-1} - u_i} \quad (7.5.4a)$$

$$b^h = \frac{f(u_i, v_{i-1}) - f(u_i, v_i)}{v_{i-1} - v_i} \quad (7.5.4b)$$

$$c^h = \frac{1}{2} \left(\frac{v_{i-1}}{u_{i-1}} + \frac{v_i}{u_i} \right), \quad d^h = \frac{2}{u_{i-1} + u_i} \quad (7.5.4c)$$

$$f^h = \frac{1}{2} (f(u_{i-1}, v_{i-1}) + f(u_i, v_i)) = \frac{1}{2} (f_{i-1} + f_i) \quad (7.5.4d)$$

a matrix A^h analogous to A is formed which satisfies (7.4.5) and obviously (7.5.1). Also the approximate eigenvectors and eigenvalues may similarly be defined using analogues to the $\mathbf{e}_k$ and λ_k ($k = 1, 2$) in (7.5.3) and so the flux splitting follows immediately (from (7.4.6)).

Using the discrete system a series of experiments were performed to compare the basic upwinding and Roe's scheme with the exact solution. This was done on a series of refined grids and an investigation into the effect of using the different limiter formulations described in 7.4 was performed.

The results from the numerical experiments are presented graphically and each run gave rise to two graphs, displaying the profiles derived from the upwind, Roe's and exact solutions with the left being the concentration component and the right the saturation component. The majority of the experiments were performed with initial conditions:

$$\mathbf{u}_r = (u_r, c_r) = (0.35, 0.0) \quad (7.5.5)$$

with the constants from (7.2.11) being $u_1 = 0.3$, $u_0 = 0.7$ and with boundary conditions of:

$$\mathbf{u}_I = (u_I, c_I) = (1.0, 1.0) \quad (7.5.6)$$

These will be taken as standard throughout the results except where a change to the initial saturation was made which will be indicated as it arises.

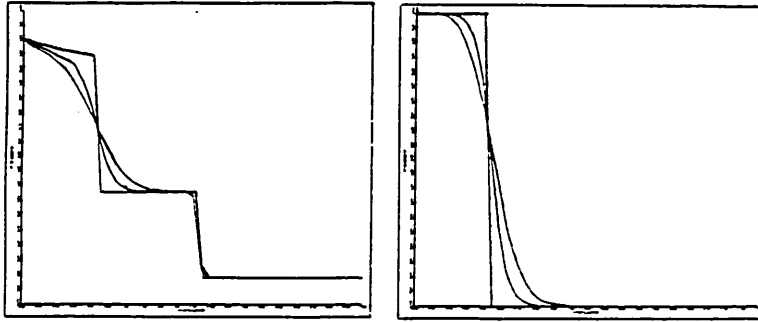


Figure 7.6: Solutions on a 40 point grid - $u_r = 0.35$

The first results, presented in Fig. 7.6, are of the three solution comparison with the limiter used being the minmod described in (7.4.13). The grid used is coarse, consisting of forty points, and the partition spacing ratio is 1/6 to ensure stability. The experiment was then repeated using refined grids of one and two hundred points to give an idea of the order of approximation of the Roe's method in comparison with the upwind scheme. These

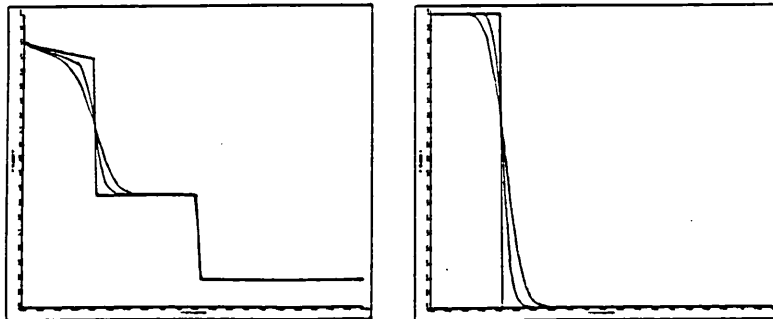


Figure 7.7: Solutions on a 100 point grid - $u_r = 0.35$

are shown in Figs. 7.7 and 7.8 respectively and as can be seen the flux corrected solution (the middle of the three curves) is more accurate than the upwind (the most diffuse curve) but there is no evidence to suggest a higher order of convergence than the theoretical $O(h)$ convergence of the upwind scheme.

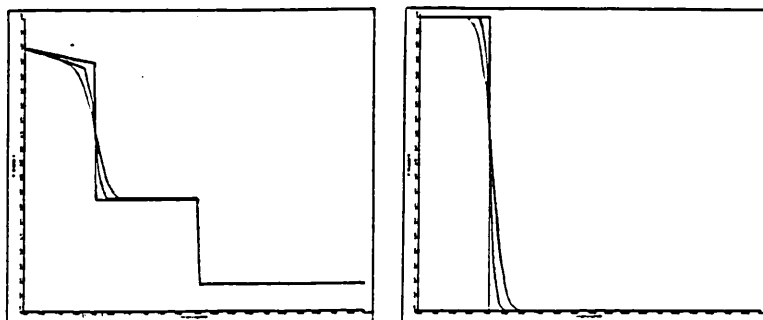


Figure 7.8: Solutions on a 200 point grid - $u_r = 0.35$

Further experiments were then performed using the different limiter formulations which produced very close results to those in the examples shown in Figs. 7.6→8. There was a slight improvement using the first of the three described limiters, with the Minmod function acting only on the first component, but insufficient to make any value judgement on the three. The final numerical investigations were then performed on a system with a different initial aqueous saturation. Here $u_r = 0.7$ which is a physically interesting situation i.e. where the chemical flood is being applied after a waterflood has been completed, and the runs presented are from runs on two fine grids, one hundred and two hundred spatial divisions, Figs. 7.9 and 7.10 respectively.

Finally throughout the numerical work the situation regarding the relative size of the eigenvalues was observed in order to discover whether the parabolic degeneracy occurred and if it had any effect. The situation never arose though as the grids were refined the minimum difference between eigenvalues became smaller and it could be clearly envisaged.

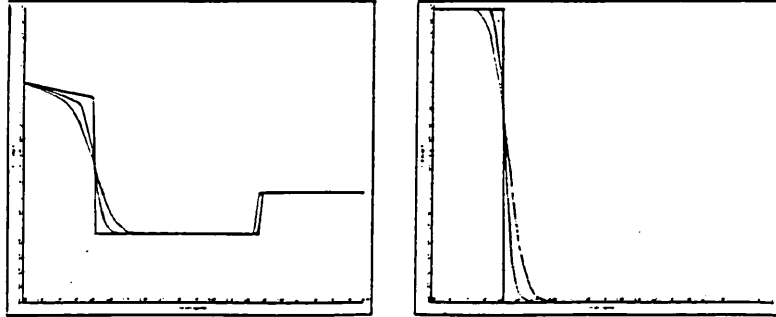


Figure 7.9: Solutions on a 100 point grid - $u_r = 0.70$

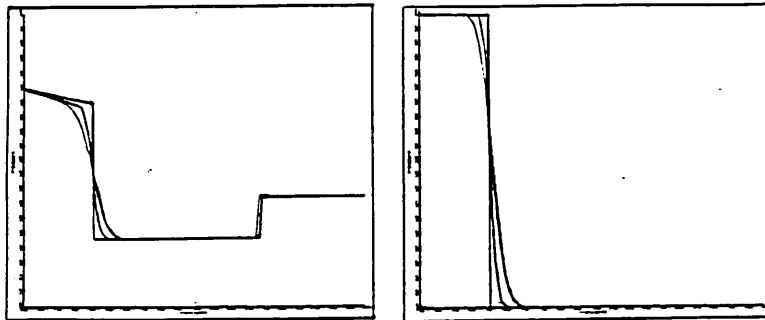


Figure 7.10: Solutions on a 200 point grid - $u_r = 0.70$

Notation and Basic Results

$$(u, v) = \int_{\Omega} uv \, dx \quad (u, v)_{\Omega_i} = \int_{\Omega_i} uv \, dx$$

$$\langle g, h \rangle = \int_{\partial\Omega} gh \, dx \quad \langle g, h \rangle_{\partial\Omega_i} = \int_{\partial\Omega_i} gh \, dx$$

$$a(u, v) \equiv (\sigma \nabla u, \nabla v)$$

$$W^{k,p}(\Omega) \equiv \{u : \nabla^i u \in L_p(\Omega), \, i = 0 \rightarrow k\}$$

$$H^k(\Omega) \equiv W^{k,2}(\Omega)$$

$$H^{k+\alpha}(\Omega) \equiv \{u \in H^k(\Omega) : \left(\int_{\Omega} \frac{|\nabla^k u(x) - \nabla^k u(y)|^p}{|x-y|^{2+2\alpha}} \, dx \, dy \right)^{\frac{1}{p}} < \infty\} \quad 0 < \alpha < 1$$

$$H_g^1(\Omega) \equiv \{u \in H^1(\Omega) : u = g \text{ on } \partial\Omega\}$$

$$\|u\|_{k,p;\Omega} \equiv \left(\sum_{i=0}^k \|\nabla^i u\|_{L_p(\Omega)}^p \right)^{\frac{1}{p}}$$

$$\|u\|_{k,\Omega} \equiv \|u\|_{k,2;\Omega}.$$

Trace inequalities (see [Kufner et al. 1977]:

$$\|D^\alpha u\|_{0,\partial\Omega} \leq c \|u\|_{\alpha+1,\Omega} \quad \forall u \in H^{\alpha+1}(\Omega)$$

specifically;

$$\|u\|_{k,\partial\Omega} \leq c \|u\|_{k+1,\Omega} \quad \forall u \in H^{k+1}(\Omega)$$

$$\left\| \frac{\partial u}{\partial \nu} \right\|_{k,\partial\Omega} \leq c \|u\|_{k+2,\Omega} \quad \forall u \in H^{k+2}(\Omega).$$

Bibliography

- [Babuška 1970] Babuška, I. (1970). Finite element methods for domains with corners. *Computing*, 6:264–273.
- [Babuška and Aziz 1972] Babuška, I. and Aziz, A. K. (1972). Survey lectures on the mathematical foundations of the finite element method. In *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations*, Aziz, A. K., editor, pages 409–474. Academic Press, New York.
- [Babuška et al. 1961] Babuška, I., Mežlik, L., and Vitasel, E. (1961). Effects of artificial cooling of concrete in a dam during its hardening. In *VII Congress des Grands Barrages*, Rome 1961, pages 1–13.
- [Babuška and Miller 1984] Babuška, I. and Miller, A. (1984). The post-processing approach in the finite element method — Part 2: The calculation of stress intensity factors. *Int. J. for Num. Meth. in Engineering*, 20:1111–1129.
- [Barrett and Elliott 1985] Barrett, J. W. and Elliott, C. M. (1985). Fitted and unfitted finite element methods for elliptic equations with smooth interfaces. Technical Report 6, Purdue University, West Lafayette, Indiana.
- [Barrett and Elliott 1987] Barrett, J. W. and Elliott, C. W. (1987). Total flux estimates for a finite element approximation of elliptic equations. *IMA J. of Numer. Anal.*, 7:129–148. Nordhoff, Leyden.

- [Blum and Dobrowolski 1982] Blum, H. and Dobrowolski, M. (1982). On finite element methods for elliptic equations on domains with corners. *Computing*, 28:53–63.
- [Bramble and Schatz 1970] Bramble, J. H. and Schatz, A. H. (1970). On the numerical solution of elliptic boundary value problems by least squares approximation of the data. In *Proceedings of the Two Symposium on the Numerical Solution of Partial Differential Equations (Synspade II)*, Hubbard, B., editor, pages 107–131. Academic Press.
- [Ciarlet 1978] Ciarlet, P. G. (1978). *The Finite Element Method for Elliptic Problems*. North Holland, Amsterdam.
- [Ciarlet and Raviart 1972] Ciarlet, P. G. and Raviart, P. A. (1972). The combines effect of curved boundaries and numerical integration in isoparametric finite element methods. In *The Mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations*, Aziz, A. K., editor, pages 409–474. Academic Press, New York.
- [Dobrowolski 1981] Dobrowolski, M. (1981). *Numerical approximation of elliptic interface and corner problems*. PhD thesis, University of Bonn.
- [Douglas et al. 1983] Douglas, J., Ewing, R. E., and Wheeler, M. F. (1983). Approximation of the pressure by a mixed method in the simulation of miscible displacement. *RAIRO Anal. Numer.*, 17:17–33.
- [Eriksson 1985a] Eriksson, K. (1985a). Improved accuracy by adapted mesh-refinements in the finite element method. *Math. Comp.*, 44(170):321–343.
- [Eriksson 1985b] Eriksson, K. (1985b). Finite element methods of optimal order for problems with singular data. *Math. Comp.*, 44(170):345–360.
- [Ewing and Wheeler 1981] Ewing, R. E. and Wheeler, M. F. (1981). Galerkin methods for miscible displacement problems with point sources and sinks – unit mobility ratio case. In *Special Year in Numerical Analysis, Lecture Notes 20*, pages 151–174. University of Maryland.

- [Ewing and Wheeler 1983] Ewing, R. E. and Wheeler, M. F. (1983). Computational aspects of mixed finite element methods. In *IMACS Transactions on Scientific Computation*, Stepleman, R., editor. North-Holland, Amsterdam.
- [Fix et al. 1973] Fix, G., Gulati, S., and Wakoff, G. I. (1973). On the use of singular functions with finite element approximations. *J. Comp. Phys.*, 13:209–238.
- [Grisvard 1975] Grisvard, P. (1975). Behaviour of the solutions of an elliptic boundary value problem in polygonal or polyhedral domain. In *Proceedings of the Third Symposium on the Numerical Solution of Partial Differential Equations (Synspade III)*, Hubbard, B., editor, pages 207–274. Academic Press, New York.
- [Grisvard 1976] Grisvard, P. (1976). Smoothness of the solution of a monotonic boundary value problem for a second order elliptic equation in a general convex domain. In *Ordinary and Partial Differential Equations, Dundee 1976, Lecture Notes in Mathematics*, pages 135–151. Springer Verlag.
- [Isaacson 1985] Isaacson, E. I. (1985). Global solution of a Riemann problem for a non-strictly hyperbolic system of conservation laws arising in enhanced oil recovery. Private communication.
- [Levine 1985] Levine, N. (1985). Superconvergent estimation of the gradient from linear finite element approximation on triangular elements. Numerical Analysis Report 3/85, University of Reading.
- [Lax 1957] Lax, P. D. (1957). Hyperbolic conservation laws II. *Comm. Pure Appl. Math.*, 10:537–566.
- [Kufner et al. 1977] Kufner, A., John, O., and Fučík, S. (1977). *Function Spaces*.
- [Nitsche and Schatz 1974] Nitsche, J. A. and Schatz, A. H. (1974). Interior estimates for Ritz–Galerkin methods. *Math. Comp.*, 28:937–958.

- [Peaceman 1977] Peaceman, D. W. (1977). *Fundamentals of Numerical Reservoir Simulation*. Elsevier, Amsterdam.
- [Pope 1980] Pope, G. A. (1980). The application of fractional flow theory to enhanced oil recovery. *Soc. of Pet. Eng. Journal*, 20.
- [Raviart and Thomas 1977] Raviart, P. A. and Thomas, J. M. (1977). A mixed finite element method for 2nd order elliptic problems. In *Mathematical Aspects of the Finite Element Method, Rome 1975, Lecture Notes in Mathematics*. Springer Verlag, Berlin.
- [Roe 1981a] Roe, P. (1981a). Numerical algorithms for the linear wave equation. *J. Comp. Phys.*, 43:357–372.
- [Roe 1981b] Roe, P. (1981b). Approximate Riemann solvers, parameter vectors and difference schemes. Technical report, Royal Aircraft Establishment.
- [Russel and Wheeler 1983] Russel, T. F. and Wheeler, M. F. (1983). Finite element methods and finite difference methods for continuous flows in porous media. In *The Mathematics of Reservoir Simulation*, Ewing, R., editor. Siam, Philadelphia.
- [Schatz 1980] Schatz, A. H. (1980). A weak discrete maximum principle and stability of the finite element method in L_∞ on plane polygonal domains. I *Math. Comp.*, 34(149):77–91.
- [Schatz and Wahlbin 1978] Schatz, A. H. and Wahlbin, L. B. (1978). Maximum norm estimates in the finite element method on plane polygonal domains. Part 1. *Math. Comp.*, 32(141):73–169.
- [Schatz and Wahlbin 1979] Schatz, A. H. and Wahlbin, L. B. (1979). Maximum norm estimates in the finite element method on plane polygonal domains. Part 2, refinements. *Math. Comp.*, 33(146):465–492.
- [Scott 1973] Scott, R. (1973). Finite element convergence for singular data. *Numer. Math.*, 21:317–327.