306319134U

Mathematical and Computational Modelling of Ultrasound Elasticity Imaging

James Southern

Keble College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2006

RADCLIFFE SCIENCE LIBRARY OXFORD
DEPOSITED
- 3 AUG 2007
THESIS

Acknowledgements

Firstly, I am very grateful for the help given to me by my supervisors Professor David Gavaghan, Professor Alison Noble and Professor Jon Chapman. Their advice and encouragement throughout my time as a D.Phil. student has been invaluable.

Next, I would like to thank Dr Jonathan Whiteley for his help throughout, especially with regard to the implementation of the finite element solution method. Also, everyone else in the Computational Biology group, particularly those in Room 306, has provided a useful source of discussion and advice.

Mike Kadour, Junbo Li and Yao Yao Cui of the Wolfson Medical Vision Laboratory provided the essential link to an experimental ultrasound elasticity imaging group necessary for this work, in particular in providing the images for the validation of the parameter recovery method. Without their help this work would have been impossible.

I am also grateful to Professor Mike Giles and Professor Endre Süli for their help in reformulating hyperbolic systems of PDEs as elliptic minimization problems.

I also acknowledge the financial support of the Engineering and Physical Sciences Research Council, through their Doctoral Training Centre program.

Finally, I am very grateful to all of my family and friends and to everyone in Keble MCR and the Life Sciences Interface DTC for making the last four years so enjoyable.

Abstract

In this thesis a parameter recovery method for use in ultrasound elasticity imaging is developed. Elasticity imaging is a method for using a series of ultrasound images (and the displacement field between them) to estimate the spatial variation of the stiffness of the tissue being imaged. Currently iterative methods are used to do this: a model of tissue mechanics is assumed and a large number of simulations using varying parameters are compared to the actual displacement field. The aim of this work is to develop a solution method that works back from the known displacement field to determine the tissue properties, reducing the number of simulations that must be performed to one.

The parameter recovery method is based on the formulation and direct solution of the 2-d linear elasticity inverse problem using finite element methods. The inverse problem is analyzed mathematically and the existence and uniqueness of solutions is described for varying numbers of displacement fields and appropriate boundary conditions. It is shown to be hyperbolic (and so difficult to solve numerically) and then reformulated as a minimization problem with hyperbolic Euler-Lagrange equations.

A finite element solution of the minimization problem is developed and implemented. The results of the finite element implementation are shown to work well in recovering the parameters used in numerical simulations of the linear elasticity forward problem so long as these are continuous. The method is shown to be robust in dealing with small errors in displacement estimation and larger errors in the boundary values of the parameters. The method is also tested on displacement fields calculated from series of real ultrasound images.

The validity of modelling the ultrasound elasticity imaging process as a 2-d problem is discussed. The assumption of plane strain is shown not to be valid and methods for extending the parameter recovery method to 3 dimensions once 3-d ultrasound becomes more widely used are described (but not implemented).

Contents

1 Introduction **1**

1.1 Motivation 1

1.2 Medical Ultrasound 2

1.2.1 Basics 3

1.2.2 Generation of an Ultrasound Image 3

1.3 Elasticity Imaging 5

1.4 Soft Tissue and Biomechanics 7

1.4.1 Soft Tissue 7

1.4.2 Biomechanics 8

1.5 Thesis Overview and Original Contributions 11

2 Theory **13**

2.1 Elementary Mechanics 13

2.1.1 Stress and Strain 14

2.1.2 The Mechanics of Fluids 20

2.1.3 The Mechanics of Solids 24

2.2 Partial Differential Equations 30

2.2.1 Characteristics and Classification 30

2.2.2 Boundary Conditions 38

2.2.3 Riemann Invariants 39

2.3 The Calculus of Variations 40

2.3.1 The Euler Equations 40

2.3.2 Integral Constraints and the Euler-Lagrange Equation 43

2.4 The Finite Element Method 44

2.4.1 Triangulation and Mesh Generation 45

2.4.2 Reformulating the System of PDEs 46

2.4.3 Numerical Methods 51

3 Previous Work **60**

3.1 Iterative Reconstruction Techniques 61

3.1.1 Newton's Method 62

3.1.2 A Bisection Method 64

3.1.3 A Method Based on the Simulated Stress Field 65

3.2 A Direct Solution Method 67

3.3 Comparing the Iterative and Direct Approaches 68

3.4 Theoretical Work 69

4 A New Parameter Recovery Method **72**

4.1 The Inverse Problem for Linear Elasticity 73

4.1.1 Formulation of the Inverse Problem 73

4.1.2 Simplifying Assumptions 74

4.2 Analysis of the Inverse Problem 76

4.2.1 Hyperbolicity and Characteristics 76

4.2.2 The Characteristics and Cauchy Data 78

4.2.3 Non-Existence of Riemann Invariants 80

4.3 Multiple Displacement Fields 83

4.3.1 Two Displacement Fields 84

4.3.2 More Than Two Displacement Fields 88

4.3.3 Determination of λ When μ is Known 92

4.3.4 Summary 93

4.4 Numerical Solution of the Inverse Problem 94

4.4.1 The Minimization Problem 94

4.4.2 The Euler-Lagrange Equations 95

4.4.3 Classification of the Euler-Lagrange Equations 97

4.4.4 Finite Element Solution 102

4.4.5 Summary 107

5 Testing the Parameter Recovery Method **109**

5.1 Explicit Solutions of the Forward Problem 110

5.1.1 Uniaxial Tension 111

5.1.2 Goodier's Solution for a Cylindrical Inclusion 112

5.2 Numerical Solutions of the Forward Problem 115
5.2.1 Smoothly Varying Lamé Constants 116
5.2.2 Lamé Constants With Discontinuous Derivative 121
5.2.3 Discontinuous Lamé Constants 125
5.3 Stress Singularities at the Corners 126
5.4 Robustness of the Method 130
5.4.1 Errors in Displacement Estimation 131
5.4.2 Sensitivity to Boundary Data 140
5.5 Conclusions 145

6 The 3-Dimensional Inverse Problem **146**
6.1 2-D Plane Strain vs 3-D Elasticity 147
6.1.1 Comparing the Forward Problems 147
6.1.2 2-D Parameter Recovery on a Cross-Section 151
6.1.3 Computational Issues 154
6.1.4 Conclusions 155
6.2 The 3-Dimensional Inverse Problem 156
6.2.1 Formulation of the Inverse Problem 157
6.2.2 Classification of the Inverse Problem 158
6.3 Multiple Displacement Fields 161
6.3.1 Two Displacement Fields 162
6.3.2 Three Displacement Fields 164
6.4 The Minimization Problem 167

7 Validation: Real Ultrasound Images **170**
7.1 Format of the Input Data 171
7.1.1 RF-Mode Images 171
7.1.2 Displacement Data 172
7.2 Results 172
7.2.1 Boundary Data and Multiple Displacement Fields 172
7.2.2 Gelatin Phantom 174
7.2.3 Scar Tissue 179
7.2.4 Invasive Tumours 181
7.3 Conclusions 183

8 Summary and Future Work **186**
8.1 Summary of Key Contributions 186
8.2 Addressing Problems With the Current Method 187
8.2.1 Dealing with Discontinuous Lamé Constants 187
8.2.2 3-Dimensional Implementation 189
8.2.3 Sensitivity to Displacement Errors 190
8.3 Possible Extensions to the Method 191
8.3.1 Improved Displacement Estimation Techniques 192
8.3.2 Realistic Parameter Values 193
8.3.3 Improved Modelling of Soft Tissue Mechanics 194
8.3.4 Incorporate Confidence in Stiffness Estimates 196
8.4 Concluding Remarks 197

A Elastic Models of Soft Tissue Mechanics **198**
A.1 Linearly Elastic Models 198
A.1.1 Classical Linear Elasticity 199
A.1.2 Azar's Pseudo-Nonlinear Model 199
A.2 Finite Elasticity 201
A.2.1 Isotropic Hyperelasticity 201
A.2.2 Modelling the Strain Energy Function 203
A.3 Pseudoelasticity and Fung Elasticity 205
A.3.1 Uniaxial Loading Experiments 205
A.3.2 Fung's 3-d Exponential Strain Energy Function Postulate 206
A.3.3 Polynomial Pseudoelasticity 208
A.3.4 Combining the Two Pseudoelastic Models 208

B Viscoelastic and Mixture Models of Soft Tissue **210**
B.1 Theory 210
B.1.1 Viscoelasticity 211
B.1.2 Mixture Theory 217
B.2 Models 220
B.2.1 Quasi-Linear Viscoelasticity 220
B.2.2 Other Viscoelastic Models 226
B.2.3 Mixture Models 231

C Methods for Solving the Equations of Plane Strain Linear Elasticity **239**
C.1 Solution by Plane Harmonic Functions 239
C.1.1 The Biharmonic Form 240
C.1.2 Harmonic Functions 241
C.1.3 Solution of Navier's Equation 243
C.1.4 Application to the Cylindrical Inclusion Problem 245
C.2 Complex Variable Methods 248
C.2.1 The Complex Potential 248
C.2.2 Analytic Extension of the Complex Potentials 252
C.2.3 The Punch Problem on the Half-Plane 253

Bibliography **255**

List of Figures

1.1 A typical ultrasound image 6

2.1 Calculating the stress vector at a surface. 14
2.2 The deformation of a region Ω under stress. 17
2.3 Velocity profile of a Newtonian 1-d shear flow. 22

4.1 Typical characteristics for the linear elasticity inverse problem 79
4.2 Extension of the domain of definition of the solution to the inverse problem by using two displacement fields 86

5.1 Solution of the inverse problem for uniaxial compression 113
5.2 Residuals for the solution of the inverse problem for uniaxial tension and with applied boundary conditions 113
5.3 Solution of the inverse problem for a cylindrical inclusion under uniaxial tension 115
5.4 Boundary conditions applied in the numerical solution of the linear elasticity forward problem. 117
5.5 Smooth input for $\lambda(\mathbf{x})$, $\mu(\mathbf{x})$. 118
5.6 Solution of the inverse problem with smooth input and applied boundary conditions 119
5.7 Residuals for the solution of the inverse problem with smooth input and applied boundary conditions 119
5.8 Solution of the inverse problem with smooth input and multiple displacement fields 120
5.9 Residuals for the solution of the inverse problem with smooth input and multiple displacement fields 120
5.10 Input $\lambda(\mathbf{x})$, $\mu(\mathbf{x})$ for a cylindrical insertion. 122

5.11 Solution of the inverse problem with a discontinuous derivative and applied boundary conditions 123
5.12 Residuals for the solution of the inverse problem with a discontinuous derivative and applied boundary conditions 123
5.13 Solution of the inverse problem with a discontinuous derivative and multiple displacement fields 124
5.14 Residuals for the solution of the inverse problem with a discontinuous derivative and multiple displacement fields 124
5.15 Discontinuous input $\lambda(\mathbf{x})$, $\mu(\mathbf{x})$. 125
5.16 Solution of the inverse problem for discontinuous Lamé constants and applied boundary conditions 127
5.17 Solution of the inverse problem for discontinuous Lamé constants and multiple displacement fields 127
5.18 Typical stress fields for the unit square under compression (fine mesh) . . . 128
5.19 Typical stress fields for the unit square under compression (coarse mesh) . 129
5.20 Solution of the inverse problem with noise having variance 5×10^{-6} and applied boundary conditions 133
5.21 Residuals for the inverse problem with noise having variance 5×10^{-6} and applied boundary conditions 133
5.22 Solution of the inverse problem with noise having variance 10^{-5} and applied boundary conditions 134
5.23 Residuals for the inverse problem with noise having variance 5×10^{-6} and applied boundary conditions 134
5.24 Solution of the inverse problem with noise having variance 5×10^{-5} and applied boundary conditions 135
5.25 Solution of the inverse problem with noise having variance 10^{-6} and multiple displacement fields 136
5.26 Residuals for the inverse problem with noise having variance 10^{-6} and multiple displacement fields 136
5.27 Solution of the inverse problem with noise having variance 5×10^{-6} and multiple displacement fields 137
5.28 Residuals for the inverse problem with noise having variance 5×10^{-6} and multiple displacement fields 137

5.29 Solution of the inverse problem with noise having variance 10^{-5} and multiple displacement fields 138
5.30 Residuals for the inverse problem with noise having variance 10^{-5} and multiple displacement fields 138
5.31 Solution of the inverse problem with noise having variance 5×10^{-5} and multiple displacement fields 139
5.32 Solution of the inverse problem with small errors in the applied boundary conditions 141
5.33 Residuals for the solution of the inverse problem with small errors on the applied boundary conditions 141
5.34 Solution of the inverse problem with large errors in the applied boundary conditions 142
5.35 Residuals for the solution of the inverse problem with large errors in the applied boundary conditions 142
5.36 Solution of the inverse problem with large, oscillating errors in the applied boundary conditions 144
5.37 Residuals for the solution of the inverse problem with large, oscillating errors in the applied boundary conditions 144

6.1 3-d deformation of a cuboid by an line indenter 149
6.2 Comparison of 2-d and 3-d deformations of a unit square 150
6.3 True values of the Lamé constants on the cross-section $x = 0$ for a 3-d displacement field 152
6.4 Solution of the 2-d inverse problem on a cross-section of a 3-d displacement field 152
6.5 Values of the Lamé constants on the cross-section $x = 0$ that conserve leave the Young's modulus and Poisson's ratio unchanged between 2-d and 3-d . 153
6.6 Solution of the 2-d inverse problem on a cross-section of a 3-d displacement field 153

7.1 Regular mesh used for parameter recovery on real ultrasound images . . . 173
7.2 Ultrasound and strain images of a gelatin phantom with a cylindrical inclusion 174
7.3 Stiffness distribution in a gelatin phantom 175
7.4 Results of applying the parameter recovery method to a series of images of a gelatin phantom, using an integral constraint 176

7.5 Results of applying the parameter recovery method to a series of images gelatin phantom, using boundary conditions 178
7.6 Ultrasound and strain images of a scar 179
7.7 Results of applying the parameter recovery method to a series of images of a scar 180
7.8 Ultrasound and strain images of an invasive tumour 181
7.9 Results of applying the parameter recovery method to a series of images of an invasive tumour 182
7.10 Ultrasound and strain images of an invasive tumour 183
7.11 Results of applying the parameter recovery method to a series of images of an invasive tumour 184

B.1 Creep in a viscoelastic solid. 212
B.2 Stress relaxation in a viscoelastic solid. 212
B.3 Three mechanical models of a viscoelastic material. 213

Chapter 1

Introduction

1.1 Motivation

Cancer is one of the biggest threats to human health in the modern world. According to Cancer Research UK [17][1] 152,857 people died of cancer in the UK in 2004. This was equivalent to 25% of all deaths, making it the second biggest killer after cardiovascular disease (37% of all deaths in 2004 according to the British Heart Foundation [16]). This proportion rises to 39% in adults under the age of 65, in whom cancer is the largest cause of death. Breast cancer is the most commonly diagnosed cancer in the UK (despite the fact that it is very rare in men), with 42,023 new cases being diagnosed in 2002. The incidence of breast cancer has increased over the last 25-30 years, from an (age-standardized) incidence rate per 100,000 women of 75 in 1979 to 116 in 2002.

Early diagnosis of breast cancer is a key factor in its successful treatment. The introduction of a national screening program in 1988 has led to more tumours being detected at an early stage and, together with the development of more effective treatments, has resulted in the number of people dying of breast cancer decreasing by 31% between 1989 and 2004 (despite the number of cases increasing as discussed above). Since the introduction of screening, many cancers have been detected by x-ray mammography before any other symptoms are apparent.

Abnormal tissue (both malignant and benign) is generally 'harder' (or 'stiffer') than nor-

[1]From where all of the cancer related statistics in this section are obtained, unless otherwise stated

mal tissue, so a second common way of detecting tumours is through manual palpation. This sensation of hardness is due to an increase in stiffness in the transition from normal to abnormal tissue. In cancerous tissue the increase in stiffness is thought to be more pronounced than in benign lesions. The reasons for this are not well understood, but it is believed that the increased stiffness is due to the presence of cancer-specific structures such as necrotic tissue and neovasculature development (see, e.g., Harris *et al.* [35]).

Currently the only time the hardness of tissue is considered is during manual palpation and, due to its subjectivity, this is used only to indicate that further investigation — in general a biopsy — is required. Similarly, if a potential tumour is detected on a mammogram (and still appears suspicious after an ultrasound scan) then a biopsy is usually performed. This work is motivated by the idea that, given a reliable method of imaging tissue stiffness and a better understanding of the mechanical properties of tumours, it might be possible to replace biopsies with a non-invasive technique for the early diagnosis of tumours, particularly in the breast where they can easily be investigated using ultrasound. If the imaging process could be performed quickly (i.e. within a few minutes at the most) this could also lead to more accurate diagnosis occurring immediately, rather than having to wait for results to come back from the laboratory, thus allowing treatment to begin earlier.

1.2 Medical Ultrasound

Ultrasound is used in medicine both as a primary imaging modality and in conjunction with other modalities. It produces images of cross-sections of regions of interest to clinicians. Since ultrasound cannot penetrate bones or pockets of gas its main function is in imaging soft tissue. Ultrasound has three key clinical advantages over alternative modalities:

- it produces images in real time;
- it does not use ionizing radiation;
- it can be used to give quantitative images of blood flow;

Medical ultrasound is particularly useful for providing images of tumours in the breast since these are generally surrounded entirely by soft tissue (no bones or lungs between them and the skin) and are not buried too deeply within the body (beyond the effective range of an ultrasound pulse).

1.2.1 Basics

An ultrasound machine (in conventional 'echo' mode) works by transmitting a pulse of sound into a region of interest and detecting the echoes received back from within the tissue. The sound is transmitted from an array of piezoelectric elements, known as the *aperture*. The time of transmission from each element of the aperture can be delayed to focus the pulse. This is called *beam-forming*. The received echoes, which are detected by the same piezoelectric elements used for transmission, are used to produce either a 3-dimensional (3-d), or more usually 2-dimensional (2-d) map of *tissue echogenicity*. The raw ultrasound data, made up of a series of sound waves (one received at each element of the aperture) corresponding to lines through the region of interest is known as *RF-mode* (radio frequency mode) ultrasound. The data is usually presented to the clinical user in *B-mode* (brightness mode). In this mode the data is presented as a greyscale image of the logarithm of the echo intensity as a function of position.

Wide bandwidth sound pulses (longitudinal pressure waves) are generally used to interrogate the region of interest. In a breast transducer, a typical ultrasound pulse might have a frequency in the range 7.5-12 MHz, a wavelength of 1.5 mm, maximum particle displacement (amplitude) of 0.0057 μm and phase velocity of 1540 ms^{-1}. The pulse emitted by an ultrasound machine is often referred to by its *point spread function* (PSF). The PSF is dependent on the acoustic signal generated at the transducer elements, the aperture and element geometries, the beam-forming applied and the form of the received signal from the medium being probed. Predicting the form of the PSF is important in modelling ultrasound. However, for an arbitrary array geometry this cannot usually be done analytically. Hence, in general the PSF is either determined by simulation or a simplified model PSF is used.

1.2.2 Generation of an Ultrasound Image

In a homogeneous medium with density ρ_0 and compressibility κ_0 it can be shown that the acoustic pressure $p(\mathbf{x}, t)$, where $\mathbf{x}$ represents position and t time, is a solution of the homogeneous wave equation

$$\nabla^2 p - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = 0 \tag{1.1}$$

where $c = \sqrt{\rho_0 \kappa_0}$ is the wave speed (see, e.g., Love [45]). This model predicts that the pressure waves are non-dispersive and travel in straight lines through the media. Tissue, however, is not a homogeneous medium — it contains interfaces between regions with different bulk properties (e.g. the cancerous lesions that are to be detected are less compressible than the healthy tissue surrounding them) — and within regions small fluctuations in the values of density and compressibility about their bulk values are seen. Both types of inhomogeneity redirect some of the energy of the acoustic waves back to the aperture, allowing the formation of ultrasound images. There are two ways that the energy can be redirected — *reflection* and *scattering.*

Suppose the region of interest contains a feature, with a scale larger than the wavelength λ of the ultrasound, within which the properties of the (otherwise homogeneous) tissue differ from those outside the feature. In particular suppose that the tissue outside the feature has density ρ_1 and compressibility κ_1, with that inside having density ρ_2 and compressibility κ_2. This defines wave speeds $c_1 = \sqrt{\rho_1 \kappa_1}$ and $c_2 = \sqrt{\rho_2 \kappa_2}$ inside and outside the feature respectively. On solving the homogeneous wave equation (1.1) in the two regions and applying matching conditions at the boundary it turns out that a proportion of the energy of the acoustic wave is reflected at the boundary of the feature. This proportion is determined by the *reflection coefficient* R, defined (for a wave entering the feature) by

$$R = \frac{\rho_2 c_2 - \rho_1 c_1}{\rho_2 c_2 + \rho_2 c_2}. \tag{1.2}$$

Ultrasound exploits this reflected sound to image tissue boundaries. However, if there is a very large difference in tissue properties across a boundary it is possible that virtually all of the energy in the sound wave could be reflected. In this case the ultrasound machine cannot 'see' past the boundary. This is the case across boundaries between tissue and air or between tissue and bone. This explains why ultrasound is primarily used as a soft tissue modality and cannot be used to image tissue in or under the lungs.

Now consider variations in density and compressibility on a scale smaller than λ. In this case some of the assumptions that are required in the formulation of equation (1.1) are not valid and a new model for the propagation of the acoustic wave is required. It turns out (see Gore & Leeman [33]) that an appropriate model is the inhomogeneous wave equation

$$\nabla^2 p - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = \nabla.(\gamma \nabla p) + \frac{\beta}{c^2}\frac{\partial^2 p}{\partial t^2} \tag{1.3}$$

where

$$\beta = \frac{\delta\kappa}{\kappa_0}, \quad \gamma = \frac{\delta\rho}{\rho} \tag{1.4}$$

and $\delta\kappa$, $\delta\rho$ are the deviations of compressibility and density respectively from their means. For a homogeneous region, equation (1.3) reduces to equation (1.1). Analysis of equation (1.3) for small spherical, elastic features, first attempted by Faran [25], shows that in an otherwise homogeneous medium sound waves radiate in all directions from these features. Hence these small features are called *scatterers*. The exact distribution of the energy amongst the possible directions (radiation pattern) depends on the properties of the scatterer — if only compressibility varies from the bulk properties then the scattered waves have equal amplitude in all directions, whereas if only density varies most of the energy is radiated back towards the source. In most cases a combination of these two effects is observed.

The Faran solution shows that increasing the size of the scatterer up to and beyond λ causes the radiation pattern to become more directional, with much more of the scattered energy travelling directly towards or directly away from the source. This is in agreement with what has already been observed to be the case for large features since reflection will be the dominant effect. The Faran model can be extended to inelastic spheres. In this case the effects of reverberation within the scatterers is included. This can significantly alter the echo response seen from the scatterer.

So far the effects of a single scatterer or an interface in the tissue have been considered. A more realistic situation is for the tissue of interest to contain a large number of small, randomly spaced scatterers. The effect of this is to sum the contributions of all of the individual scatterers giving rise to incoherent noise known as *speckle* (see, e.g., Hill *et al.* [37]). It is thought that the texture of the speckle is not dependent on the underlying tissue structure, although the local brightness does reflect local echogenicity of the underlying scatterers. Occasionally the scatterers may not be randomly arranged, but have some periodic arrangement. In this case the speckle will have a coherent component.

1.3 Elasticity Imaging

An ultrasound image like the one shown in Figure 1.1 is made up of a combination of the sharp lines seen from reflection at tissue boundaries and the textured speckle generated

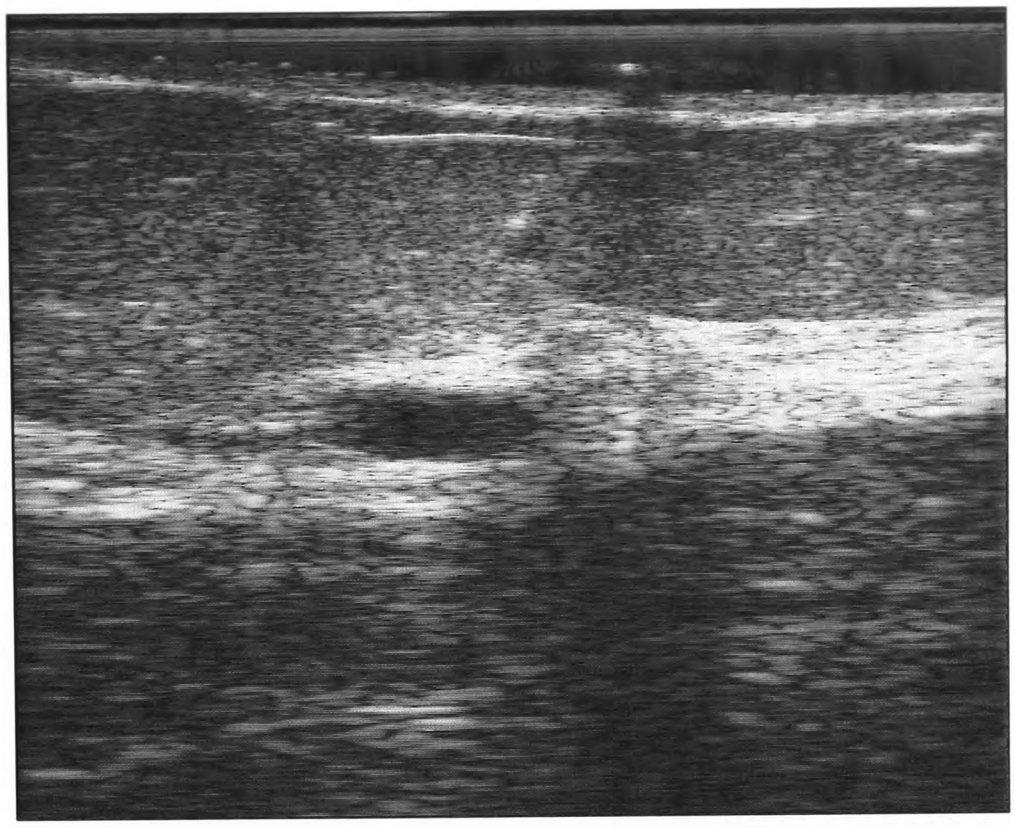


Figure 1.1: An example of a typical ultrasound image, in this case of a cyst in the breast.

everywhere else in the image. Hence, the image is inherently *noisy* (although in medical ultrasound the 'image' and the 'noise' are effectively one and the same), making it difficult to pick out the features of interest as the speckle reduces the contrast between a tumour and the surrounding healthy tissue. Various ways of increasing the contrast[2] have been suggested. One of the methods with the most potential appears to be elasticity imaging.

The aim of elasticity imaging is to construct an image of the mechanical properties of a region of interest. Definitions of these properties are given later in this work (primarily in Section 2.1), but fundamentally they are a measure of the amount of deformation (strain) an object experiences in response to the application of a known force (stress). Hence, ultrasound elasticity imaging requires at least two ultrasound images of the region of interest — one before the force is applied and at least one afterwards (for more accuracy a series of intermediate images may also be used). In addition the subsequent images must be registered to the first (or to a nominal reference image in the sequence). Once the images have been registered to one another a mathematical model of the expected behaviour of the tissue being imaged is used to estimate the mechanical properties of the tissue. Hence classic elasticity imaging is a two step process: displacement estimation and parameter recovery.

[2]Here the term 'contrast' is not used to mean the greylevel contrast, but as a measure of the 'delineation of the object boundary'.

The main focus of this work will be the modelling process for soft tissue mechanics and the solution of the resulting differential equations. It is assumed that the displacement field between two ultrasound images can be determined perfectly. In reality, this is far from being the case — ultrasound displacement estimation is very much an active field of research. Before moving on completely from displacement estimation, however, a brief summary of state-of-the-art methods is given.

According to Barron *et al.* [7], displacement estimation methods can be characterized into three main classes. *Differential* (or gradient-based) methods make the assumption that brightness does not change with time. The displacement estimation is calculated by calculating the derivatives of the intensity and often using a knowledge-based regularization procedure. This class of methods does not produce good results on ultrasound images due to their low signal-to-noise ratio. The second class of methods is based on *frequency*, exploiting either energy or phase data. The theory suggests that phase-based displacement estimation should work best for ultrasound. However, the methods exploit filters that are tuned to a velocity range and it is difficult to design optimal filters. The final class is *block-matching* displacement estimation. Here the displacement is estimated by maximizing a *similarity measure* between the images of interest. These methods are generally computationally expensive and are best suited to tracking the position of one feature (e.g. a lesion) between the images.

1.4 Soft Tissue and Biomechanics

1.4.1 Soft Tissue

The term *soft tissue* is used to describe material in the (human) body that undergoes a significant deformation in response to an applied stress. In practice this definition includes almost every material within the body with the exception of bone and pure fluids. This leads us to the observation that soft tissue can take many forms (e.g. skin, muscle, various internal organs, etc.), each responding differently to the applied stress. Each of these different tissues are, however, constructed from the same 'building blocks' — cells (of which, according to Humphrey [39], there are over 200 different types in the human body) and the extracellular matrix — with the behaviour of the tissue depending on the way in which these building blocks are arranged. *Cells* are the fundamental functional and

structural units of tissue whilst the *extracellular matrix* (ECM) holds the cells together and provides them with the aqueous environment necessary for their survival. The ECM regulates the shape, movement and orientation of the cells within the tissue of interest (although it is, in turn, constructed and maintained by cells). The overall mechanical properties of the tissue are determined by the properties of the cells and ECM that make it up and by the unique way in which they are structured within each tissue.

1.4.2 Biomechanics

Fung [30] defines *biomechanics* simply as "mechanics applied to biology." It is a modern field, having been developed primarily since the middle of the 1960s [39]. However, despite the simple definition and the relatively short time it has been in existence, biomechanics is a remarkably diverse field. That this should be the case is evident from the discussion of soft tissue in the previous section — every one of the different types of soft tissue must be considered separately and in addition biomechanics includes the study of hard tissue (such as bone) and biological fluids (e.g. blood). Humphrey [39] tells us that the study of biological material has required the development of new mathematical models and the revision of some of the basic postulates and concepts of mechanics. This observation leads Humphrey to redefine biomechanics as "the development, extension and application of mechanics for the purposes of understanding better physiology and pathophysiology as well as the diagnosis and treatment of disease and injury."

1.4.2.1 History of Biomechanics

Although biomechanics has only developed into an area of research distinct from mechanics in the last 40 years, problems in biology have been studied using mechanics for hundreds of years. Fung [30] and Humphrey [39] give several examples of these problems. Leonardo da Vinci (1452–1519) studied the flight of birds in order to determine whether man could fly, Galilieo Galilei (1562–1642) investigated the mechanical properties of bones and René Descartes (1596–1650) developed a mechanical model of animal structure.[3] Giovani Borelli

[3]Descartes' practical knowledge of physiology was limited and subsequent experiments contradicted many of the findings of his model, causing many physiologists to distrust the results of mechanical models. This emphasizes the need for theoreticians to consult with experimentalists at all stages of the modelling process in order to ensure that the model is valid for the system it is meant to represent.

(1608–1679) developed less ambitious but more successful models of the movement of various animals. In developing Hooke's law, Robert Hooke (1635–1703) performed experiments primarily on wire but also on biological materials such as hair and bone. Hooke was also the first person to use the word 'cell' in biology. Leonhard Euler (1707–1783) worked in the physiology branch of the St Petersburg Academy and formulated the governing equations for the propagation of pressure waves in arteries, which were later solved by George Friedrich Bernhard Riemann (1826–1866). The work of Thomas Young (1773–1829) on the human voice led him to consider the elasticity of materials and define the Young's modulus. Young also combined mechanics and physiology to study the mechanical function of the heart and arteries.

Interest in biological problems as a part of mechanics continued throughout the 19th and 20th centuries until, (according to Humphrey [39]) in the mid 1960s, the modern field of biomechanics was created. This development came about to a large degree as a result of the development of the theory of elasticity in the first half of the 20th century (see, e.g., Love [45]) and developments in nonlinear continuum mechanics in the middle part of the 20th century (described in Truesdell & Noll [84] and Green & Adkins [34]). Since many biological tissues behave nonlinearly under finite stresses their study could not proceed without the development of these fields. At around this time the development of the digital computer and the finite element method provided the necessary tools for the numerical solution of the complex nonlinear governing equations resulting from many biomechanical models, many of which are not analytically tractable. Finally, Humphrey [39] suggests that exploration of space from the 1950s onwards produced a requirement for understanding of how the human body would respond when subjected to the altered loads experienced in space travel.

So, by the middle of the 1960s the theoretical basis, methods, technology and motivation necessary for the study of biomechanics had been developed far beyond where they had been at any time previously. Hence it is no surprise that this period saw a rapid expansion of the field. In the last 40 years biomechanics has become a major branch of applied mathematics, with applications throughout the life sciences. For example, physiologists have developed complex models of the beating heart (see, e.g., McCulloch *et al.* [47]). Since the function of the heart is to pump blood around the body it is necessary to include a biomechanical model to describe the pumping motion. Many other current applications of biomechanics are described in some detail by Fung [30], including blood flow, bone fracture, muscle contraction and descriptions of the behaviour of cartilage, ligament and tendons.

1.4.2.2 Problem Solving in Biomechanics

The way in which biomechanics problems are studied is completely analogous to that in which problems in other areas of mechanics are approached. The problem is broken down into three general areas of study, identified by Humphrey [39] as:

- Identification of fundamental concepts, postulates and principles.
- Formulation of equations to describe the behaviour of the tissue of interest.
- Solution of initial and boundary value problems.

In addition, it is generally essential to compare theoretical results obtained in this way with results determined experimentally in order to validate the governing equations. Clearly biological materials must obey the fundamental postulates of mechanics (e.g. conservation of mass, momentum and energy) and concepts such as stress, strain and elasticity which are familiar from other fields of mechanics can be used. Hence, study of biomechanical problems can generally be reduced to the formulation of governing equations and then the solution of these equations given appropriate initial and boundary data.

Building on the above ideas Fung [30] proposes an 8 step method for the solution of a general biomechanical problem:

1. Determine the geometry of the tissue of interest. This includes determining both the initial shape of the tissue (from the morphology of the organism and anatomy of the organ) and the internal structure of the tissue (using knowledge of the types of cells present in the tissue, the way they are arranged within the ECM and the composition of the ECM itself).
2. Determine the mechanical properties of the materials making up the tissue. These can then be used to determine constitutive equations for each of the materials. In biomechanics the constitutive equations can be difficult to determine exactly for a number of reasons (e.g. tissue specimens that are too small for testing, difficulty of isolating the tissue for experimentation in its normal living condition, the tissue experiences large deformations leading to nonlinear and history-dependent constitutive equations). Usually, however, models can be built based on certain assumptions about the tissue which approximate the constitutive equation and are dependent only on a small number of numerical parameters that are left for determination in step 7.

3. Derive the governing equations of the tissue of interest in terms of the parameters identified in step 2. In general this is done using the constitutive equations derived in the previous step, conservation laws (of mass, momentum and energy) or other fundamental laws of physics (such as Maxwell's equations) and by making certain additional assumptions to simplify the model (e.g. by the identification of any terms multiplied by small parameters which can be safely neglected).

4. Determine appropriate initial and boundary conditions. This step requires an understanding of the environment and workings of the tissue of interest.

5. Solve the initial and boundary value problems.

6. Test the solutions obtained in the previous step by comparison with physiological experiments. If the results do not agree for any choice of parameter values then evaluate the assumptions made in steps 2 and 3, modify the mathematical model used and repeat steps 2 – 5.

7. Determine the values of any unknown parameters remaining from step 2 by comparison with the experimental results.

8. Use the validated model to predict the solutions of other initial and boundary value problems associated with the same constitutive equations. In this way the model can be used instead of more complicated experiments to evaluate the behaviour of the tissue of interest.

This is an iterative method — initial and boundary value problems (step 5) are solved for a first model (defined in steps 2 and 3), the model is refined (step 6) and parameter values estimated (step 7), then this process is repeated. Since the main variation between models is the form of the constitutive equations (the parameter values are a part of these) it is clear that the key step in solving a general biomechanics problem is formulating a good model of the constitutive equations of the materials that make up the tissue of interest. Both Fung [30] and Humphrey [39] emphasize this point.

1.5 Thesis Overview and Original Contributions

The structure of the rest of the thesis is outlined here. First, in Chapter 2 various areas of mathematical theory that are required later in the thesis are presented. These include

the elementary theory of mechanics (including linear elasticity theory), the classification and analysis of partial differential equations, the calculus of variations and finite element theory. In Chapter 3 a review of the existing work in this field is given. This includes discussions of both existing elasticity imaging techniques and the various models of soft tissue that are used in both elasticity imaging and more general tissue mechanics studies.

In Chapter 4 a new parameter recovery method for ultrasound elasticity imaging is developed. This method uses a linear elasticity model for soft tissue and finite element methods to determine tissue properties from given displacement fields (which are assumed to be known exactly). The results of applying this method to idealized plane strain displacement fields are described in Chapter 5. The method appears to work very well and has the further advantage of being much more computationally efficient than existing methods.

The limitations of the method with regard to modelling tissue deformation by an ultrasound probe as a plane strain problem are discussed in Chapter 6. The assumption of plane strain is shown to be invalid and a 3-d parameter recovery method is proposed as a way of avoiding this problem. Unfortunately 3-d ultrasound images are not currently available for elasticity imaging so the 3-d method is not pursued. Instead, in Chapter 7 the 2-d parameter recovery method is tested using real series of 2-d ultrasound images. Finally, in Chapter 8 some possible extensions of this work are discussed.

Chapters 4, 5, 6 and 7 constitute original contributions of this thesis.

In addition to the main body of the thesis, three appendices contain additional material that is not directly related to the development of the new parameter recovery method, but either may be used to develop it further in the future or is used in the analysis and evaluation of the method. Appendices A and B contain descriptions of alternatives to the linear elasticity model used in this work to describe soft tissue mechanics. Elastic models are found in Appendix A and non-elastic models in Appendix B. Methods for computing exact solutions of the linear elasticity forward problem are presented in Appendix C. These are used in Chapter 5 to formulate some test problems for the parameter recovery method and to account for the presence of stress singularities in simulated displacement fields also used to test the method.

Chapter 2

Theory

In this chapter various areas of mathematical theory are introduced. These will provide a framework for the formulation and solution of a general mathematical model of tissue mechanics described in subsequent chapters. In Section 2.1 some very general material models are described. These give rise to systems of partial differential equations (PDEs) that govern the motion of a material in response to the forces experienced by it. Next, in Section 2.2 methods for the classification and analysis of systems of PDEs are discussed. Section 2.3 introduces the calculus of variations — a method for reducing minimization problems to systems of PDEs. Finally, Section 2.4 explains how a system of PDEs can be solved numerically using finite element methods.

2.1 Elementary Mechanics

Here some basic mechanics that will be required later is introduced. In doing this (and throughout this thesis) the *summation convention* is used: any repetition of an index in a single term will indicate summation over that index unless explicitly stated otherwise.

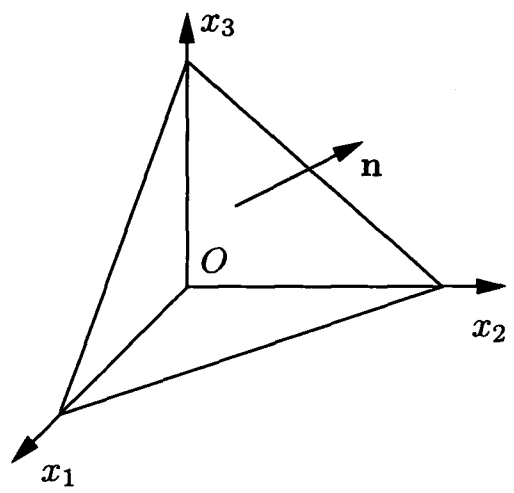


Figure 2.1: The tetrahedron used to calculate the stress vector on a surface with unit normal **n**.

2.1.1 Stress and Strain

2.1.1.1 Stress

Consider two cylindrical bars made of the same material and having the same initial length but different cross-sectional areas. Suppose a longitudinal force of magnitude F is applied to both bars. Clearly it will cause both bars to stretch. However, it is intuitively clear that the thin bar will be stretched more than the thick bar. Similarly, if the two bars have the same cross-sectional area A but have different forces applied to them then the bar on which the larger force is applied will be stretched more. It turns out that the amount the bar is stretched is a function of the force applied per unit cross-sectional area. So, for longitudinal stretching the *stress* is defined to be

$$\sigma = \frac{F}{A} \tag{2.1}$$

where F is the applied force and A the cross-sectional area of the material it is applied across.

For an arbitrary region in 3-dimensional space, define Cartesian coordinates (fixed in the body of interest) (x_1, x_2, x_3). For $i, j = 1, 2, 3$ define $\sigma_{ij}(\mathbf{x})$ to be the force in the x_i-direction per unit (deformed) area on a surface with normal in the x_j-direction. Then

$\sigma = (\sigma_{ij})$ is a rank 2 tensor[1] known as the *Cauchy stress tensor.* To calculate the stress vector on an arbitrary surface in the body with unit normal $\mathbf{n}$ consider a force balance on a tetrahedron with faces having normals in the directions of the coordinate axes and $\mathbf{n}$ as shown in Figure 2.1. On doing this it can be seen (see, e.g., Love [45]) that the stress vector has components:

$$S_i = \sigma_{ij} n_j, \quad i = 1, 2, 3. \tag{2.2}$$

Notice that since $\mathbf{S}$ and $\mathbf{n}$ are vectors, this result can also be used to show that σ is indeed a tensor. By considering an angular momentum balance on a small cube aligned with the axes it can be shown (again, see Love [45]) that the Cauchy stress tensor is symmetric, i.e.

$$\sigma_{ij} = \sigma_{ji}. \tag{2.3}$$

So, in order to completely specify the stress at any given point in the body six quantities must be specified.

It is often convenient to work with the coordinates of the undeformed body as the independent variables. Define Cartesian coordinates (X_1, X_2, X_3) describing position in the undeformed body. Then the *second Piola-Kirchoff stress tensor* $T = (T_{MN})$ is defined by setting T_{MN} to be the force in the X_M-direction per unit (undeformed) area acting on a surface with normal in the X_N-direction $(M, N = 1, 2, 3)$.[2] In much the same way as for the Cauchy stress tensor it can be shown that T is symmetric and satisfies the definition of a tensor.[3] The following relationship exists between the Cauchy and second Piola-Kirchoff stresses (see Spencer [80]):

$$\sigma = \frac{1}{\det F} F T F^T, \tag{2.4}$$

---

[1]Recall that a vector is a quantity $\mathbf{a} = (a_i)$ such that its components in a frame rotated by an orthogonal matrix T_{ij} are given by $a'_i = T_{ij} a_j$ (using the summation convention). Vectors are rank 1 tensors and the definition generalizes so that a rank 2 tensor is a set of quantities a_{ij} such that in the rotated frame $a'_{ij} = T_{il} T_{jm} a_{lm}$.

[2]In this work the following convention will be applied: lower case indices (usually i, j, k, ...) and, where possible, variable labels will represent quantities in the deformed space (defined by the (x_1, x_2, x_3) coordinates) whilst upper case indices (usually M, N, P, ...) and variable labels will refer to quantities in the undeformed space (defined by the (X_1, X_2, X_3) coordinates). Where this is not possible (e.g. in the case of density which is conventionally denoted by the Greek letter ρ) the subscript zero will be used for the undeformed quantity and no subscript used for the deformed quantity.

[3]The *first Piola-Kirchoff stress tensor* (also known as the *nominal stress tensor*) is defined to have as its (i, j)th component the force in the x_i-direction per unit(undeformed) area on a surface with normal in the X_j-direction. This tensor is *not* symmetric and is not often used.

where $F = (F_{iM})$ is the *deformation gradient tensor*, defined by

$$F_{iM} = \frac{\partial x_i}{\partial X_M}, \quad i, M = 1, 2, 3. \tag{2.5}$$

Observe that for small displacements the two stress tensors are virtually identical. Notice also that the components of the deformation gradient tensor are identical to those of the Jacobi matrix for the transformation of coordinates $(x_1, x_2, x_3) \to (X_1, X_2, X_3)$, i.e.

$$F = \frac{\partial(x_1, x_2, x_3)}{\partial(X_1, X_2, X_3)}. \tag{2.6}$$

Recall that the determinant of the Jacobi matrix is equivalent to the ratio of the volumes of a region calculated in each of the two coordinate systems. Hence it can be seen that for an incompressible material,

$$\det F \equiv 1. \tag{2.7}$$

2.1.1.2 Strain

In order to describe the deformation of a solid subjected to stress, return to the example of the bars discussed above. There are two ways to deform the bar — it can either be *stretched* (by applying a force parallel to the bar) or *sheared* (by applying a torque to the bar). For a stretch, suppose that the unstressed bar has length L_0 whilst the stressed bar has length L. Then the natural way to quantify the deformation is by using a dimensionless ratio. So, in 1-d, define the *stretch ratio* to be

$$\lambda = \frac{L}{L_0} \tag{2.8}$$

and the following *strain measures*:

$$\epsilon = \frac{L - L_0}{L_0}, \quad \epsilon' = \frac{L - L_0}{L}, \tag{2.9}$$

$$e = \frac{L^2 - L_0^2}{2L^2}, \quad E = \frac{L^2 - L_0^2}{2L_0^2}. \tag{2.10}$$

Notice that in *infinitesimal* elongation (i.e. in the limit as $L - L_0 \to 0$) all of these strain measures are identical. When a shear is applied the bar is twisted through an angle α (which varies with position in the bar). It is customary to define the *shear strain* to be

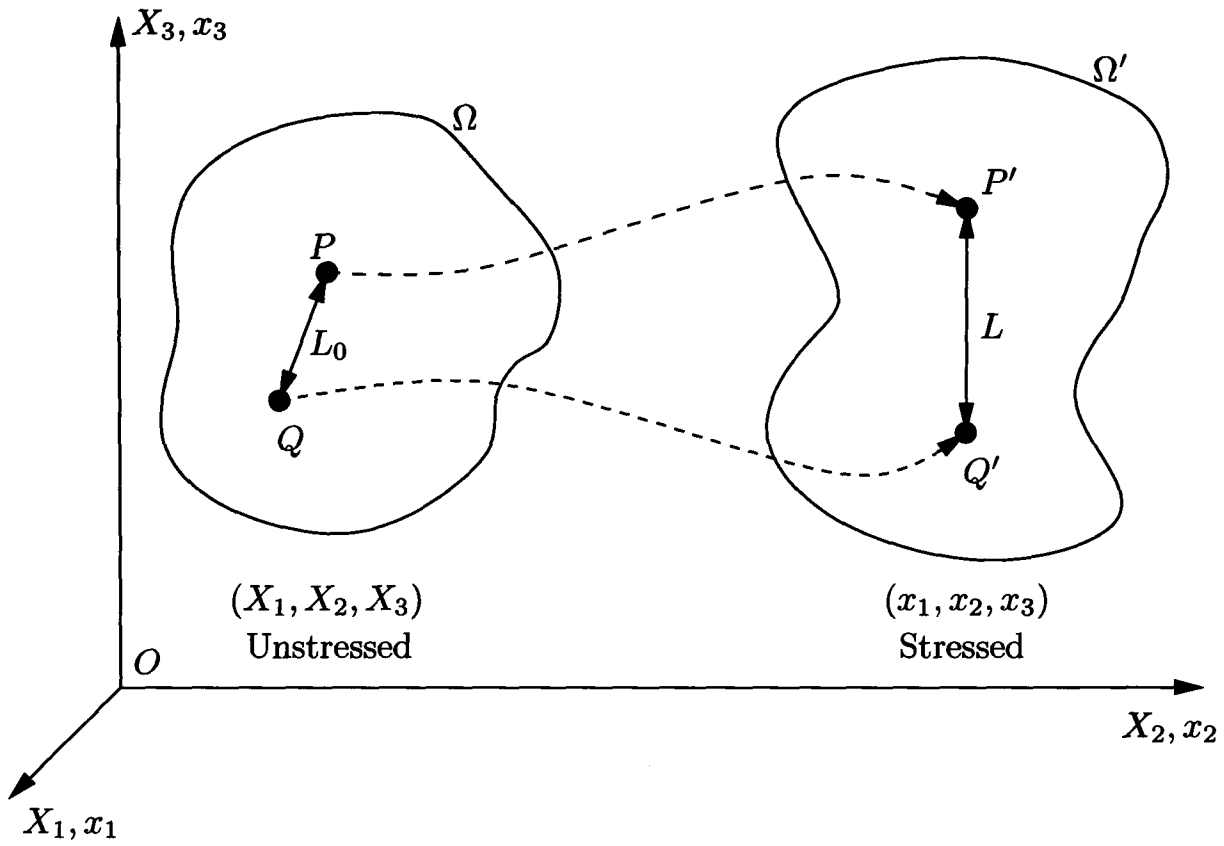


Figure 2.2: The deformation of a region Ω under stress. The line segment PQ of length L_0 is deformed to the line segment $P'Q'$ of length L. (X_1, X_2, X_3) are the coordinates of the undeformed region, (x_1, x_2, x_3) the coordinates of the deformed region.

$\frac{1}{2}\tan\alpha$. Observe that all of these strain measures are dimensionless.

In 3-d consider an unstressed body of interest occupying the region Ω. Suppose that a stress is applied, deforming the body so that it occupies the region Ω'. As before define Cartesian coordinates (X_1, X_2, X_3) in the undeformed body and (x_1, x_2, x_3) in the deformed body. For each point in the body define the *displacement vector* $\mathbf{u} = (u_1, u_2, u_3)$ by

$$u_i = x_i - X_i, \quad i = 1, 2, 3. \tag{2.11}$$

Since a rigid motion induces no stress the displacements are not directly related to the stress. Instead the displacements of various points with respect to one another must be considered. As shown in Figure 2.2, consider the line segment connecting the points $P = (X_1, X_2, X_3)$ and $Q = (X_1 + dX_1, X_2 + dX_2, X_3 + dX_3)$ in the undeformed body. This line segment has length L_0 given by

$$L_0^2 = dX_1^2 + dX_2^2 + dX_3^2. \tag{2.12}$$

The points P and Q are deformed to the points $P' = (x_1, x_2, x_3)$ and $Q' = (x_1 + dx_1, x_2 +$

$dx_2, x_3 + dx_3$) respectively. So, the deformed line segment has length L given by

$$L^2 = dx_1^2 + dx_2^2 + dx_3^2. \tag{2.13}$$

Since the deformation is a bijection between Ω and Ω' each set of coordinates can be written as a function of the other and, hence,

$$dx_i = \frac{\partial x_i}{\partial X_M} dX_M, \qquad dX_M = \frac{\partial X_M}{\partial x_j} dx_j. \tag{2.14}$$

Defining the Kronecker delta (a rank 2 tensor) by

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}, \tag{2.15}$$

it can be seen that

$$L_0^2 = \delta_{MN} dX_M dX_N = \delta_{MN} \frac{\partial X_M}{\partial x_i} \frac{\partial X_N}{\partial x_j} dx_i dx_j, \tag{2.16}$$

$$L^2 = \delta_{ij} dx_i dx_j = \delta_{ij} \frac{\partial x_i}{\partial X_M} \frac{\partial x_j}{\partial X_N} dX_M dX_N. \tag{2.17}$$

On introducing the matrices

$$e_{ij} = \frac{1}{2}\left(\delta_{ij} - \frac{\partial X_M}{\partial x_i} \frac{\partial X_M}{\partial x_j}\right), \tag{2.18}$$

$$E_{MN} = \frac{1}{2}\left(\frac{\partial x_i}{\partial X_M} \frac{\partial x_i}{\partial X_N} - \delta_{MN}\right), \tag{2.19}$$

the difference between the squares of the line segments lengths is seen to be given by the expressions

$$L^2 - L_0^2 = 2E_{ij} dX_i dX_j, \tag{2.20}$$

$$L^2 - L_0^2 = 2e_{ij} dx_i dx_j. \tag{2.21}$$

So, the two expressions (2.18) and (2.19) are exactly analogous to the 1-d strain measures given in equation (2.10).

Both $e = (e_{ij})$ and $E = (E_{MN})$ are symmetric tensors called *Almansi's strain tensor*

and *Green's strain tensor* respectively. They can be written in terms of the deformation gradient tensor as

$$e = \tfrac{1}{2}\big(I - F^{-T}F^{-1}\big) = \tfrac{1}{2}\big(I - B^{-1}\big), \tag{2.22}$$

$$E = \tfrac{1}{2}\big(F^{T}F - I\big) = \tfrac{1}{2}\big(C - I\big), \tag{2.23}$$

where $B = FF^T$ and $C = F^TF$ are known as the *left* and *right Cauchy-Green deformation tensors* respectively. Next consider the case of infinitesimal elongation, which is recalled to be the limit $L - L_0 \to 0$. Equation (2.11) can be used to express the components of the deformation gradient tensor (and of its inverse) in terms of derivatives of the displacement components. Since the displacements are small, second order terms can be neglected. Hence, e and E reduce to *Cauchy's infinitesimal strain tensor* $\epsilon = (\epsilon_{ij})$, defined by[4]

$$\epsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right). \tag{2.24}$$

Finally, observe that for a rigid body motion the deformation gradient tensor reduces to the identity tensor of rank 2. Hence, as expected, all of the strain tensors reduce to the zero tensor in the case of a rigid body motion ($L - L_0$ constant for all line segments). Conversely, a deformation resulting in a strain represented by the zero tensor can be seen to be rigid.

2.1.1.3 Strain Rate

In a fluid, strains do not persist over time (since the molecules making up the fluid can rearrange themselves internally to eliminate strain). The velocity field and the resulting *rate of change* of strain must be considered instead of the displacement field and strain used for solids in previous sections. Let $\mathbf{v} = (v_1, v_2, v_3)$ be the velocity field of the fluid of interest[5] and consider the velocities of two fluid elements P and Q located instantaneously at (x_1, x_2, x_3) and $(x_1 + dx_1, x_2 + dx_2, x_3 + dx_3)$ respectively. Then the difference in the

[4]Note that in older literature the shear strains (ϵ_{ij} for $i \neq j$) are often defined to be twice those given here (analogous to defining the 1-d shear strain as $\tan\alpha$). Using this definition, ϵ does not satisfy the definition of a tensor, hence the definition used here. Care must be taken over this point when reading the literature.

[5]Note that in the literature of fluid dynamics the velocity field is generally denoted by $\mathbf{u} = (u_1, u_2, u_3)$. Here $\mathbf{v}$ is used instead in order to avoid confusion with the displacement field $\mathbf{u}$ for the deformation of a solid.

velocities of the two particles is

$$dv_i = \frac{\partial v_i}{\partial x_j} dx_j. \tag{2.25}$$

On defining the *strain rate tensor* V_{ij} and the *vorticity tensor* Ω_{ij} by

$$V_{ij} = \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right), \tag{2.26}$$

$$\Omega_{ij} = \frac{1}{2}\left(\frac{\partial v_j}{\partial x_i} - \frac{\partial v_i}{\partial x_j}\right), \tag{2.27}$$

it can be seen that difference in velocities can be written as

$$dv_i = V_{ij} dx_j - \Omega_{ij} dx_j. \tag{2.28}$$

Observe that V is symmetric and Ω antisymmetric (i.e. $\Omega_{ij} = -\Omega_{ji}$). Notice that V is analogous to Cauchy's infinitesimal strain tensor ϵ, defined in equation (2.24) with instantaneous velocity replacing infinitesimal displacement. The reason for this can be seen by consideration of an infinitesimal time interval dt. In this time interval the instantaneous velocity $\mathbf{v}$ gives rise to an infinitesimal displacement $\mathbf{u} = \mathbf{v}dt$ so, in fact, the concepts of infinitesimal displacement and instantaneous velocity are interchangeable.

2.1.2 The Mechanics of Fluids

2.1.2.1 Inviscid Fluids

A fluid in which no shear stresses are present is called *inviscid.* Since pressure acts equally in all directions in a fluid, the stress tensor for an inviscid fluid has the form

$$\sigma_{ij} = -p\delta_{ij}, \tag{2.29}$$

where p is the pressure in the fluid. The pressure is related to the density ρ and temperature T of the fluid by an *equation of state* of the form

$$f(p, \rho, T) = 0. \tag{2.30}$$

For a real fluid the equation of state may be difficult to determine. However, two special cases exist. For an *incompressible* fluid the equation of state is (see, e.g., Batchelor [8])

$$\rho = \text{constant}. \tag{2.31}$$

It can be shown (see, e.g., Batchelor [8]) that conservation of mass in an arbitrary fluid element leads to the condition

$$\nabla.\mathbf{v} = 0. \tag{2.32}$$

For an *ideal gas*, defined by Batchelor [8] to be a fluid in which the molecules exert no force on one another (except in collisions) and have negligible volume, the equation of state is (again, see Batchelor [8])

$$\frac{p}{\rho} = RT, \tag{2.33}$$

where R is the gas constant. Many real fluids can be approximated either as incompressible or ideal gases.

In order to determine the governing equations for the motion of an inviscid fluid introduce the *material derivative*:

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + (\mathbf{v}.\nabla). \tag{2.34}$$

This can be shown (see, e.g., Acheson [1]) to be the rate of change of a quantity of interest at a point *fixed in the fluid.* Hence, the acceleration of an arbitrary fluid element Ω (fixed in the fluid) is given by

$$\int_\Omega \rho \frac{D\mathbf{v}}{Dt} dV. \tag{2.35}$$

The governing equation is derived by applying Newton's law to the fluid element Ω. This gives,

$$\int_\Omega \rho \frac{Dv_i}{Dt} dV = \int_{\partial\Omega} \sigma_{ij} n_j + \int_\Omega \rho f_i, \quad i = 1, 2, 3, \tag{2.36}$$

where $\mathbf{n} = (n_1, n_2, n_3)$ is the unit outward pointing normal to $\partial\Omega$ and $\mathbf{f} = (f_1, f_2, f_3)$ is the vector of external forces. On substituting from equation (2.29), applying the divergence theorem, removing the integration since it is over an *arbitrary* fluid element and switching to vector notation it can be seen that the governing equation is

$$\frac{D\mathbf{v}}{Dt} = -\frac{1}{\rho}\nabla p + \mathbf{f}. \tag{2.37}$$

This is known as *Euler's equation.* Then, assuming that ρ and T are given throughout

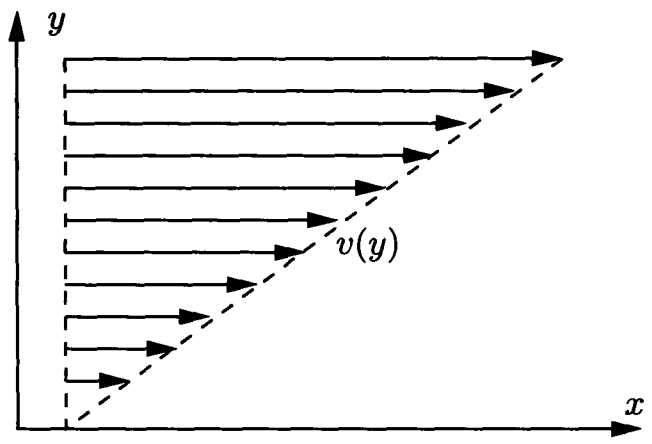


Figure 2.3: Velocity profile of a Newtonian 1-d shear flow.

the fluid, Euler's equation and the equation of state define a system of one vector and one scalar equation that can be used to determine the one unknown vector quantity ($\mathbf{v}$) and one unknown scalar quantity (p). For an inviscid fluid an appropriate boundary condition is that there should be no flow across a solid boundary.

2.1.2.2 Newtonian Viscous Fluids

Next, consider fluids in which shear stresses are present. The simplest shear flow is in 1-d where fluid flows in the x-direction with velocity $v(y)$. Suppose the flow has constant velocity gradient as shown in Figure 2.3. For this flow Newton proposed the following relationship between the shear stress σ and the velocity gradient (see, e.g. Batchelor [8]):

$$\sigma = \mu \frac{dv}{dy}, \tag{2.38}$$

where μ is a constant of proportionality known as the *coefficient of viscosity.* Note also that in 1-d the velocity gradient and strain rate are identical so that equation (2.38) provides a relationship between shear stress and rate of strain.

The stress-strain relationship of a fluid can be written (according to Batchelor [8]) as the sum of an isotropic part $-p\delta_{ij}$, where p is the *static pressure*, and a deviatoric part σ'_{ij} with the property $\sigma'_{kk} = 0$. To generalize the 1-d shear flow described above to a 3-d *Newtonian viscous fluid* the deviatoric stress tensor is assumed to depend linearly on the components of the strain rate tensor so that:

$$\sigma_{ij} = -p\delta_{ij} + A_{ijkl}V_{kl}, \tag{2.39}$$

where A is a rank 4 tensor containing the viscosity coefficients of the fluid as a function of temperature (but *not* of stress or of strain rate). The term $-p\delta_{ij}$ represents the stress in a fluid at rest and, as for the inviscid fluid, the relationship between p, ρ and T is determined by an equation of state. As a rank 4 tensor, A has 81 components; however not all of these are independent (for example, observe that since σ and δ are symmetric then A must be symmetric in i and j).

A material in which there is no preferential direction of deformation is referred to as *isotropic*. The Kronecker delta function, δ, is an isotropic tensor and, according to Batchelor [8], all rank 4 isotropic tensors in a 3-d Euclidean space that are symmetric in i and j can be written as a linear combination of two rank 4 tensors with (i,j,k,l)th components $\delta_{ij}\delta_{kl}$ and $\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk}$ respectively [8]. So, for an isotropic Newtonian viscous fluid,

$$A_{ijkl} = \lambda\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}). \tag{2.40}$$

Notice that this means that, for an isotropic fluid, A is symmetric in k and l in addition to i and j. Further, the constitutive equation can be written as (see, e.g., Batchelor [8])

$$\sigma_{ij} = -p\delta_{ij} + \lambda V_{kk}\delta_{ij} + 2\mu V_{ij}. \tag{2.41}$$

Since the deviatoric stress tensor has zero trace, the mean normal stress $\frac{1}{3}\sigma_{kk}$ is equal to the static pressure p. Then, since a contraction of equation (2.39) gives

$$\sigma_{kk} = -3p + (3\lambda + 2\mu)V_{kk}, \tag{2.42}$$

it can be seen that the following relationship holds:

$$3\lambda + 2\mu = 0. \tag{2.43}$$

So, the constitutive equation becomes

$$\sigma_{ij} = -p\delta_{ij} + 2\mu V_{ij} - \tfrac{2}{3}\mu V_{kk}\delta_{ij}. \tag{2.44}$$

A fluid with this constitutive equation is known as a *Stokes fluid*. Note that only one constant — the coefficient of viscosity μ — is necessary to describe a Stokes fluid.

Suppose finally that the fluid is *incompressible*. Then $\nabla.\mathbf{v} = V_{kk} = 0$ (see, e.g., Batchelor

[8]) and so the constitutive equation becomes:

$$\sigma_{ij} = -p\delta_{ij} + 2\mu V_{ij}. \tag{2.45}$$

Notice that if $\mu = 0$ this reduces to the constitutive equation for the inviscid fluid (2.29).

As for the inviscid fluid, the governing equation of motion for a Newtonian viscous fluid is given by applying Green's theorem to equation (2.36) and removing the integrals over an arbitrary fluid volume to give

$$\rho\frac{Dv_i}{Dt} = \frac{\partial\sigma_{ij}}{\partial x_j} + \rho f_i, \quad i = 1, 2, 3, \tag{2.46}$$

where $\mathbf{f}$ is the vector of external forces as in the inviscid case. So, for a general Newtonian viscous fluid, substituting from equation (2.39) gives the governing equation

$$\rho\frac{Dv_i}{Dt} = -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j}(A_{ijkl}V_{kl}) + \rho f_i, \quad i = 1, 2, 3. \tag{2.47}$$

Governing equations for each of the special cases discussed above can be written down by substituting from the appropriate constitutive equations. In particular, for the incompressible Stokes fluid the governing equation is, in vector notation,

$$\rho\frac{D\mathbf{v}}{Dt} = -\nabla p + \mu\nabla^2\mathbf{v} + \rho\mathbf{f}. \tag{2.48}$$

This is a second order system of PDEs and is known as the *Navier Stokes equations.* Appropriate boundary conditions for a Newtonian viscous fluid are to specify that there is no flow across, and no slip along, solid boundaries.

2.1.3 The Mechanics of Solids

2.1.3.1 Linear Elastic Solids

Recall (from, e.g., Love [45]) that in 1-d Hooke's law states that strain is directly proportional to the applied stress. For a stretch,

$$\sigma = E_Y\epsilon, \tag{2.49}$$

where $E_Y > 0$ is the *Young's modulus* of the material, and for a shear,

$$\sigma = G \tan \alpha, \tag{2.50}$$

where $G > 0$ is the *shear modulus* (or *modulus of rigidity*). A material that obeys Hooke's law is said to be *linearly elastic* (or *Hookean*). In general, materials behave in a linearly elastic fashion only when they undergo small deformations — for large deformations Hooke's law is no longer valid.

Hooke's law can be extended to the 3-d case in a similar way to the method used in Section 2.1.2.2 for a Newtonian viscous fluid — by the introduction of a tensor as the constant of proportionality. In 3-d, define a *linearly elastic solid* to be a material in which the Cauchy stress tensor σ is related to Cauchy's infinitesimal strain tensor ϵ by

$$\sigma_{ij} = A_{ijkl}\epsilon_{kl}, \tag{2.51}$$

where A is a rank 4 tensor of the moduli of elasticity. It is assumed that A is independent of both stress and strain. Since σ is symmetric then A must be symmetric in i and j.

Much of the analysis of Section 2.1.2.2 can be adapted to produce analogous results for linearly elastic solids as for Newtonian viscous fluids. So, for an isotropic solid the constitutive equation becomes (see, e.g., Love [45])

$$\sigma_{ij} = \lambda \epsilon_{kk}\delta_{ij} + 2\mu\epsilon_{ij}, \tag{2.52}$$

where λ and μ are referred to as the *Lamé constants*. The Lamé constant μ can be identified as being equivalent to the shear modulus in 1-d and is often written as G. Equation (2.52) can be inverted to show that

$$\epsilon_{ij} = \frac{1+\nu}{E_Y}\sigma_{ij} - \frac{\nu}{E_Y}\sigma_{kk}\delta_{ij}, \tag{2.53}$$

where

$$E_Y = \frac{\mu(3\lambda + 2\mu)}{\lambda + \mu} = \frac{\lambda(1+\nu)(1-2\nu)}{\nu} = 2G(1+\nu), \tag{2.54}$$

$$G = \mu = \frac{\lambda(1-2\nu)}{2\nu} = \frac{E_Y}{2(1+\nu)}, \tag{2.55}$$

$$\nu = \frac{\lambda}{2(\lambda+\mu)} = \frac{E_Y}{2G} - 1. \tag{2.56}$$

E_Y can be identified as the 3-d equivalent of the Young's modulus in the 1-d case and is referred to as such; ν is called the *Poisson's ratio* of the material and is the ratio of lateral contraction to longitudinal extension of a bar under tension. Since $\mu(= G) > 0$ it can be seen that $\nu < \frac{1}{2}$ for all materials. Further, since it is expected that the cross-sectional area of a bar will decrease when it is stretched under tension, $\nu > 0$ for *most* materials.[6] It can be shown (see, e.g., [45]) that as $\nu \to \frac{1}{2}$ a material becomes less compressible and that a material with a Poisson's ratio of $\frac{1}{2}$ is incompressible.

The governing equation for the deformation of a linear elastic solid is determined in much the same way as for the fluids in Sections 2.1.2.1 and 2.1.2.2. First, apply Newton's 2nd law to an arbitrary region of the material in equilibrium:

$$\int_{\partial\Omega} \sigma_{ij} n_j dS + \int_{\Omega} \rho f_i dV = 0, \quad i = 1, 2, 3, \tag{2.57}$$

where ρ is the density of the material, $\mathbf{f}$ the vector of external forces acting on the region of interest and $\mathbf{n}$ a unit outward pointing normal to the surface of the region of interest. Then, on applying Green's theorem, the integration (since it is over an arbitrary region within the body) can be removed to obtain the governing equation

$$\frac{\partial \sigma_{ij}}{\partial x_j} + \rho f_i = 0, \quad i = 1, 2, 3. \tag{2.58}$$

Using equation (2.51) to substitute for the stress in terms of the strain it can be seen that the governing equation becomes:

$$\frac{\partial}{\partial x_j}(A_{ijkl}\epsilon_{kl}) + \rho f_i = 0, \quad i = 1, 2, 3. \tag{2.59}$$

[6]Note, however, that materials with negative Poisson's ratio *do* exist. The cross-sectional area of a bar of such a material increases when it is stretched under tension.

In the case of the isotropic linear elastic solid for which the form of A is known the governing equation is

$$\mu\left(\frac{\partial^2 u_i}{\partial x_j^2}+\frac{\partial^2 u_j}{\partial x_i \partial x_j}\right)+\lambda\frac{\partial u_j}{\partial x_i \partial x_j}\delta_{ij}+\rho f_i=0, \quad i=1,2,3, \tag{2.60}$$

or, in vector notation,

$$\mu\nabla^2\mathbf{u}+(\lambda+\mu)\nabla(\nabla.\mathbf{u})+\rho\mathbf{f}=\mathbf{0}. \tag{2.61}$$

This is a linear second order system of PDEs and is known as the *Navier equation.* Appropriate boundary conditions for a problem in linear elasticity are to specify either $\mathbf{u}$ or the *traction* $\sigma_{ij}n_j$ on the boundary $\partial\Omega$.

It is relatively straightforward to amend the force balance to account for the motion of a solid not in equilibrium. It can be seen that if the material is not in equilibrium, equation (2.58) becomes

$$\rho\frac{\partial^2 u_i}{\partial t^2}=\frac{\partial \sigma_{ij}}{\partial x_j}+\rho f_i, \quad i=1,2,3. \tag{2.62}$$

Clearly for an isotropic linear elastic solid, the Navier equation (2.61) becomes

$$\rho\frac{\partial^2 \mathbf{u}}{\partial t^2}=\mu\nabla^2\mathbf{u}+(\lambda+\mu)\nabla(\nabla.\mathbf{u})+\rho\mathbf{f}. \tag{2.63}$$

Notice that if $\lambda+\mu=0$ this reduces to the vector wave equation, for which a general solution is known in terms of the superposition of travelling plane waves.

Some exact solutions of the equations of linear elasticity for materials with constant (or piecewise constant) coefficients exist in the literature. The best known of these are the solutions of the *plane strain* and *anti-plane strain* for materials with constant Lamé constants. These solutions are described in more detail in, e.g., Love [45]. In Appendix C some examples of the techniques that can be used to solve (plane strain) linear elasticity problems for materials with piecewise constant Lamé constants are discussed. There do not, however, appear to be any solutions of the equations of linear elasticity for a material with *continuously varying* Lamé constants in the literature.

2.1.3.2 Nonlinear Elasticity

As mentioned in Section 2.1.3.1 most materials only exhibit linear elastic properties for small deformations. For materials experiencing large deformations it is necessary to con-

sider *nonlinear elasticity.* Recall also that for large deformations there are multiple stress and strain tensors (and, in particular, Cauchy's infinitesimal strain tensor is not a valid measure of strain). In order to track the motion of the material of interest over large deformations it is necessary to choose coordinates that deform with the material (i.e. (X_1, X_2, X_3) from Section 2.1.1.1). Hence, it is appropriate to use the second Piola-Kirchoff stress tensor T and Green's strain tensor E. Using the second Piola-Kirchoff stress tensor the governing equation (2.62) becomes (see, e.g., Malvern [46]):

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial}{\partial X_M}\left(T_{MN}\frac{\partial x_i}{\partial X_N}\right) + \rho f_i, \quad i = 1, 2, 3. \tag{2.64}$$

This can be shown to be the case by using equation (2.4) to substitute for the Cauchy stress tensor and recalling the definition of the deformation gradient tensor from equation (2.5). As in the case of linear elasticity a constitutive equation is required relating the stress to the strain in order to solve the governing equation (since strain, but not stress, is an explicit function of displacement). In nonlinear elasticity the constitutive equation takes the form

$$f(T, E) = 0. \tag{2.65}$$

As was the case for linear elasticity appropriate boundary conditions are the specification of the displacement $\mathbf{u}$ or the traction $T_{MN}N_N$ (where $\mathbf{N} = (N_1, N_2, N_3)$ is the unit outward pointing normal to the boundary in the (X_1, X_2, X_3) coordinate system) on the boundary of the region of interest. The study of nonlinear elastic materials subjected to finite (i.e. not infinitesimal) strains is known as *finite elasticity.*

2.1.3.3 The Strain Energy Function

Observe that for a 1-d linear elastic material under tension the *strain energy* of the material is defined to be

$$\text{SE} = \frac{1}{2}E_Y \int_0^L \left(\frac{du}{dx}\right)^2 dx = \frac{1}{2}E_Y \int_0^L \epsilon^2 dx = \frac{1}{2}\int_0^L \sigma\epsilon dx. \tag{2.66}$$

In this case the *strain energy function* at any give point within the material is defined to be

$$W(x) = \frac{1}{2}\sigma(x)\epsilon(x) \tag{2.67}$$

so that its integral over the region of interest is equal to the total strain energy of the material. This idea can be generalized to a 3-d linear elastic material by defining the *strain energy function* in 3-d to be

$$W = \sigma_{ij}\epsilon_{ij}. \tag{2.68}$$

Then, in this case, the total strain energy of a region of interest Ω is given by

$$\text{SE} = \frac{1}{2}\int_{\Omega} W dV. \tag{2.69}$$

Observe that since

$$\sigma_{ij} = \frac{\partial W}{\partial \epsilon_{ij}} \tag{2.70}$$

specifying the form of the strain energy function in terms of the strain alone is equivalent to giving a constitutive equation for the material. Use of the strain energy function also gives a framework for a variational approach to the solution of linear elasticity problems (see Pathmanathan [60]).

For nonlinear elasticity suppose that the *strain energy function* W is defined such that the following relationship between the second Piola-Kirchoff stress and the Lagrange-Green strain holds (see Green and Adkins [34]):

$$T_{MN} = \frac{1}{2}\left(\frac{\partial W}{\partial E_{MN}} + \frac{\partial W}{\partial E_{NM}}\right). \tag{2.71}$$

Since E is, by definition, symmetric this can be simplified to

$$T_{MN} = \frac{\partial W}{\partial E_{MN}}. \tag{2.72}$$

Notice the similarity between equations (2.70) and (2.72). On integrating equation (2.72) it can be seen that the strain energy function is given by

$$W = T_{MN}E_{MN}, \tag{2.73}$$

exactly analogously to equation (2.68) for the linear case. Since $T \to \sigma$ and $E \to \epsilon$ as the displacements tend to zero it can be seen that W has exactly the same physical interpretation as in the linear case, i.e. its integral over the region of interest is equal to twice the total strain energy of the material. As in the linear case, specification of W as a

function of E is equivalent to giving a constitutive equation for the material.

2.2 Partial Differential Equations

Definition 2.1. A *partial differential equation* (often abbreviated to PDE) is an equation relating a function u of two or more independent variables and its partial derivatives.
Definition 2.2. A PDE (or system) is said to be *semi-linear* if it is linear in the highest order derivatives, e.g. a second-order PDE for $u(x, y)$ is semi-linear if it can be written as

$$a(x, y)u_{xx} + 2b(x, y)u_{xy} + c(x, y)u_{yy} = f(x, y, u, u_x, u_y), \tag{2.74}$$

where the subscripts denote partial differentiation.

In this section some of the theory used in the solution of systems of partial differential equations is introduced. For simplicity, the analysis presented here is restricted to semi-linear PDEs. A method for classifying the system in terms of the number of *characteristics* it possesses is introduced and the way in which appropriate boundary conditions for a system depends on its type is discussed. Further, in the case of *hyperbolic* equations a method for solving the system of equations by looking for quantities that are conserved along a characteristic is described. Throughout, the equations will be assumed to be defined on a domain $\Omega \subseteq \mathbb{R}^2$ and systems are assumed to consist of two PDEs. This reflects the type of problem to be solved in this work. The theory is more complex in higher dimensions and with larger systems but some of the fundamental ideas still hold (e.g. hyperbolicity and ellipticity). More details on the theory of PDEs can be found in, e.g., Ockendon *et al.* [53].

2.2.1 Characteristics and Classification

In this section a theory for the classification of first- and second-order PDEs and systems is given. The classification is based on the identification of *characteristics* — curves across which the highest derivatives of the solution are allowed to be discontinuous. The number and reality of the characteristics possessed by a PDE or system will determine its type.

2.2.1.1 First-Order PDEs

Consider a first-order semi-linear PDE of the form

$$a(x,y)\frac{\partial u}{\partial x} + b(x,y)\frac{\partial u}{\partial y} = f(x,y,u), \quad \forall (x,y) \in \Omega, \tag{2.75}$$

where a, b and f are continuously differentiable functions of their arguments. An alternative way of writing this equations in in matrix form as

$$(a,b,f).(u_x,u_y,-1) = 0, \tag{2.76}$$

where again subscripts denote partial differentiation. Now, $(u_x,u_y,-1)$ is (by definition) the normal to the solution surface in (x,y,u)-space, so (a,b,f) is everywhere tangent to the solution surface. So, a curve $(x(t),y(t),u(t))$ constructed by solving the system of ordinary differential equations

$$\dot{x} = a(x,y), \tag{2.77}$$

$$\dot{y} = b(x,y), \tag{2.78}$$

$$\dot{u} = f(x,y,u), \tag{2.79}$$

where dots denote differentiation with respect to the parameter t, will be everywhere parallel to the solution surface. Suppose that $(x(0),y(0),u(0))$ is specified to be a point in the solution surface. Then the curve that solves equations (2.77) – (2.79) and passes through $(x(0),y(0),u(0))$ lies within the solution surface. Further, a family of curves $(x(s,t),y(s,t),u(s,t))$, each lying within the solution surface, can be generated by setting

$$x(s,0) = x_0(s), \quad y(s,0) = y_0(s), \quad u(s,0) = u_0(s), \tag{2.80}$$

providing that the lines s = constant, $\Gamma = \{(s,t) : t = 0\}$ intersect at most once in Ω (if they intersect more than once then contradictory data may be imposed). The surface defined by $u(s,t)$ is now a subset of the solution surface. Hence, a 'solution' to equation (2.75) has been constructed from the initial data given on Γ in equation (2.80) in some neighbourhood of the curve $t = 0$.

Equation (2.80) is said to define *Cauchy data* on the curve $t = 0$ and the solutions of equations (2.77) – (2.79) are referred to as *characteristics* of equation (2.75). The projections

of the characteristics onto the (x, y)-plane are known as the *characteristic projections* and have gradient given by

$$\frac{dy}{dx} = \frac{\dot{y}}{\dot{x}} = \frac{b(x,y)}{a(x,y)}, \tag{2.81}$$

which depends only on position in the (x, y)-plane and is real-valued for real-valued a and b (i.e. the characteristic projections always exist for a single first-order PDE).[7] These characteristic projections can be seen to be non-intersecting in the (x, y)-plane except possibly at a *critical point* where $a = b = 0$. Observe that the condition that the curves $s =$ constant (characteristics) and Γ intersect at most once is equivalent to requiring that $(x(s, 0), y(s, 0))$ is never parallel to a characteristic projection (a full justification for this can be found in Ockendon *et al.* [53]), so Cauchy data for $u(x, t)$ must be given on a curve Γ that is nowhere parallel to a characteristic projection.

This argument for the construction of a solution of the PDE (2.75) subject to the Cauchy data given in equation (2.80) is plausible. However, to prove that it is a *classical solution* is is necessary to verify that the surface is continuous and has continuous first derivatives. Now, since a, b and f are continuous and do not depend on t, Picard's theorem[8] can be applied to show that equations (2.77) – (2.79) have unique solutions (as functions of t alone) which are continuous and have continuous derivatives with respect to t in some neighbourhood of the initial curve Γ. Further, if x_0, y_0 and u_0 are continuously differentiable then it can be shown that x, y and u are continuous and have continuous first derivatives as functions of both s and t in a neighbourhood of Γ (since they depend continuously on the initial data). Now, $u(s, t)$ is a continuously differentiable function of $x(s, t)$, $y(s, t)$ if

[7]Note that if $a(x, y) = 0$ (and $b(x, y) \neq 0$) then the gradient of the characteristic projections becomes infinite. This does not cause a problem, however, as a characteristic is a curve in (x, y)-space and there is no requirement for y to be a single-valued function of x. If both a and b are zero at some point in the (x, y)-plane then the gradient of the characteristic through that point can be calculated using L'Hôpital's rule.

[8]PICARD'S THEOREM
Suppose that $f(\mathbf{x}, t)$ is Lipschitz continuous (i.e. there exists an L such that $|f(\mathbf{x}, t) - f(\mathbf{y}, t) \leqslant L|\mathbf{x} - \mathbf{y}|$) for all $\mathbf{x}$, $\mathbf{y}$, t in a neighbourhood of $(\mathbf{x}_0, t_0) \in \{(\mathbf{x}, t) : |\mathbf{x} - \mathbf{x}_0| < b, |t - t_0| < a\}$. Then the initial value problem $\dot{\mathbf{x}}(t) = f(\mathbf{x}, t)$, $\mathbf{x}(t_0) = \mathbf{x}_0$ has a unique solution on the interval $(t_0 - \delta, t_0 + \delta)$, where $\delta = \min(a, b/B)$ and $B = \sup |f(\mathbf{x}, t)|$. Further, the solution is continuous, has a continuous first derivative and depends continuously on the initial data. For a proof of Picard's theorem see, e.g., Boyce and DiPrima [15].

the transformation $(s,t) \to (x,y)$ has a unique inverse, i.e. if

$$\det J = \det \begin{pmatrix} x' & \dot{x} \\ y' & \dot{y} \end{pmatrix} \neq 0, \infty, \tag{2.82}$$

$$\Leftrightarrow \quad ay' - bx' \neq 0, \infty, \tag{2.83}$$

$$\Leftrightarrow \quad \frac{y'}{x'} = \frac{dy}{dx} \neq \frac{b}{a}, \tag{2.84}$$

where dots continue to denote differentiation with respect to t (i.e. along a characteristic projection) and dashes denote differentiation with respect to s (i.e. along Γ). Since Γ is prohibited from being parallel to a characteristic this condition is automatically satisfied on Γ and (by the continuity of x and y as functions of s and t) in some neighbourhood of Γ. Hence, there exists a unique classical solution of equation (2.75) subject to Cauchy data defined in equation (2.80) in some neighbourhood of Γ. Hence, an alternative definition of the *characteristics* is to say that they are the curves in (x,y,u)-space in which the system defined by the PDE (2.75) and Cauchy data specified on the characteristic does not have a unique solution. The characteristic projections in the (x,y)-plane are usually referred to simply as characteristics. This terminology is used throughout the rest of this work.

2.2.1.2 First-Order Systems

Consider a general system of first-order PDEs:

$$A(x,y)\frac{\partial \mathbf{u}}{\partial x} + B(x,y)\frac{\partial \mathbf{u}}{\partial y} = \mathbf{f}(x,y,\mathbf{u}), \quad \forall (x,y) \in \Omega, \tag{2.85}$$

where the matrices A and B and vector $\mathbf{f}$ are continuously differentiable functions of their arguments and $\mathbf{u} = (u,v)^T$. In order to determine the characteristics consider curves in which it is possible for the partial derivatives of u_i, $i = 1, \ldots, N$, to be discontinuous. On a candidate curve $\Gamma = (x(s), y(s))$

$$x'\frac{\partial u}{\partial x} + y'\frac{\partial u}{\partial y} = u', \tag{2.86}$$

$$x'\frac{\partial v}{\partial x} + y'\frac{\partial v}{\partial y} = v', \tag{2.87}$$

A characteristic is a curve on which the system of equations (2.85), (2.86), (2.87) does not have a unique solution, i.e. a curve on which the left-hand-sides are linearly dependent. So, on a characteristic,

$$\det \begin{pmatrix} a_{11} & a_{12} & b_{11} & b_{12} \\ a_{21} & a_{22} & b_{21} & b_{22} \\ x' & 0 & y' & 0 \\ 0 & x' & 0 & y' \end{pmatrix} = 0, \tag{2.88}$$

where a_{ij} and b_{ij}, $i, j = 1, 2$, are the (i, j)-th entries of A and B respectively. Then on setting $y = kx'$ and replacing columns 5 and 6 with ((column 5) $-$ k(column 3) $+$ k^2(column 1)) and ((column 6) $-$ k(column 4) $+$ k^2(column 2)) respectively it can be seen that the necessary condition becomes:[9]

$$\det \begin{pmatrix} a_{11} & a_{12} & b_{11} - ka_{11} & b_{12} - ka_{12} \\ a_{21} & a_{22} & b_{21} - ka_{21} & b_{22} - kb_{21} \\ x' & 0 & 0 & 0 \\ 0 & x' & 0 & 0 \end{pmatrix} = 0. \tag{2.89}$$

This condition is equivalent to

$$\det(B - kA) = 0, \tag{2.90}$$

which is the characteristic equation for a first-order system and the characteristics are the curves $\Gamma = (x(s), y(s))$ defined by

$$\frac{dy}{dx} = k. \tag{2.91}$$

Observe that equation (2.90) is a quadratic equation in k. So, the characteristics are not guaranteed to have real gradients k. The behaviour of the system depends crucially on whether or not it has real characteristics and leads to the *classification* of the first-order system as:

- *Hyperbolic* if it has two sets of real characteristics.
- *Parabolic* if it has one (repeated) set of real characteristics.
- *Elliptic* if it has no real characteristics.

[9]These replacements are elementary column operations and do not affect the determinant of the matrix they are performed upon.

2.2.1.3 Second-Order PDEs

For a second-order PDE

$$a(x,y)\frac{\partial^2 u}{\partial x^2} + 2b(x,y)\frac{\partial^2 u}{\partial x \partial y} + c(x,y)\frac{\partial^2 u}{\partial y^2} = f(x,y,u,u_x,u_y), \tag{2.92}$$

where a, b, c and f are continuous, search for characteristics as curves on which the system given by the PDE and Cauchy data does not have a unique solution for $\mathbf{u}$. Cauchy data for a second-order equation is the specification of both u and its normal derivative on $\Gamma = (x(s), y(s))$. It can be shown (see, e.g. Ockendon *et al.* [53]) that this completely determines all directional derivatives of u on Γ. So, differentiating the first derivatives of u with respect to x and y along Γ to obtain equations for the second derivatives gives:

$$x'\frac{\partial^2 u}{\partial x^2} + y'\frac{\partial^2 u}{\partial x \partial y} = u'_x, \tag{2.93}$$

$$x'\frac{\partial^2 u}{\partial x \partial y} + y'\frac{\partial^2 u}{\partial y^2} = u'_y. \tag{2.94}$$

The necessary condition for these equations to be linearly dependent (and hence not have a unique solution for $\mathbf{u}$) is

$$\det \begin{pmatrix} a & 2b & c \\ x' & y' & 0 \\ 0 & x' & y' \end{pmatrix} = 0, \tag{2.95}$$

$$ay'^2 - 2bx'y' + cx'^2 = 0, \tag{2.96}$$

$$a - 2bk + ck^2 = 0, \tag{2.97}$$

where $k = dy/dx$ is the gradient of the characteristic. Again this is a quadratic equation for k so there can be either one, two or zero real solutions for k. Solving for k gives

$$k = \frac{b \pm \sqrt{b^2 - ac}}{a}, \tag{2.98}$$

and the solution is said to be:

- *Hyperbolic* if $b^2 > ac$, i.e. if there are two families of real characteristics.
- *Parabolic* if $b^2 = ac$, i.e. there is one family of real characteristics.

- *Elliptic* if $b^2 < ac$, i.e. if there are no real characteristics.

2.2.1.4 Second-Order Systems

Consider a system of two second-order PDEs of the form:

$$A(x,y)\frac{\partial^2 \mathbf{u}}{\partial x^2} + 2B(x,y)\frac{\partial^2 \mathbf{u}}{\partial x \partial y} + C(x,y)\frac{\partial^2 \mathbf{u}}{\partial y^2} = \mathbf{f}(x, y, \mathbf{u}, \mathbf{u}_x, \mathbf{u}_y), \tag{2.99}$$

where A, B, C and $\mathbf{f}$ are continuous. Again, a characteristic is a curve $\Gamma = (x(s), y(s))$ on which the PDE and the Cauchy data do not uniquely determine $\mathbf{u}$. Differentiating u_x along Γ gives

$$x'\frac{\partial^2 u}{\partial x^2} + y'\frac{\partial^2 u}{\partial x \partial y} = u'_x. \tag{2.100}$$

Similar relationships can be derived from the continuity of the s-derivatives of u_y, v_x and v_y along the characteristics. For these four equations and the PDE not to have a unique solution their left-hand-sides must be linearly dependent, i.e.

$$\det \begin{pmatrix} a_{11} & a_{12} & 2b_{11} & 2b_{12} & c_{11} & c_{12} \\ a_{21} & a_{22} & 2b_{21} & 2b_{22} & c_{21} & c_{22} \\ x' & 0 & y' & 0 & 0 & 0 \\ 0 & x' & 0 & y' & 0 & 0 \\ 0 & 0 & x' & 0 & y' & 0 \\ 0 & 0 & 0 & x' & 0 & y' \end{pmatrix} = 0, \tag{2.101}$$

where a_{ij}, b_{ij} and c_{ij}, $i, j = 1, 2$, are the (i, j)-th entries of A, B and C respectively. By setting $y' = kx'$ (i.e. setting k to be the gradient of the characteristics) and replacing columns 5 and 6 with ((column 5) $-\, k$(column 3) $+\, k^2$(column 1)) and ((column 6) $-\, k$(column 4) $+$ k^2(column 2)) respectively it can be seen that the necessary condition becomes:

$$\det \begin{pmatrix} a_{11} & a_{12} & 2b_{11} & 2b_{12} & c_{11} - 2kb_{11} + k^2 a_{11} & c_{12} - 2kb_{12} + k^2 a_{12} \\ a_{21} & a_{22} & 2b_{21} & 2b_{22} & c_{21} - 2kb_{21} + k^2 a_{21} & c_{22} - 2kb_{22} + k^2 a_{22} \\ x' & 0 & kx' & 0 & 0 & 0 \\ 0 & x' & 0 & kx' & 0 & 0 \\ 0 & 0 & x' & 0 & 0 & 0 \\ 0 & 0 & 0 & x' & 0 & 0 \end{pmatrix} = 0. \tag{2.102}$$

This condition is equivalent to

$$\det(C - 2kB + k^2 A) = 0, \tag{2.103}$$

which is the characteristic equation for a system of two second order PDEs. Observe that since A, B and C are 2-by-2 matrices this equation is 4th order in k, so there are at most four families of characteristics. Analogously with lower order systems and scalar second order equations, define a system of two second order PDEs to be *hyperbolic* if it has four distinct families of real characteristics and to be *elliptic* if it does not have any real characteristics. If the system has two repeated families of characteristics it is said to be *parabolic.*

Example: Linear Elasticity

Consider the equations of linear elasticity as given in Section 2.1.3.1. Then the matrices A, B and C in equation (2.99) are given by

$$A = \begin{pmatrix} \lambda + 2\mu & 0 \\ 0 & \mu \end{pmatrix}, \tag{2.104}$$

$$B = \begin{pmatrix} 0 & \lambda + \mu \\ \lambda + \mu & 0 \end{pmatrix}, \tag{2.105}$$

$$C = \begin{pmatrix} \mu & 0 \\ 0 & \lambda + 2\mu \end{pmatrix}. \tag{2.106}$$

where λ and μ are functions of x and y. Then the characteristic equation is

$$\det \begin{pmatrix} \mu + (\lambda + 2\mu)k^2 & -(\lambda + \mu)k \\ -(\lambda + \mu)k & (\lambda + 2\mu) + \mu k^2 \end{pmatrix} = 0, \tag{2.107}$$

$$\Rightarrow \quad k^4 + 2k^2 + 1 = 0. \tag{2.108}$$

Hence,

$$k^2 = -1. \tag{2.109}$$

So the equations of linear elasticity are *elliptic.*

2.2.2 Boundary Conditions

In constructing a solution of the first-order PDE in Section 2.2.1.1 Cauchy data was specified on an initial curve Γ that is nowhere parallel to a characteristic. Recall, however that the construction method involved extending the solution along the characteristics. Hyperbolic equations and systems will always have enough real characteristics to construct the solution in this way,[10] but parabolic and elliptic equations will not. Hence, Cauchy data is only appropriate for hyperbolic problems. In this section the type of boundary conditions that are required to solve elliptic and parabolic equations is discussed. In order to do this it is necessary to define the properties that the solution of the resulting *boundary value problem* should have. This motivates the following definition:

Definition 2.3. A boundary value problem (PDE plus boundary conditions) is said to be *well-posed* if:

- A solution of the boundary value problem exists.
- The solution is unique.
- The solution depends continuously on the boundary data (i.e. a small change in boundary data gives rise to a correspondingly small change in the solution).

By adapting the argument used in Section 2.2.1.1 the Cauchy problem for hyperbolic equations can be shown to be well-posed (see, e.g. Ockendon *et al.* [53]).

The equivalent of the Cauchy problem for an elliptic PDE (of second-order or below) is the *Dirichlet problem.* This involves the specification of $\mathbf{u}$ *everywhere* on the boundary of the domain. Proofs that the Dirichlet problem is well-posed can be found in, e.g., Ockendon *et al.* [53]. For second-order elliptic PDEs and systems alternatives to the Dirichlet problem are the *Neumann problem* — specification of the normal derivative of $\mathbf{u}$ everywhere on the boundary — and the *Robin problem* — specification of a linear relationship between $\mathbf{u}$ and its normal derivative everywhere on the boundary. Note, however, that the Neumann and Robin problems may be ill-posed (again, see Ockendon *et al.* [53] for full details).

For parabolic equations a combination of the types of boundary data used for hyperbolic and elliptic equations is used. Cauchy data is specified on a characteristic Γ.[11] If Γ is

[10] All first-order PDEs with real coefficients are hyperbolic since the characteristic equation is linear in the characteristic gradient k.

[11] Crucially, observe that this is the opposite of the requirement that Γ not be a characteristic for hyperbolic equations.

infinite then this is enough to make the problem well-posed on one side of Γ, for example the 2-d heat equation

$$\frac{\partial u}{\partial t} = \kappa(x,y)\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) + f(x,y), \tag{2.110}$$

with u prescribed on the (characteristic) plane $t = 0$ is well-posed in the half space $t > 0$.[12] If, however, Γ is finite then additional data is required along boundary curves that intersect the ends of Γ and are nowhere parallel to a characteristic (the planes $t = 0$ in the case of the 2-d heat equation). The boundary data given on these curves is of the form given for elliptic equations, i.e. it may be Dirichlet, Neumann or Robin. Proofs that these problems are well-posed are again not given here, but can be found in Ockendon *et al.* [53].

2.2.3 Riemann Invariants

Consider a hyperbolic first-order system of n PDEs of the form

$$A\frac{\partial \mathbf{u}}{\partial x} + B\frac{\partial \mathbf{u}}{\partial y} = \mathbf{c}, \tag{2.111}$$

where A, B are matrices of appropriate size and $\mathbf{c}$ is a vector. Then a characteristic of this equation is (by definition) a curve on which the system of equations defined by equation (2.111) and

$$x'\frac{\partial \mathbf{u}}{\partial x} + y'\frac{\partial \mathbf{u}}{\partial y} = \dot{\mathbf{u}}, \tag{2.112}$$

(where dots denote derivatives along the characteristics), does not have unique solutions for $\mathbf{u}_x$ and $\mathbf{u}_y$. Hence the left-hand-sides of these equations are linearly dependent. One consequence of this is the characteristic equation:

$$\det(B - kA) = 0, \tag{2.113}$$

on a characteristic where k is the gradient of the characteristic. Suppose $\mathbf{l}$ is a left eigenvector of the characteristic equation, corresponding to an eigenvalue k.[13] Then

$$\mathbf{l}^T(B - kA) = \mathbf{0}^T. \tag{2.114}$$

[12] This is Cauchy data for t since there is only one t-derivative in the PDE. If two t-derivatives were present then it would also be necessary to specify u_t on $t = 0$.

[13] Note that since the PDE is hyperbolic there will be n distinct real eigenvalues k and, hence, n linearly independent real eigenvectors.

Suppose further that the PDE (2.111) is also multiplied by the left eigenvector to give

$$\mathbf{l}^T A\left(\frac{\partial \mathbf{u}}{\partial x} + k\frac{\partial \mathbf{u}}{\partial y}\right) = \mathbf{l}^T \mathbf{c}, \tag{2.115}$$

since $\mathbf{l}^T B = \mathbf{l}^T kA$. So, on a characteristic $C = (x(t), y(t))$,

$$\mathbf{l}^T A\dot{\mathbf{u}} = \mathbf{l}^T \mathbf{c}\dot{x}. \tag{2.116}$$

In certain cases (e.g. if the PDE has constant coefficients) it may be possible to integrate equation (2.116) along the characteristic to give

$$\mathbf{l}^T A\mathbf{u} - \mathbf{l}^T \int^x \mathbf{c}dx' = \text{constant}, \tag{2.117}$$

on the characteristic. Then the (scalar) quantity on the LHS of equation (2.117) is known as a *Riemann invariant.* There are n linearly independent left eigenvectors, so potentially n Riemann invariants $R_1(u_i, \ldots, u_n)$, $i = 1, \ldots, n$, where $\mathbf{u} = (u_1, \ldots, u_n)^T$. These define n algebraic equations for the n components of $\mathbf{u}$ which can be solved to give $\mathbf{u}$ at any point in the domain. This provides a simple method for solving a hyperbolic first-order system of PDEs, assuming that enough Riemann invariants exist.

2.3 The Calculus of Variations

2.3.1 The Euler Equations

Consider the problem of minimizing the functional

$$I(y) = \int_a^b F(x, y(x), y'(x))dx, \tag{2.118}$$

over all functions y that are twice continuously differentiable on $[a, b]$ and satisfy the boundary conditions $y(a) = c$, $y(b) = d$ for some real c and d. Suppose $\eta(x)$ is a function that is twice continuously differentiable on $[a, b]$ with the further property that $\eta(a) = 0 = \eta(b)$. Such functions are known as *test functions.* In order to solve the minimization problem it is necessary to define some properties of functionals:

Definition 2.4. The functional I has a *local minimum* at y if there exists $\delta > 0$ such that $I(z) \geqslant I(y)$ for all functions z in the domain of I that satisfy $\sup |y(x) - z(x)| < \epsilon$.

Definition 2.5. The directional derivative of I at y in the direction of a test function η, known as the *Gâteaux derivative*, is given by

$$\partial_\eta I(y) = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} [I(y + \epsilon\eta)]. \tag{2.119}$$

Definition 2.6. The functional I is *stationary* at y if $\partial_\eta I(y) = 0$ for all test functions η defined on $[a, b]$.

Lemma 2.1. *If y is a local minimum of I then I is stationary at y.*

Proof. Suppose η is any test function. Then, since η is twice continuously differentiable and $\eta(a) = 0 = \eta(b)$:

$$I(y + \epsilon\eta) \geqslant I(y) \quad \text{for small } \epsilon. \tag{2.120}$$

$$\Rightarrow \quad f : \epsilon \to I(y + \epsilon\eta) \quad \text{has a local minimum at 0.} \tag{2.121}$$

$$\Rightarrow \quad \frac{d}{d\epsilon}\bigg|_{\epsilon=0} [I(y + \epsilon\eta)] = 0. \tag{2.122}$$

$$\Rightarrow \quad \partial_\eta I(y) = 0. \tag{2.123}$$

Since this holds for all test functions, I is stationary at y. □

So, the strategy for minimizing the functional I is to try to find functions y that satisfy equation (2.122) for all test functions η on $[a, b]$. It turns out that this can be written in the form of a differential equation for y:

Theorem 2.1. *Suppose I is stationary at y. Then y satisfies the* Euler equation*:*

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = \frac{\partial F}{\partial y}, \tag{2.124}$$

subject to boundary conditions

$$\frac{\partial F}{\partial y'} = 0, \quad \text{at } x = a \text{ and } x = b. \tag{2.125}$$

Proof. Since I is stationary at y, consider the ϵ-derivatives of $I(y + \epsilon\eta)$, where η is any test

function, evaluated at $\epsilon = 0$.

$$I(y + \epsilon\eta) = \int_a^b F(x, y(x) + \epsilon\eta(x), y'(x) + \epsilon\eta'(x))dx. \tag{2.126}$$

$$\Rightarrow \quad \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} [I(y + \epsilon\eta)] = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \int_a^b F(x, y(x) + \epsilon\eta(x), y'(x) + \epsilon\eta'(x))dx, \tag{2.127}$$

Then, on applying Taylor's theorem, it can be seen that

$$\begin{aligned} 0 = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \int_a^b \Big(& F(x, y(x), y'(x)) + \epsilon\eta(x)\frac{\partial F}{\partial y}(x, y(x), y'(x)) \\ & + \epsilon\eta'(x)\frac{\partial F}{\partial y'}(x, y(x), y'(x)) + O(\epsilon^2) \Big) dx, \end{aligned} \tag{2.128}$$

$$= \int_a^b \left(\eta(x)\frac{\partial F}{\partial y}(x, y(x), y'(x)) + \eta'(x)\frac{\partial F}{\partial y'}(x, y(x), y'(x)) + O(\epsilon) \right) dx \Big|_{\epsilon=0}, \tag{2.129}$$

$$= \int_a^b \left(\frac{\partial F}{\partial y}(x, y(x), y'(x)) - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}(x, y(x), y'(x)) \right) \right) \eta(x)dx + \eta\frac{\partial F}{\partial y'}\Big|_{x=a}^{x=b}, \tag{2.130}$$

where the last equation is obtained by integrating by parts. The term that is not under the integral sign on the right-hand-side vanishes if F satisfies the boundary conditions specified in equation (2.125). Since $\eta(x)$ can be any test function it can be shown (by contradiction) that the remainder of the integrand must be equal to zero.[14] □

It is relatively straightforward to generalize theorem 2.1. Suppose that the functional depends on more than one function, i.e. $I = I(y_1(x), \ldots, y_n(x))$. Then the appropriate Euler equations are

$$\frac{d}{dx}\left(\frac{\partial F}{\partial y_i'}\right) = \frac{\partial F}{\partial y_i}, \quad i = 1, \ldots, n. \tag{2.131}$$

If y is a function of a vector $\mathbf{x} = (x_1, \ldots, x_n)$ then the Euler equations for a functional $I(y(\mathbf{x}))$ can be shown to be

$$\nabla_x \cdot (\nabla_p F) = \frac{\partial F}{\partial y}, \tag{2.132}$$

where

$$\nabla_x = \left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right)^T, \quad \nabla_p = \left(\frac{\partial}{\partial p_1}, \ldots, \frac{\partial}{\partial p_n}\right)^T, \tag{2.133}$$

[14] This observation is known as the *fundamental lemma of the calculus of variations.*

and

$$p_i = \frac{\partial y}{\partial x_i'}. \tag{2.134}$$

In both cases appropriate boundary conditions can be determined by working through the argument used in the proof of theorem 2.1.

2.3.2 Integral Constraints and the Euler-Lagrange Equation

Suppose now that the functional in equation (2.118) is to be minimized subject to the integral constraint

$$J(y) = \int_a^b G(x, y(x), y'(x))dx = c, \tag{2.135}$$

for some constant c. In general the constrained minimization problem will not be solved by the function given by the Euler equations (since this won't satisfy the constraint). The solution requires the introduction of a Lagrange multiplier.

Lemma 2.2. *Suppose y is a local minimum of equation (2.118) subject to the constraint given in equation (2.135). Then, if y is not stationary for J, there exists a unique real-valued Lagrange multiplier k (independent of η) such that*

$$\partial_\eta I(y) = k\partial_\eta J(y), \tag{2.136}$$

for all test functions η.

Proof. Since J is not stationary at y there exists some test function ξ such that $\partial_\xi J(y) \neq 0$. Now, suppose that η is any test function and set

$$f(\epsilon_1, \epsilon_2) = I(y + \epsilon_1\xi + \epsilon_2\eta), \tag{2.137}$$

$$g(\epsilon_1, \epsilon_2) = J(y + \epsilon_1\xi + \epsilon_2\eta). \tag{2.138}$$

Since y is a local minimum for the I subject to J it can be seen that $(\epsilon_1, \epsilon_2) = (0, 0)$ minimizes f subject to g. Then, provided that $\nabla g(0, 0) \neq \mathbf{0}$, Lagrange's theorem can be applied to show that there exists a k satisfying

$$\frac{\partial f}{\partial \epsilon_i}(0, 0) = k\frac{\partial g}{\partial \epsilon_i}(0, 0), \quad i = 1, 2. \tag{2.139}$$

Observe that

$$\frac{\partial g}{\partial \epsilon_1} = \frac{d}{d\epsilon_1}\bigg|_{\epsilon_1=0} J(y + \epsilon_1 \xi) = \partial_\xi J(y) \neq 0, \tag{2.140}$$

so that $\nabla g \neq 0$ and equation (2.139) is valid. Then taking $i = 2$ in equation (2.139) gives equation (2.136), while taking $i = 1$ determines k uniquely and independently of η (this again relies on equation (2.140) being satisfied). □

Equation (2.136) is equivalent to the statement $\partial_\eta(I - kJ)(y) = 0$ for all test functions η. Hence,

$$I - kJ = \int_a^b (F - kG)dx, \tag{2.141}$$

is stationary at y. So, y satisfies the Euler equation for $I - kJ$ and, hence,

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\left(\frac{\partial F}{\partial y'}\right) = k\left(\frac{\partial G}{\partial y} - \frac{d}{dx}\left(\frac{\partial G}{\partial y'}\right)\right). \tag{2.142}$$

This equation is known as the *Euler-Lagrange equation.* The Euler-Lagrange equation generalizes to higher dimensions for x and y in exactly the same way as the Euler equation does.

2.4 The Finite Element Method

Very often the systems of PDEs resulting from mathematical models of physical situation cannot be solved analytically. Hence, methods for solving these systems numerically are sought. In this section one such method — that of *finite elements* — is described. The method can be split into three basic parts: first the region of interest must be converted into a computational domain; secondly the PDEs must be reformulated and discretized on the computational domain; finally the resulting equations must be solved. These three steps are described in Sections 2.4.1, 2.4.2 and 2.4.3 respectively. The description that is given here concentrates on solving (elliptic and hyperbolic) linear second-order systems of PDEs. This is because the models of Section 2.1 give rise to equations of this type. A more complete description of the implementation and underlying theory can be found in the textbooks by Johnson [40] and Reddy [71].

Throughout this section the problem of solving the governing equations resulting from steady state, 2-d linear elasticity (with no body forces) is used as an example to illustrate

how to formulate and solve a system of PDEs using finite elements. Expanding the matrix equation (2.61) gives the governing system

$$\frac{\partial}{\partial x}\left((\lambda+2\mu)\frac{\partial u}{\partial x}+\lambda\frac{\partial v}{\partial y}\right)+\frac{\partial}{\partial y}\left(\mu\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)\right)=0, \tag{2.143}$$

$$\frac{\partial}{\partial x}\left(\mu\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)\right)+\frac{\partial}{\partial y}\left(\lambda\frac{\partial u}{\partial x}+(\lambda+2\mu)\frac{\partial v}{\partial y}\right)=0, \tag{2.144}$$

where x, y are Cartesian coordinates, $\lambda(x,y)$, $\mu(x,y)$ are the (prescribed) Lamé constants and $u(x,y)$, $v(x,y)$ are the displacements in the x- and y-directions respectively. Here, the material is assumed to occupy the unit square, i.e. $(x,y)\in[0,1]\times[0,1]$ and the problem has the following boundary conditions:

$$u=0,\quad v=0,\quad \text{on } y=0, \tag{2.145}$$

$$u=0,\quad v=1-\delta,\quad \text{on } y=1, \tag{2.146}$$

$$\sigma_{11}=\sigma_{12}=0,\quad \text{on } x=0,1, \tag{2.147}$$

where σ is defined in terms of the displacements by equations (2.24) and (2.52). Clearly this system is not analytically tractable for general λ and μ, so it is necessary to solve it numerically.

2.4.1 Triangulation and Mesh Generation

The finite element method calculates numerical solutions using a mesh. Hence the first step in any implementation is to generate the mesh to be used. In this work the mesh will be made up of triangular (in 2-d) or tetrahedral (in 3-d) *elements*. The elements are assumed to be disjoint and their union covers the whole of the computational domain.[15] Since for a domain Ω with curved edges it is impossible to cover the whole domain (and nothing else) using only triangles (or tetrahedra) the computational domain may differ slightly from the physical domain. Usually the vertices of the elements are defined to be the *nodes* of the mesh, i.e. the points at which the solution is approximated. Additional points within an element or on its edges (or faces) may also be defined to be nodes. The solution to the differential equation is approximated first at the nodes and then extended

[15]Since the elements are conventionally assumed to be open it is more technically correct to say that the union of their closures covers the whole domain.

to all other points in the domain. For computational reasons it is usually necessary to generate the mesh in such a way that if a given point is a member of two elements then it must be either a node of both or of neither. Recently, new finite element methods have been developed that allow the solution to be discontinuous and can deal with points that are on the boundary of two elements but only a node of one. These methods are not considered in this thesis, except for in Section 8.2.1.1 as a possible future extension to deal with discontinuous Lamé constants.

Triangulation of a domain can be a difficult problem in itself, however commercial packages are available. In this work the DistMesh package developed by Persson and Strang [61] for use in Matlab is used to generate the required meshes. This method uses the built-in Delaunay triangulation routine in Matlab and optimizes node locations by applying a force-based smoothing procedure.

2.4.2 Reformulating the System of PDEs

Consider a general linear second-order linear PDE of the form

$$-\nabla.(A(\mathbf{x})\nabla u) + \mathbf{b}(\mathbf{x}).\nabla u + c(\mathbf{x})u = f(\mathbf{x}), \quad \forall \mathbf{x} \in \Omega, \tag{2.148}$$

where u is the variable to be solved for, $\mathbf{x}$ defines position within the (open) domain Ω ($\subseteq \mathbb{R}^n$ for some integer $n > 0$) and the matrix $A(\mathbf{x})$, vector $\mathbf{b}(\mathbf{x})$ and scalars $c(\mathbf{x})$, $f(\mathbf{x})$ contain the coefficients of the equation. Suppose that a numerical solution of equation (2.148) is sought subject to the following general linear boundary conditions

$$p(\mathbf{x})\frac{\partial u}{\partial n} + q(\mathbf{x})u = g(\mathbf{x}), \quad \forall \mathbf{x} \in \partial\Omega, \tag{2.149}$$

for some $p(\mathbf{x})$, $q(\mathbf{x})$, $g(\mathbf{x})$ such that at least one of p and q is nonzero at every point on the boundary and where $\partial/\partial n$ denotes the derivative in the direction normal to the boundary at any given point. Converting this boundary value problem into a form that is numerically tractable (in this case a system of linear algebraic equations) is a two step process. First the problem must be converted to *weak form* and then discretized into *finite element form.*

2.4.2.1 Weak Form

Definition 2.7. A function $u(\mathbf{x})$ that is twice continuously differentiable on Ω, continuously differentiable on $\partial\Omega$ and satisfies equations (2.148) and (2.149) is referred to as a *classical solution* of the boundary value problem.

According to the theory of PDEs (see, e.g., Ockendon *et al.* [53]) an elliptic boundary value problem will have a unique classical solution so long as A, $\mathbf{b}$, c, f, p, q, g are sufficiently smooth. However, in searching for a numerical solution on a domain made up of a series of discrete nodes it is desirable to weaken the constraints on the smoothness of a solution given in Definition 2.7.[16] Motivated by this desire, the following definition is made:

Definition 2.8. Suppose that $v(\mathbf{x})$ is a function with integrable first derivatives on Ω. Then v is a *test function* on Ω.

Now suppose that the PDE is multiplied by an arbitrary test function and integrated over the domain Ω, i.e.

$$-\int_\Omega \nabla.(A\nabla u)v dV + \int_\Omega (\mathbf{b}.\nabla u)v dV + \int_\Omega cuv dV = \int_\Omega fv dV, \quad i = 1, \ldots, M. \tag{2.150}$$

On applying the divergence theorem to the first term this becomes

$$\int_\Omega A\nabla u.\nabla v dV - \int_{\partial\Omega} A\frac{\partial u}{\partial n} v dS + \int_\Omega (\mathbf{b}.\nabla u)v dV + \int_\Omega cuv dV = \int_\Omega fv dV. \tag{2.151}$$

And, since the normal derivatives of u can be rewritten in terms of u, p, q and g according to the boundary condition equation (2.149) the problem can be further rewritten as

$$\int_\Omega A\nabla u.\nabla v dV - \int_{\partial\Omega} A\left(\frac{g - qu}{p}\right) v dS + \int_\Omega (\mathbf{b}.\nabla u)v dV + \int_\Omega cuv dV = \int_\Omega fv dV. \tag{2.152}$$

Observe that for this expression to make sense it is only necessary for u to have integrable first derivatives on Ω and be itself integrable on $\partial\Omega$. This is a much weaker condition than that imposed on a classical solution and gives rise to the following definitions:

Definition 2.9. Equation (2.152) is the *weak form* of the boundary value problem defined

[16] Note that for a finite element solution, in which the solution is defined at every point in the numerical domain (even those that are not nodes) by interpolation using appropriate basis functions the numerical solution may, in fact be twice continuously differentiable. However, it is still not a classical solution as it does not satisfy equation (2.148), but instead the finite element approximation to it (see later in this section).

by equations (2.148) and (2.149).

Definition 2.10. By extension, a function that has integrable first derivatives on Ω, is integrable on $\partial\Omega$ and satisfies equation (2.152) for all test functions v on Ω is a *weak solution* of the boundary value problem defined by equations (2.148) and (2.149).

The idea of a weak solution can trivially be extended to linear systems of PDEs: a weak form of each member equation of the system is generated and all of the analysis of this section remains valid (however, the notation is less clear). Theory to prove the existence and uniqueness of weak solutions can be found in the finite element literature, e.g. in Johnson [40] or Reddy [71]. It can also be shown that any classical solution of a boundary value problem is a weak solution and that any twice continuously differentiable weak solution is a classical solution.

Example: Linear Elasticity

Multiplying equations (2.143) and (2.144) by an arbitrary test function $\phi(x, y)$ and integrating over the unit square (Ω) gives

$$\int_\Omega \frac{\partial}{\partial x}\left((\lambda+2\mu)\frac{\partial u}{\partial x}+\lambda\frac{\partial v}{\partial y}\right)\phi dxdy+\int_\Omega \frac{\partial}{\partial y}\left(\mu\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)\right)\phi dxdy=0, \tag{2.153}$$

$$\int_\Omega \frac{\partial}{\partial x}\left(\mu\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)\right)\phi dxdy+\int_\Omega \frac{\partial}{\partial y}\left(\lambda\frac{\partial u}{\partial x}+(\lambda+2\mu)\frac{\partial v}{\partial y}\right)\phi dxdy=0. \tag{2.154}$$

Stokes' theorem in the plane can be applied to each of these equations, moving the derivatives to the test functions and adding boundary integrals, i.e.

$$\begin{aligned}0=&\int_\Omega\left((\lambda+2\mu)\frac{\partial u}{\partial x}+\lambda\frac{\partial v}{\partial y}\right)\frac{\partial\phi}{\partial x}dxdy+\int_\Omega\left(\mu\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)\right)\frac{\partial\phi}{\partial y}dxdy\\&+\int_{\partial\Omega}\left((\lambda+2\mu)\frac{\partial u}{\partial x}+\lambda\frac{\partial v}{\partial y}\right)\phi dy-\int_{\partial\Omega}\left(\mu\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)\right)\phi dx,\end{aligned} \tag{2.155}$$

$$\begin{aligned}0=&\int_\Omega\left(\mu\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)\right)\frac{\partial\phi}{\partial x}dxdy+\int_\Omega\left(\lambda\frac{\partial u}{\partial x}+(\lambda+2\mu)\frac{\partial v}{\partial y}\right)\frac{\partial\phi}{\partial y}dxdy\\&+\int_{\partial\Omega}\left(\mu\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)\right)\phi dy-\int_{\partial\Omega}\left(\lambda\frac{\partial u}{\partial x}+(\lambda+2\mu)\frac{\partial v}{\partial y}\right)\phi dx.\end{aligned} \tag{2.156}$$

The quantities in the two terms which are integrated with respect to y can be identified as σ_{11} (first equation) and σ_{12} (second equation). But, the boundary conditions prescribe these to be zero on the parts of the boundary along which y varies. Hence both of these

contributions to the boundary integral vanish. Further, since ϕ can be any test function, suppose that it is one that vanishes on the lines $y = 0, 1$. Then the remaining boundary integrals also vanish since the only regions of the boundary where x varies are the lines $y = 0, 1$.[17] So, the weak form of the 2-d linear elasticity boundary value problem defined by equations (2.143) – (2.147) is

$$\int_\Omega \left((\lambda + 2\mu)\frac{\partial u}{\partial x} + \lambda \frac{\partial v}{\partial y} \right) \frac{\partial \phi}{\partial x} dxdy + \int_\Omega \left(\mu \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) \right) \frac{\partial \phi}{\partial y} dxdy = 0, \tag{2.157}$$

$$\int_\Omega \left(\mu \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) \right) \frac{\partial \phi}{\partial x} dxdy + \int_\Omega \left(\lambda \frac{\partial u}{\partial x} + (\lambda + 2\mu)\frac{\partial v}{\partial y} \right) \frac{\partial \phi}{\partial y} dxdy = 0. \tag{2.158}$$

2.4.2.2 Finite Element Form

Having converted the boundary value problem to weak form, the next step in a finite element solution method is to discretize the problem. Suppose that there are N nodes in the finite element mesh, with coordinates $\mathbf{x}_j$, $j = 1, \ldots, N$, respectively, introduce the vector $\mathbf{U} = (U_1, \ldots, U_N)^T$ such that

$$U_j = u(\mathbf{x}_j), \tag{2.159}$$

and make the following definition:

Definition 2.11. A set of *finite element basis functions* on a numerical domain Ω^h consisting of N nodes is a set of functions $\{\phi_1, \ldots, \phi_N\}$ such that:

- $\phi_i(\mathbf{x}_j) = \delta_{ij}$, $i, j = 1, \ldots, N$, where δ_{ij} is the Kronecker delta function.
- ϕ_i is continuously differentiable everywhere in Ω^h except at the intersections of (the closures of) two or more elements, $i = 1, \ldots, N$.
- $\phi_i \equiv 0$ on all elements that (have closures that) do not contain node i.

Observe that each of the finite element basis functions are test functions on the numerical domain Ω^h (since, although potentially discontinuous, they will always have integrable first

[17] This approach is only valid on parts of the domain where Dirichlet boundary conditions are given (i.e. where the values of the dependent variable(s) are prescribed). If other forms of boundary condition are given and the value of the test function is set to zero on the boundary then the uniqueness of the weak solution is lost. Hence this approach could not have been used here to ensure that the derivatives with respect to y were zero.

derivatives). Suppose that u is approximated at any point in the numerical domain by

$$u^h(\mathbf{x}) = U_j \phi_j(\mathbf{x}). \tag{2.160}$$

Then replacing u with u_h and v with each of the ϕ_i's in turn in equation (2.152) gives

$$\begin{aligned} \int_\Omega A U_j \nabla\phi_i . \nabla\phi_j dV - \int_{\partial\Omega} A\phi_i \left(\frac{g - qU_j\phi_j}{p} \right) dS \\ + \int_\Omega U_j \phi_i (\mathbf{b}.\nabla\phi_j) v dV + \int_\Omega c U_j \phi_i \phi_j dV = \int_\Omega f \phi_i dV, \end{aligned} \tag{2.161}$$

for $i = 1, \ldots, N$. Notice that this is a linear system of N algebraic equations for the N (constant) unknowns U_j, $j = 1, \ldots, N$. Hence, it is simple to solve the system numerically (assuming that the integrals defining the coefficients can be evaluated either analytically or numerically).

Definition 2.12. Equation (2.161) is the *finite element form* of the boundary value problem defined by equations (2.148) and (2.149).

Definition 2.13. By extension, a function $u^h(\mathbf{x})$ given by equation (2.160), where the U_j satisfy equation (2.161) for $i = 1, \ldots, N$, is a *finite element solution* of the boundary value problem defined by equations (2.148) and (2.149).

Again, the theory can be trivially extended to systems of PDEs. Consider a system of M PDEs for variables u_i. Then the vector $\mathbf{U}$ is redefined so that

$$U_{(i-1)N+j} = u_i(\mathbf{x}_j), \quad i = 1, \ldots, M, \quad j = 1, \ldots, N, \tag{2.162}$$

and the method results in a linear system of MN equations for the MN unknown components of $\mathbf{U}$.

Example: Linear Elasticity

Suppose that the linear elasticity boundary value problem is to be solved on a mesh containing N nodes located at (x_j, y_j), $j = 1, \ldots, N$. Introduce the vectors $\mathbf{U}$, $\mathbf{V}$ such that

$$U_j = u(x_j, y_j), \quad V_j = v(x_j, y_j). \tag{2.163}$$

Recall that in the weak form of the linear elasticity boundary value problem the test functions were assumed to vanish on the lines $y = 0, 1$. So, define a set of N^* test functions $\{\phi_i : y_i \neq 0, 1\}$ satisfying the conditions of Definition 2.11. Observe that $N^* < N$ (since some of the mesh points will be on the lines $y = 0, 1$). Introduce the finite element approximations to u and v:

$$u^h(x, y) = U_j \phi_j(x, y), \quad v^h(x, y) = V_j \phi_j(x, y), \quad \text{if } y \neq 0, 1. \tag{2.164}$$

The sums in these expressions run from 0 to N, but exclude the values of j for which $y_j = 0, 1$. So, replacing u by u^h, v by v^h and ϕ by each of the ϕ_i's in turn in equations (2.157) and (2.158) gives the $2N^*$ equations

$$\int_\Omega \left(\left((\lambda + 2\mu) \frac{\partial \phi_i}{\partial x} \frac{\partial \phi_j}{\partial x} + \mu \frac{\partial \phi_i}{\partial y} \frac{\partial \phi_j}{\partial y} \right) U_j + \left(\lambda \frac{\partial \phi_i}{\partial x} \frac{\partial \phi_j}{\partial y} + \mu \frac{\partial \phi_i}{\partial y} \frac{\partial \phi_j}{\partial x} \right) V_j \right) dxdy = 0, \tag{2.165}$$

$$\int_\Omega \left(\mu \frac{\partial \phi_i}{\partial x} \frac{\partial \phi_j}{\partial y} + \left(\lambda \frac{\partial \phi_i}{\partial y} \frac{\partial \phi_j}{\partial x} \right) U_j + \left(\mu \frac{\partial \phi_i}{\partial x} \frac{\partial \phi_j}{\partial x} + (\lambda + 2\mu) \frac{\partial \phi_i}{\partial y} \frac{\partial \phi_j}{\partial y} \right) V_j \right) dxdy = 0, \tag{2.166}$$

for $i = 1, \ldots, N$, but excluding values of i for which $y_i = 0, 1$. In order to solve the linear system the number of equations must be increased to $2N$. This is done by using the boundary conditions at $y = 0, 1$ to give the additional $2(N - N^*)$ equations

$$U_i = 0, \quad \text{if } y_i = 0, 1, \quad i = 1, \ldots, N, \tag{2.167}$$

$$V_i = 0, \quad \text{if } y_i = 0, \quad i = 1, \ldots, N, \tag{2.168}$$

$$V_i = 1 - \delta, \quad \text{if } y_i = 1 \quad i = 1, \ldots, N. \tag{2.169}$$

Equations (2.165) – (2.169) define the finite element form of the boundary value problem given by equations (2.143) – (2.147).

2.4.3 Numerical Methods

The system of equations (2.161) can be written in matrix form as

$$A\mathbf{U} = \mathbf{f}, \tag{2.170}$$

where A is known as the *global stiffness matrix*, $\mathbf{f}$ as the *global load vector* and they have entries given by

$$A_{ij} = \int_\Omega (A\nabla\phi_i.\nabla\phi_j + \phi_i(\mathbf{b}.\nabla\phi_j) + c\phi_i\phi_j)dV + \int_{\partial\Omega} \frac{Aq\phi_i\phi_j}{p} dS, \quad i,j = 1,\ldots,N, \quad (2.171)$$

$$f_i = \int_\Omega f\phi_i dV + \int_{\partial\Omega} \frac{Ag\phi_i}{p} dS, \quad i = 1,\ldots,N. \quad (2.172)$$

Many software packages exist for solving linear matrix equations of this form. In this work the built-in functions of Matlab (in particular the `mldivide`, or 'backslash', function[18]) are used.

2.4.3.1 Assembly of the Global Stiffness Matrix and Load Vector

The entries of the global stiffness matrix and load vector are calculated by adding the contributions from each element in turn. Recall that the numerical domain is split into discrete elements. Suppose there are m elements, labelled Δ_k, $k = 1,\ldots,m$ in the mesh. Suppose further that element k contains $N_k (< N)$ nodes, numbered n_i, $1 \leqslant n_i \leqslant N$, for $i = 1,\ldots,N_k$. Then define the *element stiffness matrices* A^k and *element load vectors* $\mathbf{f}^k$ to be the matrices and vectors with entries

$$A^k_{ij} = \int_{\Delta_k} (A\nabla\phi_{n_i}.\nabla\phi_{n_j} + \phi_{n_i}(\mathbf{b}.\nabla\phi_{n_j}) + c\phi_{n_i}\phi_{n_j})dV + \int_{\partial\Delta_k\cap\partial\Omega} \frac{Aq\phi_{n_i}\phi_{n_j}}{p} dS, \quad (2.173)$$

$$f^k_i = \int_{\Delta_k} f\phi_{n_i} dV + \int_{\partial\Delta_k\cap\partial\Omega} \frac{Ag\phi_{n_i}}{p} dS, \quad (2.174)$$

for $i,j = 1,\ldots,N_k$ and $k = 1,\ldots,m$. Then, since the elements are disjoint and (their closures) cover the whole of Ω^h the global stiffness matrix and load vector are given by

$$A = \sum_{k=1}^{m} L^k A^k \left(L^k\right)^T, \quad (2.175)$$

$$\mathbf{f} = \sum_{k=1}^{m} L^k \mathbf{f}^k, \quad (2.176)$$

where L^k is an N-by-N_k matrix with all zero entries, except that $L^k_{i,n_i} = 1$ for $i = 1,\ldots,N_k$.

[18]The `mldivide` function in Matlab solves linear systems by simple Gaussian elimination. Full details of how it works can be obtained by typing `help mldivide` in Matlab.

In order to construct the element stiffness matrices and load vectors it is necessary to evaluate the various integrals in equations (2.173) and (2.174). In order to do this efficiently each element is mapped onto a *canonical element* on which the values of the basis functions and their derivatives can easily be calculated (and, in general, are the same for all elements of the mesh). The integrations can then be performed on this canonical element and then mapped back to the original matrix. In 2-d each element is mapped onto the *canonical triangle* $\bar{\Delta} = \{(\xi, \eta) : \xi > 0, \eta > 0, \xi + \eta < 1\}$. Then, the mapping back from the canonical triangle to an arbitrary triangle Δ with corners located at $\mathbf{x}_i = (x_i, y_i)$, $i = 1, 2, 3$, is given by

$$\mathbf{x} = (1 - \xi - \eta)\mathbf{x}_1 + \xi\mathbf{x}_2 + \eta\mathbf{x}_3, \tag{2.177}$$

for a general point $\mathbf{x} \in \Delta$. This transformation has Jacobi matrix

$$J = \frac{\partial(x, y)}{\partial(\xi, \eta)} = \begin{pmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{pmatrix}. \tag{2.178}$$

Hence, an integral over Δ is given in terms of the equivalent integral over the canonical triangle by

$$\int_{\Delta} f_x(x, y) dx dy = \det J \int_{\bar{\Delta}} f_\xi(\xi, \eta) d\xi d\eta, \tag{2.179}$$

where $f_x(x, y) = f_\xi(\xi, \eta)$ when x, y, ξ and η satisfy equation (2.177). Note that the vertices must be numbered so that the Jacobian determinant $\det J$ is positive. In practice this means that they are numbered anti-clockwise around $\bar{\Delta}$, starting with the node at $(0, 0)$.

In 3-d, the *canonical tetrahedron* is $\bar{\Delta} = \{(\xi, \eta, \zeta) : \xi > 0, \eta > 0, \zeta > 0, \xi + \eta + \zeta < 1\}$, the mapping back to a general tetrahedron Δ with corners at $\mathbf{x}_i = (x_i, y_i, z_i)$, $i = 1, \ldots, 4$ is

$$\mathbf{x} = (1 - \xi - \eta - \zeta)\mathbf{x}_1 + \xi\mathbf{x}_2 + \eta\mathbf{x}_3 + \zeta\mathbf{x}_4, \tag{2.180}$$

for a general point $\mathbf{x} \in \Delta$, the Jacobi matrix is

$$J = \frac{\partial(x, y, z)}{\partial(\xi, \eta, \zeta)} = \begin{pmatrix} x_2 - x_1 & x_3 - x_1 & x_4 - x_1 \\ y_2 - y_1 & y_3 - y_1 & y_4 - y_1 \\ z_2 - z_1 & z_3 - z_1 & z_4 - z_1 \end{pmatrix}, \tag{2.181}$$

and an integral over Δ in terms of the equivalent integral over Δ is given by

$$\int_{\Delta} f(x,y,z)dxdy = \det J \int_{\bar{\Delta}} f(\xi,\eta,\zeta)d\xi d\eta d\zeta, \tag{2.182}$$

where $f_x(x,y,z) = f_\xi(\xi,\eta,\zeta)$ when x, y, z, ξ, η and ζ satisfy equation (2.180). Again the vertices must be numbered to ensure that $\det J > 0$. This is done by first numbering the nodes in the triangle in the plane $z = 0$ in an anti-clockwise direction starting with the node at $(0,0,0)$. The node $(0,0,1)$ is then numbered 4.

2.4.3.2 Finite Element Basis Functions on the Canonical Triangle

The most commonly used finite element basis functions are *piecewise polynomial.* i.e. functions that are polynomials (and, hence, infinitely differentiable) on each element, but not necessarily across interfaces between elements. The degree of the polynomials used as basis functions, together with the dimension of the elements, determines the number of nodes per element required in the mesh — one node is required for each coefficient of the polynomial to ensure that the basis functions are uniquely defined on each element (e.g. for a quadratic basis function on an element in 2-d six nodes are required since the quadratic can be written as $\phi(x,y) = c_1x^2 + c_2y^2 + c_3xy + c_4x + c_5y + c_6$).[19] Since all of the elements of the mesh can be mapped to the canonical element it is necessary only to define the finite element basis functions for the nodes of the canonical element in order to completely determine the basis functions for all of the nodes of the mesh. In this section some simple sets of piecewise polynomial basis functions on the canonical elements ($\bar{\Delta}$) in 2-d and 3-d are defined. The basis functions used in simulations later in this work are selected only from those introduced here.

Piecewise Constant Basis Functions

The simplest of all basis functions are those that are constant on each element. For a constant only one node is required to determine its value everywhere on the triangle. So,

[19] Recall that $\phi_i(\mathbf{x}_j) = \delta_{ij}$, $i,j = 1,\ldots,N_k$, where $\mathbf{x}_j$ are the locations of the nodes, so each node gives one piece of data about ϕ_i on the element.

on the canonical element the constant basis function is

$$\phi(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \bar{\Delta}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.183}$$

These functions are unique amongst those described here in that they are discontinuous between elements. Their first derivatives are also zero, which makes them unsuited to the solution of second order PDEs. Hence they can only used to interpolate coefficient functions that are differentiated once in the PDEs (i.e. not at all in weak form). An example of such functions are the Lamé constants λ and μ in equations (2.143) and (2.144) for linear elasticity.

Piecewise Linear Basis Functions

The next set of piecewise polynomial functions are those that are linear on each element. On the canonical triangle in 2-d three nodes are required, usually located at the vertices of the triangle, and the basis functions are given by

$$\phi_1(\xi, \eta) = 1 - \xi - \eta, \tag{2.184}$$

$$\phi_2(\xi, \eta) = \xi, \tag{2.185}$$

$$\phi_3(\xi, \eta) = \eta, \tag{2.186}$$

on the canonical triangle. In 3-d, four nodes are required, again they are conventionally at the vertices of the tetrahedron, and the basis functions are

$$\phi_1(\xi, \eta, \zeta) = 1 - \xi - \eta - \zeta, \tag{2.187}$$

$$\phi_2(\xi, \eta, \zeta) = \xi, \tag{2.188}$$

$$\phi_3(\xi, \eta, \zeta) = \eta, \tag{2.189}$$

$$\phi_4(\xi, \eta, \zeta) = \zeta, \tag{2.190}$$

on the canonical tetrahedron. These basis functions are continuous on Ω^h, but have discontinuous first derivatives.

Piecewise Quadratic Basis Functions

The final set of basis functions used in this work are those that are quadratic on each element. In 2-d, six nodes are required per element and these are conventionally located at the vertices of the triangles and the midpoints of their edges. So, the basis functions on the canonical triangle are

$$\phi_1(\xi,\eta) = 2(1-\xi-\eta)(\tfrac{1}{2}-\xi-\eta), \tag{2.191}$$
$$\phi_2(\xi,\eta) = 4\xi(1-\xi-\eta), \tag{2.192}$$
$$\phi_3(\xi,\eta) = 2\xi(\xi-\tfrac{1}{2}), \tag{2.193}$$
$$\phi_4(\xi,\eta) = 4\xi\eta, \tag{2.194}$$
$$\phi_5(\xi,\eta) = 2\eta(\eta-\tfrac{1}{2}), \tag{2.195}$$
$$\phi_6(\xi,\eta) = 4\eta(1-\xi-\eta), \tag{2.196}$$

where the nodes are again numbered anti-clockwise around the triangle, starting at $(0,0)$. In 3-d, ten nodes are required, again located at the vertices and mid-points of the edges of the tetrahedral elements. On the canonical tetrahedron the basis functions are given by

$$\phi_1(\xi,\eta,\zeta) = 2(1-\xi-\eta-\zeta)(\tfrac{1}{2}-\xi-\eta-\zeta), \tag{2.197}$$
$$\phi_2(\xi,\eta,\zeta) = 4\xi(1-\xi-\eta-\zeta), \tag{2.198}$$
$$\phi_3(\xi,\eta,\zeta) = 2\xi(\xi-\tfrac{1}{2}), \tag{2.199}$$
$$\phi_4(\xi,\eta,\zeta) = 4\xi\eta, \tag{2.200}$$
$$\phi_5(\xi,\eta,\zeta) = 2\eta(\eta-\tfrac{1}{2}), \tag{2.201}$$
$$\phi_6(\xi,\eta,\zeta) = 4\eta(1-\xi-\eta-\zeta), \tag{2.202}$$
$$\phi_7(\xi,\eta,\zeta) = 4\zeta(1-\xi-\eta-\zeta), \tag{2.203}$$
$$\phi_8(\xi,\eta,\zeta) = 4\xi\zeta, \tag{2.204}$$
$$\phi_9(\xi,\eta,\zeta) = 4\eta\zeta, \tag{2.205}$$
$$\phi_{10}(\xi,\eta,\zeta) = 2\zeta(\zeta-\tfrac{1}{2}), \tag{2.206}$$

where the nodes are numbered 1–6 anti-clockwise around the triangle in the plane $z=0$, starting from $(0,0,0)$, then 7–9 anti-clockwise around the triangle in the plane $z=\frac{1}{2}$, starting from $(\frac{1}{2},0,0)$ and finally the node at $(0,0,1)$ is numbered 10.

2.4.3.3 Numerical Integration: Gauss Quadrature

With the values of all of the terms in the integrands of equations (2.173) and (2.174) being determined by the selection of the basis functions, the final step in the finite element solution of a boundary value problem is the calculation of the integrals. For some forms of the coefficients $A(\mathbf{x})$, $\mathbf{b}(\mathbf{x})$, $c(\mathbf{x})$, $f(\mathbf{x})$, $g(\mathbf{x})$, $p(\mathbf{x})$ and $q(\mathbf{x})$ (e.g. if they are constants on each element) it may be possible to calculate the integrals exactly. However, in general it is necessary to compute them numerically. In this work numerical integration is performed using *Gauss quadrature*. A full description of the application of Gauss quadrature can be found in, e.g., Süli and Mayers [81], but the basic principle is that an integral can be approximated using a weighted average of the values of the integrand at various points throughout the region of integration.

By appropriate selection of the (n per spatial dimension) points and weights all the integrals of all polynomials of degree $2n - 1$ and below can be determined exactly by Gauss quadrature. Such points are referred to as *Gauss points*. In this work either two or three Gauss points per dimension are used (i.e. four or nine points in 2-d, eight or 27 in 3-d). This gives exact values of integrals for polynomials up to degree 3 or 4. The locations of the Gauss points on the canonical elements and their corresponding weights are shown in Tables 2.1 – 2.4. The derivations of these locations and weights can be found in Süli and Mayers [81].

Point	ξ_p	η_p	w_p
1	0.166667	0.211325	0.197169
2	0.622008	0.211325	0.197169
3	0.044658	0.788675	0.052831
4	0.166667	0.788675	0.052831

Table 2.1: Location (ξ_p, η_p) and weights w_p, $p = 1, \ldots, 4$ of the Gauss points for a 4-point Gaussian quadrature scheme on the canonical triangle $\bar{\Delta}$ in 2-d.

Point	ξ_p	η_p	w_p
1	0.100000	0.112702	0.068464
2	0.443649	0.112702	0.109543
3	0.787298	0.112702	0.068464
4	0.056351	0.500000	0.061728
5	0.250000	0.500000	0.098765
6	0.443649	0.500000	0.061728
7	0.012702	0.887298	0.008696
8	0.056351	0.887298	0.013914
9	0.100000	0.887298	0.008696

Table 2.2: Location (ξ_p, η_p) and weights w_p, $p = 1, \ldots, 9$ of the Gauss points for a 9-point Gaussian quadrature scheme on the canonical triangle $\bar{\Delta}$ in 2-d.

Point	ξ_p	η_p	ζ_p	w_p
1	0.166667	0.211325	0.131446	0.061320
2	0.622008	0.211325	0.035221	0.016431
5	0.166667	0.211325	0.490563	0.061320
6	0.622008	0.211325	0.131446	0.016431
3	0.044658	0.788675	0.035221	0.004403
4	0.166667	0.788675	0.009437	0.001180
7	0.044658	0.788675	0.131446	0.004403
8	0.166667	0.788675	0.035221	0.001180

Table 2.3: Location (ξ_p, η_p, ζ_p) and weights w_p, $p = 1, \ldots, 8$, of the Gauss points for an 8-point Gaussian quadrature scheme on the canonical tetrahedron $\bar{\Delta}$ in 3-d.

Point	ξ_p	η_p	ζ_p	w_p
1	0.100000	0.112702	0.088730	0.017967
2	0.443649	0.112702	0.050000	0.016199
3	0.787298	0.112702	0.011270	0.002282
4	0.100000	0.112702	0.393649	0.017967
5	0.443649	0.112702	0.221825	0.016199
6	0.787298	0.112702	0.050000	0.002282
7	0.100000	0.112702	0.698569	0.017967
8	0.443649	0.112702	0.393649	0.016199
9	0.787298	0.112702	0.088730	0.002282
10	0.056351	0.500000	0.050000	0.009128
11	0.250000	0.500000	0.028175	0.008230
12	0.443649	0.500000	0.006351	0.001159
13	0.056351	0.500000	0.221825	0.009128
14	0.250000	0.500000	0.125000	0.008230
15	0.443649	0.500000	0.028175	0.001159
16	0.056351	0.500000	0.393649	0.009128
17	0.250000	0.500000	0.221825	0.008230
18	0.443649	0.500000	0.050000	0.001159
19	0.012702	0.887298	0.011270	0.000289
20	0.056351	0.887298	0.006351	0.000261
21	0.100000	0.887298	0.001431	0.000036
22	0.012702	0.887298	0.050000	0.000289
23	0.056351	0.887298	0.028175	0.000261
24	0.100000	0.887298	0.006351	0.000036
25	0.012702	0.887298	0.088730	0.000289
26	0.056351	0.887298	0.050000	0.000261
27	0.100000	0.887298	0.011270	0.000036

Table 2.4: Location (ξ_p, η_p, ζ_p) and weights w_p, $p = 1, \ldots, 27$, of the Gauss points for a 27-point Gaussian quadrature scheme on the canonical tetrahedron $\bar{\Delta}$ in 3-d.

Chapter 3

Previous Work

In this chapter some existing methods for ultrasound elasticity imaging are discussed. Almost all of the published techniques for reconstructing the distribution of tissue stiffness assume that the deformation is linearly elastic, plane strain (see Section 4.1.2 for a full description of the plane strain assumption) and use some form of iterative technique, i.e. a candidate stiffness distribution is assumed and the deformation predicted by the model using this distribution is compared to the observed deformation. The main differences between the methods are:

- The method used to update the estimated stiffness distribution between iterates.
- How the similarity between calculated and observed displacement fields is calculated.

These methods are discussed in Section 3.1. Then, in Section 3.2 the one direct solution method that does exist is described and in Section 3.3 the advantages and disadvantages of each of the solution methods are considered. There is a small amount of theoretical work on directly solving for the stiffness in terms of the observed displacements, which is discussed in Section 3.4. It is this theory that is used as a basis for developing a new method for elasticity imaging in Chapter 4. The validity of the linear elasticity assumption and some alternative elastic models of soft tissue mechanics are discussed in Appendix A. In Appendix B some possible viscoelastic and mixture models of soft tissue are introduced.

3.1 Iterative Reconstruction Techniques

In this section a selection of the work on iterative elasticity imaging techniques is reviewed. Elasticity imaging (often referred to as *elastography*) has been an active field of research since the late 1980's. However, most of the techniques developed in this time have the same basic framework, which consists of the following five-step process:

(i) Calculate (or at least estimate) a displacement field $\mathbf{u}_{\text{image}}(x, y)$ from the ultrasound images.

(ii) Initialize the distributions of $\lambda(x, y)$ and $\mu(x, y)$ (or just μ if the material is assumed to be incompressible).

(iii) Simulate the linear elasticity forward problem using the current distributions of λ and μ. This gives a displacement field $\mathbf{u}_{\text{simulated}}(x, y)$.

(iv) Calculate $S(\mathbf{u}_{\text{image}}, \mathbf{u}_{\text{simulated}})$, where $S(\mathbf{u}, \mathbf{v})$ is a similarity measure, defined so that $S(u, v)$ is smaller the more similar the two displacement fields $\mathbf{u}$ and $\mathbf{v}$ are.

(v) If $|S(\mathbf{u}_{\text{image}}, \mathbf{u}_{\text{simulated}})| < \delta$ for some given value of δ, then stop and return the current λ and μ. Otherwise, update λ and μ and goto (iii).

Various different schemes to update λ and μ exist. These converge at different rates and give solutions that have varying levels of smoothness. Similarly, different similarity measures can be used. The choice of similarity measure defines exactly whether certain types of difference between the displacement fields are more acceptable than others (e.g. larger errors may be allowable close to the boundary).

In order to construct an elasticity image using an iterative method it is necessary to solve the linear elasticity forward problem many times. Using a direct solution method (i.e. only numerically solving one system of equations) would greatly reduce the time that it takes to generate the image. Developing such a method is the aim of this work, so the discussion of iterative techniques given here is relatively brief and aims to provide a broad overview of the field rather than concentrating on the finer details of largely similar imaging methods. The discussion is split into two parts: in Section 3.1.1 early methods based on Newton-Raphson iterations are described; then in Section 3.1.2 a bisection method is discussed; finally, a method that depends directly on the simulated stress field is considered in Section 3.1.3. A more complete description of the field can be found in the review paper by Ophir

et al. [56].

3.1.1 Newton's Method

Many of the first iterative solution methods, such as those of Parker *et al.* [58], Ophir *et al.* [57] and Céspedes *et al.* [18] assumed some degree of prior knowledge of the physiology of the tissue (i.e. the presence and approximate location and geometry of a tumour). Using prior knowledge in this way makes it easier to determine exactly how to deform the region of interest in order to give the clearest elasticity image (since the strain field can be chosen in such a way as to improve the contrast in the region where the tumour is located). These methods provided a useful foundation from which more advanced techniques of elasticity imaging could be developed, but they suffered from severe limitations in terms of diagnostic power — real tumours may be located anywhere in the region of interest (or not be present at all) and are very likely to be irregular in shape. The early methods were a mixture of finite element and finite difference methods. The more recent methods are almost exclusively finite element methods.

One of the most successful of the early methods was developed by Kallel and Bertrand [41]. The method uses only axial displacements, i.e. displacements in the direction in which the tissue is compressed by the ultrasound probe, (primarily because lateral displacement estimates are less reliable) and assumes that the tissue is nearly incompressible — having a fixed Poisson's ratio of 0.495. The Young's modulus distribution is updated at each iteration using Newton's method (see, e.g. Süli and Mayers [81]).[1] The similarity measure used is the least squares error (LSE), i.e.

$$S(u_{\text{image}}, u_{\text{simulated}}) = ||u_{\text{image}} - u_{\text{simulated}}||^2, \tag{3.1}$$

where $||u(x,y)||$ represents the 2-norm of the displacement field u, defined as

$$||u(x,y)||^2 = \int_\Omega [u(x,y)]^2 dxdy, \tag{3.2}$$

and where Ω is the region being imaged. Choosing this similarity measure has the effect of penalizing large displacements extremely heavily. It is one of the most commonly used

[1] Newton's method is sometimes referred to as the Newton-Raphson or Newton-Fourier method in the literature. All of these terms refer to exactly the same iterative technique for solving systems of nonlinear algebraic equations.

similarity measures in elasticity imaging.

In order to derive the appropriate Newton's method iteration, suppose that a numerical estimate of the axial displacement is $\mathbf{U} = (U_j)$, where $U_j \approx u_{\text{simulated}}(x_j, y_j)$ and (x_j, y_j) is the coordinate of the jth node of the finite element mesh. Suppose further that $\hat{\mathbf{U}}$ is the equivalent for the displacement estimate derived from the images. Then the LSE can be approximated as

$$S(\hat{\mathbf{U}}, \mathbf{U}) \approx |\hat{\mathbf{U}} - \mathbf{U}|^2. \tag{3.3}$$

Similarly suppose that the Young's modulus is approximated by $\mathbf{E} = (E_j)$, where $E_j \approx E(x_j, y_j)$ and define the Jacobi matrix $J = (J_{ij})$ by setting

$$J_{ij} = \frac{\partial U_i}{\partial E_j}. \tag{3.4}$$

Then the update scheme resulting from Newton's method can be seen to be

$$\mathbf{E}_{k+1} = \mathbf{E}_k - \Delta\mathbf{E}_k, \tag{3.5}$$

where

$$\Delta\mathbf{E}_k = \left[J_k^T J_k\right]^{-1} J_k^T (\hat{\mathbf{U}} - \mathbf{U}_k), \tag{3.6}$$

and the subscript k denotes the estimates at iteration k.

Kallel and Bertrand [41] tested their method on both ideal displacement fields and those calculated from real ultrasound images. In general, the matrices $J_k^T J_k$ turn out to be ill-conditioned. However, regularization techniques have been included in the method to allow them to be inverted. The method is very effective in reconstructing the input parameters for the ideal displacement fields. Although less successful for the real displacement fields stiffer regions can still be seen in the expected locations. The method was extended by Doyley *et al.* [22] to include lateral as well as axial displacements (lateral displacement estimation techniques had improved sufficiently between 1996 and 2000 to make this worthwhile). The quality of the resulting elasticity images for displacement fields estimated from real images of gelatin phantoms was noticeably better than those of Kallel and Bertrand [41]. This indicates that it is important to include as much of the available displacement data as possible in order to produce the best image.

3.1.2 A Bisection Method

Although Newton's method is a widely used technique for finding zeros of a nonlinear function (the similarity measure in elasticity imaging) and converges comparatively quickly in the region of the zero it does have some disadvantages over alternative methods. Most importantly, it is not guaranteed to succeed. If the initial guess for the parameter distributions is not close enough to the solution then the method may not converge. This explains why some *a priori* knowledge of the solution is required in the method of Kallel and Bertrand [41].

Fu *et al.* [26] have proposed a bisection method for updating the parameter values, based on the observation that the Young's modulus of soft tissue generally falls within a known range of values (1 kPa to 200 kPa). At each iteration a dynamic linear elasticity forward problem is solved in which the tissue is vibrated along one edge. The amplitude of the resulting oscillations throughout the tissue are calculated and compared to the observed amplitudes in ultrasound images of the tissue being investigated being subjected to similar vibrations.[2] The similarity measure used is then defined to be

$$S(x, y) = A_e(x, y) - A_m(x, y), \tag{3.7}$$

where A_e is the estimated (from the simulations) amplitude and A_m is the measured (from the images) amplitude. Since stiffer regions would be expected to experience oscillations with smaller amplitudes the sign of S indicates whether the current estimate for the Young's modulus is too low or too high.

The iterative scheme works by defining a numerical mesh on which the solution is to be calculated and gradually narrowing the range within which the value of the Young's modulus may lie at each point of the mesh. Termination occurs when the difference between the upper and lower limits reaches some threshold (related to their absolute values) or after a given number of iterations. More explicitly, the method can be summarized as follows:

(i) For each mesh point, choose initial values of $E_{\min}$ and $E_{\max}$. Typically these values are defined by the theoretical range given above, i.e. $E_{\min} = 1$ kPa, $E_{\max} = 200$ kPa. The Poisson's ratio takes the fixed value $\mu = 0.495$.

[2]The vibrations are generated by the imaging process itself: high energy ultrasound is used so that the motion of the ultrasound probe as the pulse is emitted perturbs the tissue surface. This type of elasticity imaging is sometimes referred to as *vibration sonoelastography*.

(ii) At each point, define an estimated Young's modulus $E_e = \frac{1}{2}(E_{\min} + E_{\max})$.

(iii) Solve a linear elasticity forward problem to determine the amplitude of the oscillations in terms of the Young's modulus throughout the domain. Calculate $S(x, y)$ at each point of the mesh.

(iv) Update the range $[E_{\min}, E_{\max}]$ at each point as follows:

- If $S < 0$, then set $E_{\min} = E_e$ and leave $E_{\max}$ unchanged.
- If $S > 0$, then set $E_{\max} = E_e$ and leave $E_{\min}$ unchanged.

(v) Repeat steps (ii), (iii) and (iv) until the value of $(E_{\max} - E_{\min})/E_e$ is less than the given threshold or until the iteration number reaches a predefined value N.

Observe that the range $[E_{\min}, E_{\max}]$ halves at each iteration, so the method is guaranteed to converge and will terminate after at most N iterations (the predefined value in step (v)), so the maximum computational time can be limited before the reconstruction begins.

Fu *et al.* [26] tested their method on ultrasound images of gelatin phantoms subjected to vibration. The results show good agreement with the known stiffnesses of the phantoms, although the method is not tested on images of real tissue. The authors identify problems with displacement estimation as the most significant limitation on the effectiveness of their model. Issues relating to the 2-d approximation and the validity of the linear assumption are also considered. In particular, it is suggested that a viscoelastic model may be more appropriate.

3.1.3 A Method Based on the Simulated Stress Field

Both Newton's method and the bisection method are general techniques for finding zeros of nonlinear functions — they are not optimized to the characteristics of the elasticity imaging problem. It seems reasonable to think that the convergence of the iterative techniques can be improved by using more novel methods, developed with elasticity imaging in mind, for updating the parameter values between successive iterations.

One such method is due to Yamakawa and Shiina [91]. This method again assumes that the material is incompressible (i.e. the Poisson's ratio is set to be a fixed value close to 0.5 — although this value is not specified). At each iteration a simulation of the deformation using the current estimated Young's modulus distribution is performed and the Young's

modulus updated in a way that depends on the stress field calculated from the simulation and on the strain field measured from the ultrasound images. Unusually, the simulation at each iteration was performed in 3-d, rather than by assuming plane strain. However, it is not made clear how the displacement away from the plane of the ultrasound image is calculated. The paper makes it clear that a linear array transducer was used for the imaging (i.e. only 2-d images were available). It is possible (and, perhaps, most likely) that multiple slices were imaged, although there is no evidence to support this possibility. Alternatively, the displacement field could have been extended out of the plane of the image in some way — possibly by assuming that the displacement is the same in all planes or that it gradually decays to zero away from the plane of the image.

By using equations (2.54) and (2.55) to substitute for E and ν in the constitutive equation for linear elasticity (defined in equation (2.51)) with $i = j = 3$ it can be seen that the Young's modulus, Poisson's ratio, stress and axial strain are related by:

$$E = \frac{\sigma_{33} - \nu(\sigma_{11} + \sigma_{22})}{\epsilon_{33}}. \tag{3.8}$$

Observe that the axial strain can be directly calculated from the displacement field estimated from the ultrasound images. The stress field, however, cannot be calculated in this way (since it requires knowledge of the Young's modulus to derive it from the strain field). Nevertheless, equation (3.8) suggests a method for updating the Young's modulus. Suppose that the current Young's modulus distribution is E^k and a stress field $\hat{\sigma}^k$ is calculated from a simulated deformation using E^k. Then, define the updated Young's modulus distribution by

$$E^{k+1} = \frac{\hat{\sigma}^k_{33} - \nu(\hat{\sigma}^k_{11} + \hat{\sigma}^k_{22})}{\epsilon_{33}}, \tag{3.9}$$

where ϵ_{33} is the axial strain estimated from the ultrasound images. The method does not use an explicit similarity measure, instead it terminates when $\int_\Omega |E^{k+1} - E^k| dxdy < \delta$, for some given value of δ. The assumption underlying the method is that using the *real* displacement field rather than the *simulated* one to calculate the strain will result in a better approximation for the Young's modulus.

Yamakawa and Shiina [91] present results of applying their method to both simulated displacement fields and those resulting from ultrasound images of phantoms. No results using real tissue are given. They also compared the accuracy of the method when 2-d and 3-d simulations of the forward problem were performed at each iteration. The reconstructed

Young's modulus distributions showed good agreement with the true values for both the simulated displacement fields and the phantom experiments. Further, it was demonstrated that the results obtained using the 3-d simulations were significantly more accurate than their 2-d equivalents. However, the authors acknowledge that there is no guarantee that the scheme defined by equation (3.9) will converge (although it does do so in their examples).

3.2 A Direct Solution Method

One direct solution method for elasticity imaging that does exist is due to Raghavan and Yagle [69]. This is a fairly rudimentary technique, relying on a finite difference solution of a hyperbolic system of PDEs and, as such, has limited potential. Nevertheless, as it appears to be the only direct solution technique developed to date it is worth discussing here.

The method of Raghavan and Yagle [69] assumes that the material under the ultrasound probe is a 2-d region undergoing plane strain and that soft tisuue is linearly elastic. Using the finite difference approximation the equations of linear elasticity can be written in matrix form as $A\mathbf{x} = \mathbf{b}$, where A is a matrix containing terms dependent on the tissue parameters, $\mathbf{x} = (u, v)^T$, where u and v are the displacements in the x- and y-directions respectively and $\mathbf{b}$ is a constant vector dependent on the boundary conditions. Solutions of the linear elasticity equations with appropriate boundary conditions are compared with real axial displacement fields calculated from ultrasound images of a gelatin phantom.[3] Reasonable agreement is seen between the simulated results and the displacements seen in the phantom. This is not entirely surprising as gelatin is incompressible and the deformations are within its linearly elastic range.

In order to solve the inverse problem the linear system of equations is reformulated as $A'\mathbf{x}' = \mathbf{b}'$, where A' is a matrix containing terms dependent on u and v, $\mathbf{x}'$ is a vector consisting of the elasticity parameters (here, the Young's modulus and a pressure term to enforce the incompressibility) and $\mathbf{b}'$ is a constant vector dependent on the boundary conditions. In order to be able to specify boundary conditions the method assumes that

[3]Lateral displacements are not considered since the method assumes that soft tissue is incompressible, so it has only one degree of freedom. However, comparing simulated and observed lateral displacements would have provided evidence for the validity (or not) of the incompressibility assumption. This is generally assumed to hold but the assumption does not appear to be tested.

the elasticity parameters are known on the boundary. This will not, in general, be the case for real tissue. The method is tested on numerically calculated displacement fields and is shown to successfully recover the parameter values on coarse, regular grids. However, it does not work well on fine grids. This is not an encouraging sign — well constructed finite difference methods should be expected to converge to the exact solution as the grid spacing is reduced.

Raghavan and Yagle suggest that the method fails on fine grids because the condition number of the matrix A' becomes large. Whilst the condition number does become large for fine grids, a more likely reason for the failure of the method is that the linear elasticity inverse problem is *hyperbolic*[4] and the finite difference method used is not appropriate for solving PDEs of this type (see, e.g., Morton and Mayers [50]). It is possible that an alternative finite difference scheme could be proposed that would overcome this problem. However, such a scheme would depend on the characteristics of the inverse problem — so different schemes would be required for different displacement fields. The finite difference method also uses a regular grid with uniform spacing. This makes the method inflexible and difficult to apply to irregularly shaped domains. For these reasons a finite element solution method would be preferable (and, indeed, most other elasticity imaging methods use finite elements).

3.3 Comparing the Iterative and Direct Approaches

The iterative approach has three main advantages over a direct solution method. First, the forward problem is easier to solve numerically than the inverse problem — primarily because it is elliptic. Indeed, commercial finite element packages often include robust solvers for the linear elasticity forward problem. Secondly, since the distributions of λ and μ are specified before each simulation it is easier to deal with discontinuities in the Lamé constants between neighbouring elements. Finally, the parameter values determined by an iterative method are certain to fall within the physically allowable range (in particular, the shear modulus and Young's modulus should be positive) since the values can be checked before simulation of the forward problem is attempted and forced to lie within the correct

[4]This is shown to be the case for the incompressible inverse problem by Barbone and Gokhale [6] and, more generally, for the compressible case in Section 4.2.1 of this work.

range.[5] For a direct solution method errors in the displacement estimate may lead to the theoretical solution of the inverse problem containing regions in which the values of the parameters are physically unrealistic. However, there are methods that can be applied to constrain the range of allowable values of the parameters returned by the direct solution method (e.g. the addition of a term that takes extremely large values whenever the physical constraints are breached in the inverse problem).

On the other hand, there is a clear advantage in terms of speed in using a direct approach — only one simulation is required compared to thousands in most iterative methods. Further, if the inverse problem is solved directly then PDE and finite element theory can be used to provide information on the uniqueness and accuracy of the solution that is not available for an iterative solution. The strengths and weaknesses of the two approaches are summarized in Table 3.1. On the whole, the iterative techniques are easier to implement but the direct solution methods are very much quicker and should give solutions that can be used with greater confidence. Nevertheless, the relative simplicity of implementing iterative methods has so far outweighed the theoretical advantages and greater speed of direct methods. So, to date very few direct solution methods have been published.

3.4 Theoretical Work

In this section the existing theoretical work on solving the (linear) elasticity inverse problem for elasticity imaging is discussed. This work is limited, largely due to the fact that in elasticity theory it is usually impossible to know the displacement field away from the boundary. For example, in a typical elastic solid (e.g. a metal bar) the interior of the material is hidden from view. So, it is only with the advent of techniques for imaging the interior of solid bodies and estimating the displacements between images that the solution of the elasticity inverse problem has had any practical application. To date, elasticity imaging has been performed on an *ad hoc* basis and no underlying theory has yet emerged. Hence, little theoretical analysis of the behaviour of the elasticity inverse problem exists at the present time.

[5] In any case, the 'solutions' of the forward problem using invalid parameters are often highly irregular — for example, they are usually not at all smooth. So, an iterative solution method is unlikely to ever converge to such a solution — even if it is the true global minimum of the similarity measure (which the direct solution method will find). Instead, it is likely to converge to a local minimum of the similarity measure that corresponds to a smooth solution of the forward problem.

Iterative	Direct
Forward problem is elliptic, i.e. easier to solve numerically	Inverse problem is hyperbolic, i.e. difficult to solve numerically
Constraints can be used to ensure realistic stiffness values	Results may not be physically realistic
Can deal with possible discontinuities in the parameter values	Can be difficult to calculate discontinuous solutions directly (since classical solutions of PDEs must be continuous)
Requires multiple simulations (slow)	Only one simulation necessary (fast)
Uniqueness of the solution is not guaranteed	Uniqueness proof from PDE and finite element theory can be applied

Table 3.1: Comparison of the direct and iterative methods for calculating the stiffness parameters from a displacement estimate for use in ultrasound elasticity imaging. Strengths of each method are shown in black, weaknesses in red.

There are two major obstacles to overcome in solving the linear elasticity inverse problem directly for use in elasticity imaging, both of which were first identified by Skovoroda *et al.* [79]. First, the problem is hyperbolic, which makes it difficult to solve numerically. This is due to the fact that the solution at each point in the domain is dependent only on the solution in certain regions (the *domain of dependence*). To be effective a numerical scheme must be chosen in such a way that the numerical domain of dependence of each point is as similar as possible to the true domain of dependence (see, e.g., Morton and Mayers [50]). However, the domain of dependence is defined by the characteristics of the PDE, which are different for any two displacment fields. For an elliptic PDE, which does not have any real characteristics, this problem does not arise. The second problem in solving the elasticity inverse problem is the difficulty in obtaining appropriate boundary data to ensure that the solution is unique.

In order to overcome this difficulty in specifying boundary data Skovoroda *et al.* [79] suggested a method for determining contours along which the elasticity parameters are approximately constant (in the incompressible case). The solution within these contours can then be calculated *relative* to the value on the contour. This type of boundary condition

is known as *Goursat data.* Unfortunately, it has been shown by Barbone and Bamber [5] that the method does not work for a general displacement field. Further, Barbone and Gokhale [6] have shown that even when such contours can be found the resulting boundary value problem may still be ill-posed — it does not have continuous dependence on the given data, i.e. a small change in the boundary data or displacement field can give rise to a very large change in the solution.[6] McLaughlin and Yoon [48] have shown that the same problems with boundary conditions exist in the compressible case.

Barbone and Gokhale [6] have shown that by using multiple displacement fields much less boundary data is required in order to solve the incompressible linear elasticity inverse problem uniquely. Using only two displacement fields (i.e. three images rather than the usual two) the value of the shear modulus is only required to be known at four points. If four fields are used then the shear modulus is uniquely determined up to a multiplicative constant. The work of Barbone and Gokhale [6] comprises an important building block for the work presented in this thesis and is described in more detail in Section 4.3, where it is also extended to the case of a compressible material.

[6]This is likely to be due to the hyperbolicity of the inverse problem. Goursat data is equivalent to specifying Dirichlet data all around the boundary of a domain (the region enclosed by the contour). However, as discussed in Section 2.2.2, Cauchy data and not Dirichlet data constitutes an appropriate boundary condition for a hyperbolic system of PDEs.

Chapter 4

A New Parameter Recovery Method

In this chapter a method for determining the Lamé constants of a linearly elastic material given a displacement field (or series of related displacement fields) is described. The method is based on reformulating the usual governing equations of linear elasticity with the Lamé constants as independent variable and the displacements as known parameters and then minimizing an appropriate residual over the region of interest. This is an original contribution of the thesis.

In Section 4.1 the parameter recovery problem is formulated in terms of the inverse problem of linear elasticity. The mathematical assumptions underpinning the method are discussed and a system of first-order PDEs that can be solved to determine the Lamé constants are derived. This system is then analyzed for a general strain field in Section 4.2. It is shown to be hyperbolic and to have characteristics that are everywhere parallel to the principal strains of the deformation. Appropriate Cauchy data is discussed — in particular, the way in which the geometry of the characteristics makes it difficult to identify any obvious curve on which the data could be imposed. The possibility of using Riemann invariants to construct a solution is ruled out.

One alternative to applying boundary data is to consider the strain fields resulting from more than one deformation of the same material. This is done in Section 4.3, where it is shown that using two displacement fields means that only four pieces of data about the solution are required. If the number of displacement fields that are given is increased to four then the solution is determined completely up to a multiplicative constant. Further increase in the number of displacement fields to eliminate this constant is shown to be

impossible as no set of five displacement fields can ever be linearly independent.

The development of a numerical method for the solution of the linear elasticity inverse problem is addressed in Section 4.4. Hyperbolic systems of PDEs are not easy to solve numerically so the system is reformulated as a minimization problem with linear, elliptic, second-order Euler-Lagrange equations. The numerical solution of an elliptic system of PDEs is much simpler than in the hyperbolic case. Observe that the conversion to an elliptic equation is at the cost of making the system second-order rather than first-order. A finite element formulation of the minimization problem is described, reducing the problem to a linear system of algebraic equations that can be solved trivially. The results of solving the finite element formulation are described in Chapter 5.

4.1 The Inverse Problem for Linear Elasticity

In this section the theory of linear elasticity introduced in Section 2.1.3.1 is revisited. In Section 4.1.1 it is noted that the governing equations for a linearly elastic material can be solved either for the displacements in terms of known Lamé constants or for the Lamé constants in terms of known displacements. These two problems are defined as the *forward* and *inverse problems* respectively. The formulation of boundary conditions to make the inverse problem well-posed is postponed until later in this chapter. Some simplifying assumptions for use in ultrasound elasticity imaging are discussed in Section 4.1.2. In particular, the deformation is assumed to be *quasi-static* (i.e. there is no time dependence) and to take the form of *plane strain* (i.e. displacements are confined to the plane of the ultrasound image).

4.1.1 Formulation of the Inverse Problem

As in Section 2.1.3.1, consider a material occupying the region $\Omega \in \mathbb{R}^3$. Define $\mathbf{x} = (x_1, x_2, x_3) = (x, y, z)$ to be the (Cartesian) spatial coordinate system and the corresponding displacement field of the material to be $\mathbf{u}(\mathbf{x}) = (u_1, u_2, u_3) = (u, v, w)$. Then Cauchy's infinitesimal strain tensor is given by

$$\epsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \quad i, j = 1, 2, 3. \tag{4.1}$$

The material is assumed to be linearly elastic and isotropic, so its Cauchy stress $\sigma = (\sigma_{ij})$ is related to the strain by the constitutive equation:

$$\sigma_{ij} = \lambda \epsilon_{kk} \delta_{ij} + 2\mu \epsilon_{ij}, \quad i, j = 1, 2, 3, \tag{4.2}$$

where $\lambda(\mathbf{x})$, $\mu(\mathbf{x})$ are the Lamé constants (constant at any given point in the material as it deforms, not necessarily constant between different points), δ is the Kronecker delta and the summation convention over repeated indices is used. Recall that on applying a momentum balance a constitutive equation of this form gives rise to the following governing equations for the material:

$$\rho \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial}{\partial x_i}\left(\lambda \frac{\partial u_k}{\partial x_k}\right) + \frac{\partial}{\partial x_j}\left(\mu \frac{\partial u_i}{\partial x_j}\right) + \frac{\partial}{\partial x_j}\left(\mu \frac{\partial u_j}{\partial x_i}\right), \quad i = 1, 2, 3, \tag{4.3}$$

where $\rho(\mathbf{x})$ is the density of the material at $\mathbf{x} \in \Omega$.

At this stage it is convenient to formally define the forward and inverse problems of linear elasticity.

Definition 4.1 (The Forward Problem)**.** Given suitable boundary conditions defined on $\partial\Omega$ and $\mu(\mathbf{x})$, $\lambda(\mathbf{x})$, $\rho(\mathbf{x})$ defined throughout Ω find $\mathbf{u}(\mathbf{x}, t)$ such that equation (4.3) is satisfied.

Definition 4.2 (The Inverse Problem)**.** Given $\mathbf{u}(\mathbf{x}, t)$, ρ defined on $\Omega \times (0, T > 0)$ find $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$ such that equation (4.3) is satisfied.

Note that this statement of the inverse problem does not refer to any boundary conditions. The problem as stated here is ill-posed. Part of this work is to determine whether appropriate boundary conditions can be formulated in an elasticity imaging context and if not whether any alternative information can be given instead (e.g. the use of more than one displacement field).

4.1.2 Simplifying Assumptions

In the context of elasticity imaging it can usually be assumed that the deformation is *quasi-static* (i.e. the deformation is applied instantaneously and the tissue allowed to reach a steady state before imaging begins). In this situation the left-hand-side of equation (4.3) is zero. For the remainder of this chapter this is assumed to be the case.

Further, there are only two Lamé constants, but three governing equations, making the inverse problem appear to be over-specified in 3-d, even if appropriate boundary conditions are specified (the inverse problem in 3-d is discussed in detail in Section 6.2.1). However, ultrasound images are generally 2-d (although 3-d ultrasound[1] machines are beginning to be introduced), making it impossible to see any out of plane displacements between a pair of images. For this reason it is generally assumed in elasticity imaging that the deformation only occurs in the plane of the image, i.e. $\mathbf{u} = (u(x,y), v(x,y), 0)$. Mathematically deformations of this form are known as *plane strain* and have been studied as a special case of elasticity theory (see, e.g. Love [45]). In this work it is initially assumed that the material is undergoing a plane strain deformation. However, in Chapter 6 the validity of this assumption is considered and a general 3-d theory developed.

The governing equations for a material subject to a quasi-static, plane strain deformation are:

$$0 = \frac{\partial}{\partial x_i}\left(\lambda \frac{\partial u_k}{\partial x_k}\right) + \frac{\partial}{\partial x_j}\left(\mu \frac{\partial u_i}{\partial x_j}\right) + \frac{\partial}{\partial x_j}\left(\mu \frac{\partial u_j}{\partial x_i}\right), \quad i = 1,2, \tag{4.4}$$

where the sums over the repeated indices j and k are now for $j,k = 1,2$. For clarity it is often easiest to consider these equations written out in full in terms of the components of strain, i.e.

$$0 = \frac{\partial}{\partial x}(\lambda(\epsilon_{11} + \epsilon_{22}) + 2\mu\epsilon_{11}) + \frac{\partial}{\partial y}(2\mu\epsilon_{12}), \tag{4.5}$$

$$0 = \frac{\partial}{\partial x}(2\mu\epsilon_{12}) + \frac{\partial}{\partial y}(\lambda(\epsilon_{11} + \epsilon_{22}) + 2\mu\epsilon_{22}). \tag{4.6}$$

Or, in matrix form,

$$A\begin{pmatrix}\lambda_x \\ \mu_x\end{pmatrix} + B\begin{pmatrix}\lambda_y \\ \mu_y\end{pmatrix} = \mathbf{c}, \tag{4.7}$$

[1]The term '3-d ultrasound' can refer either to multiple 2-d slices through a region of interest using a linear-array transducer or to an image acquired using a 2-d matrix-array transducer. In this context the second of these definitions is intended.

where the subscripts denote partial differentiation and

$$A = \begin{pmatrix} \epsilon_{11} + \epsilon_{22} & 2\epsilon_{11} \\ 0 & 2\epsilon_{12} \end{pmatrix}, \tag{4.8}$$

$$B = \begin{pmatrix} 0 & 2\epsilon_{12} \\ \epsilon_{11} + \epsilon_{22} & 2\epsilon_{22} \end{pmatrix}, \tag{4.9}$$

$$\mathbf{c} = -\begin{pmatrix} \lambda(\epsilon_{11} + \epsilon_{22})_x + 2\mu((\epsilon_{11})_x + (\epsilon_{12})_y) \\ \lambda(\epsilon_{11} + \epsilon_{22})_y + 2\mu((\epsilon_{12})_x + (\epsilon_{22})_y) \end{pmatrix}. \tag{4.10}$$

These are the basic equations for which a solution is sought in the remainder of this chapter.

4.2 Analysis of the Inverse Problem

In this section some simple properties of the (plane strain) inverse problem are discussed. In Section 4.2.1 the governing equations are classified as hyperbolic and their characteristics are determined. The impact of the characteristics on the application of Cauchy data is discussed in Section 4.2.2. It is shown that for a square domain there is no appropriate initial curve on which Cauchy data can be specified in order to determine the Lamé constants throughout the domain. Finally, the possibility of constructing a solution using Riemann invariants is considered (and rejected) in Section 4.2.3. The results of this section provide the motivation for considering more than one displacement field in Section 4.3 and for reformulating the governing equations to make them more numerically tractable in Section 4.4.1.

4.2.1 Hyperbolicity and Characteristics

First, the inverse problem is shown to be hyperbolic and its characteristics are determined. This section uses the theory of the classification of first-order systems of PDEs introduced in Section 2.2.1.2. This is done in the form of the following lemma:

Lemma 4.1. *Given* $\mathbf{u}(\mathbf{x})$ *and* $\epsilon(\mathbf{x})$ *computed from equation (4.1), if* ϵ *is non-zero and regular everywhere inside* Ω *then:*

(i) Equation (4.7) is hyperbolic.

(ii) The characteristics of equation (4.7) are parallel to the principal axes of strain.

Proof. Equation (4.7) has characteristics given by $y'(x) = k$ where k satisfies the characteristic equation $\det(B - kA) = 0$, i.e.

$$0 = \det \begin{pmatrix} -k(\epsilon_{11} + \epsilon_{22}) & 2(\epsilon_{12} - k\epsilon_{11}) \\ \epsilon_{11} + \epsilon_{22} & 2(\epsilon_{22} - k\epsilon_{12}) \end{pmatrix}, \tag{4.11}$$

$$= \epsilon_{12}k^2 + (\epsilon_{11} - \epsilon_{22})k - \epsilon_{12}. \tag{4.12}$$

Hence, the characteristics of equation (4.7) have gradient

$$\frac{dy}{dx} = -\frac{\epsilon_{11} - \epsilon_{22} \pm \sqrt{(\epsilon_{11} - \epsilon_{22})^2 + 4\epsilon_{12}^2}}{2\epsilon_{12}}. \tag{4.13}$$

Since $(\epsilon_{11}-\epsilon_{22})^2+4\epsilon_{12}^2 > 0$ for any $\epsilon \neq 0$ there are always exactly two distinct characteristics at any point in the domain, i.e. the system is hyperbolic.

Next, it is verified that the characteristics are parallel to the principal axes of strain by direct calculation. Strain has principal values p given by $\det(\epsilon - pI) = 0$. So,

$$2p = \epsilon_{11} + \epsilon_{22} \pm \sqrt{(\epsilon_{11} - \epsilon_{22})^2 + 4\epsilon_{12}^2}. \tag{4.14}$$

We check that these eigenvalues (principal values) have corresponding eigenvectors (principal strains) parallel to the characteristics.

(a) Characteristic direction $y'(x) = -(\epsilon_{11} - \epsilon_{22} + s)/2\epsilon_{12}$, where $s = \sqrt{(\epsilon_{11} - \epsilon_{22})^2 + 4\epsilon_{12}^2}$:

$$\begin{pmatrix} \epsilon_{11} & \epsilon_{12} \\ \epsilon_{12} & \epsilon_{22} \end{pmatrix} \begin{pmatrix} 2\epsilon_{12} \\ \epsilon_{22} - \epsilon_{11} - s \end{pmatrix} = \begin{pmatrix} \epsilon_{12}(\epsilon_{11} + \epsilon_{22} - s) \\ 2\epsilon_{12}^2 + \epsilon_{22}^2 - \epsilon_{11}\epsilon_{22} - \epsilon_{22}s \end{pmatrix} \tag{4.15}$$

$$= \tfrac{1}{2}(\epsilon_{11} + \epsilon_{22} - s) \begin{pmatrix} 2\epsilon_{12} \\ \epsilon_{22} - \epsilon_{11} - s \end{pmatrix} \tag{4.16}$$

$$= p^{(-)} \begin{pmatrix} 2\epsilon_{12} \\ \epsilon_{22} - \epsilon_{11} - s \end{pmatrix}. \tag{4.17}$$

(b) Characteristic direction $y'(x) = -(\epsilon_{11} - \epsilon_{22} - s)/2\epsilon_{12}$:

$$\begin{pmatrix} \epsilon_{11} & \epsilon_{12} \\ \epsilon_{12} & \epsilon_{22} \end{pmatrix} \begin{pmatrix} 2\epsilon_{12} \\ \epsilon_{22} - \epsilon_{11} + s \end{pmatrix} = \begin{pmatrix} \epsilon_{12}(\epsilon_{11} + \epsilon_{22} + s) \\ 2\epsilon_{12}^2 + \epsilon_{22}^2 - \epsilon_{11}\epsilon_{22} + \epsilon_{22}s \end{pmatrix} \tag{4.18}$$

$$= \tfrac{1}{2}(\epsilon_{11} + \epsilon_{22} + s) \begin{pmatrix} 2\epsilon_{12} \\ \epsilon_{22} - \epsilon_{11} + s \end{pmatrix} \tag{4.19}$$

$$= p^{(+)} \begin{pmatrix} 2\epsilon_{12} \\ \epsilon_{22} - \epsilon_{11} + s \end{pmatrix}. \tag{4.20}$$

This verification completes the proof. □

Since equation (4.7) is hyperbolic, the theory of Sections 2.2.1 and 2.2.2 shows that it is possible to formulate appropriate boundary conditions to make the inverse problem well-posed in the form of Cauchy data on some given curve Γ in Ω — i.e. specifying the values of λ and μ on Γ. In order for the solution to be uniquely determined throughout Ω it is necessary for each of the characteristics through any given point in Ω to intersect with Γ exactly once. Note that this means that a characteristic is excluded from being a candidate for Γ (see Sections 2.2.1 and 2.2.2 or, e.g. Ockendon *et al.* [53] for more details on Cauchy data for hyperbolic equations).

4.2.2 The Characteristics and Cauchy Data

In order to determine an appropriate curve Γ for the specification of Cauchy data it is necessary to identify the characteristics of equation (4.7). An approximation to the type of deformation seen in elasticity imaging (when the ultrasound probe is pressed into the skin) is to consider a unit square with $u = 0$ on $y = 0, 1$, $v = 0$ on $y = 0$, $v = -v_0$ on $y = 1$ for some v_0 and zero stress on $x = 0, 1$.

The characteristics resulting from a deformation of this form are shown in Figure 4.1. These are calculated numerically by first determining the strain field across the domain by a finite element solution of the linear elasticity forward problem (as described in Section 2.4.2.2) and then calculating the characteristic gradients using equation (4.13). Any curve that is nowhere parallel to a characteristic through the same point is a candidate for Γ. The obvious candidate is the line $y = 1$ (or, equivalently $y = 0$). This is the line that is in contact with the ultrasound probe in elasticity imaging, and possibly the only part of

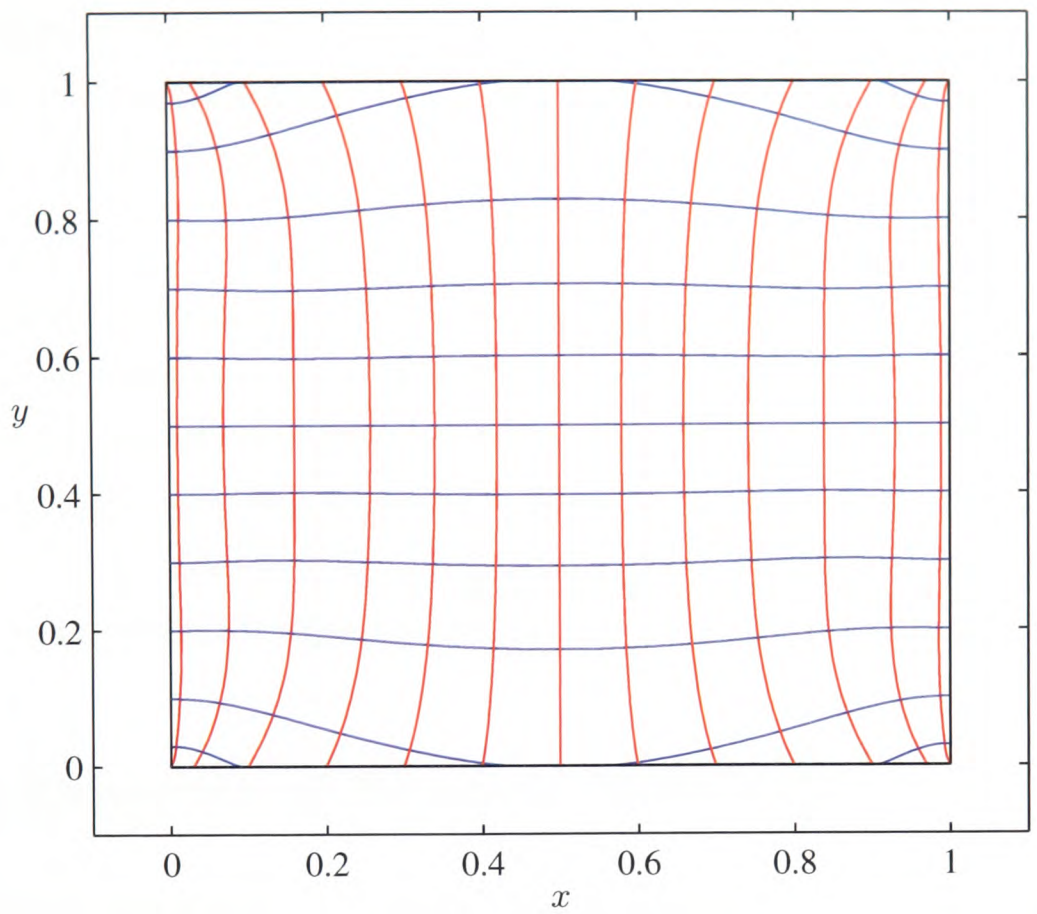


Figure 4.1: Characteristic curves for the linear elasticity inverse problem on a unit square with a numerically generated displacement field corresponding to zero stress boundary conditions on the boundaries $x = 0, 1$, $u = 0$ on $y = 0, 1$, $v = 0$ on $y = 0$ and $v = -0.1$ on $y = 1$. The Lamé constants are assumed to be constant throughout the domain, with $\lambda = 1$, $\mu = 2/3$. One family of characteristics is shown in red, the other in blue.

the image for which there is any realistic hope of knowing the stiffness before the imaging process begins. However, $y = 1$ is parallel to a characteristic at $x = 0.5$. Further, the domain of definition (the set of all points for which both characteristics through the point intersect with the curve on which Cauchy data is given) of a solution with Cauchy data given on $y = 1$ is very small — extending only just past $y = 0.9$ at its furthest extreme. The situation is even worse on the edges $x = 0, 1$, which are almost parallel to a characteristic for their whole length and would give rise to a solution with a domain of definition reaching only a very small distance into the unit square.

In fact, there is no curve Γ which would give rise to a solution defined throughout the region of interest for the deformation shown in Figure 4.1. For example, in order to have a domain of definition covering all four corners of the square any appropriate curve would have to pass through each of them (since one characteristic through each of the corners

does not intersect with any other point in the square). However, there is a characteristic that passes through the two corners with $x = 0$ (and another passing through the two with $x = 1$). Since Γ cannot intersect the same characteristic twice it cannot pass through more than two corners. Hence, there is no curve Γ with domain of definition covering the unit square.

Although this analysis is specific to the deformation given in Figure 4.1 it highlights the difficulty in even determining what appropriate boundary conditions should be given for the linear elasticity inverse problem, let alone solving the problem subject to the data. Using multiple displacement fields instead of boundary conditions is discussed in Section 4.3, whilst in Section 4.4 the equation is reformulated as an elliptic problem allowing simple boundary data to be given all around the edge of the region of interest.

4.2.3 Non-Existence of Riemann Invariants

In Section 2.2.3 the possibility of constructing a solution to a system of hyperbolic equations using Riemann invariants (quantities that are constant along a characteristic) was discussed. It has been shown that equation (4.7) is hyperbolic so, in this section, the possibility of it having Riemann invariants is investigated. If they did exist then the solution of equation (4.7) would become straightforward (since the Cauchy data would specify the values of the Riemann invariants on each characteristic, leaving an algebraic equation to solve for λ and μ at each point in Ω). With this in mind, search for left eigenvectors of $B - kA$, i.e. vectors $\mathbf{l}$ such that

$$\mathbf{l}^T(B - kA) = \mathbf{0}^T. \tag{4.21}$$

$$\Rightarrow \begin{pmatrix} l_1 & l_2 \end{pmatrix} \begin{pmatrix} -k(\epsilon_{11} + \epsilon_{22}) & 2(\epsilon_{12} - k\epsilon_{11}) \\ \epsilon_{11} + \epsilon_{22} & 2(\epsilon_{22} - k\epsilon_{12}) \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}. \tag{4.22}$$

Writing this matrix equation as a system of algebraic equations for l_1, l_2 in terms of k it can be seen that

$$0 = l_2 - kl_1, \tag{4.23}$$

$$0 = l_1(\epsilon_{12} - k\epsilon_{11}) + l_2(\epsilon_{22} - k\epsilon_{12}). \tag{4.24}$$

Hence,

$$\mathbf{l} = (1, k)^T. \tag{4.25}$$

Observe that equation (4.24) is satisfied if $\mathbf{l}$ is given by equation (4.25) as it reduces to

$$0 = -\epsilon_{12}k^2 - (\epsilon_{11} - \epsilon_{22})k + \epsilon_{12}, \tag{4.26}$$

which is the characteristic equation and so is automatically satisfied by k. Note also that equation (4.25) defines *two* families of eigenvectors since there are two values of k.

Then recall from Section 2.2.3 that the Riemann invariants (if there are any) are given by integrating

$$\mathbf{l}^T A\dot{\boldsymbol{\lambda}} = \mathbf{l}^T \mathbf{c}\dot{x}. \tag{4.27}$$

$$\text{i.e.} \quad \begin{pmatrix} 1 & k \end{pmatrix} \begin{pmatrix} \epsilon_{11}+\epsilon_{22} & 2\epsilon_{11} \\ 0 & 2\epsilon_{12} \end{pmatrix} \begin{pmatrix} \dot{\lambda} \\ \dot{\mu} \end{pmatrix} = -\begin{pmatrix} 1 & k \end{pmatrix} \begin{pmatrix} \lambda(\epsilon_{11}+\epsilon_{22})_x + 2\mu((\epsilon_{11})_x + (\epsilon_{12})_y) \\ \lambda(\epsilon_{11}+\epsilon_{22})_y + 2\mu((\epsilon_{12})_x + (\epsilon_{22})_y) \end{pmatrix} \dot{x}, \tag{4.28}$$

$$\begin{aligned}(\epsilon_{11}+\epsilon_{22})\dot{\lambda} + 2(\epsilon_{11}+k\epsilon_{12})\dot{\mu} = &-\lambda((\epsilon_{11}+\epsilon_{22})_x + k(\epsilon_{11}+\epsilon_{22})_y)\dot{x} \\ &- 2\mu((\epsilon_{11})_x + (\epsilon_{12})_y + k((\epsilon_{12})_x + (\epsilon_{22})_y))\dot{x},\end{aligned} \tag{4.29}$$

where, as in Section 2.2.3, the dots denote differentiation with respect to t, the parameter along a characteristic. Observe that

$$\frac{d}{dt}(\lambda(\epsilon_{11}+\epsilon_{22})) = \dot{\lambda}(\epsilon_{11}+\epsilon_{22}) + \lambda((\epsilon_{11}+\epsilon_{22})_x\dot{x} + (\epsilon_{11}+\epsilon_{22})_y\dot{y}), \tag{4.30}$$

$$= \dot{\lambda}(\epsilon_{11}+\epsilon_{22}) + \lambda((\epsilon_{11}+\epsilon_{22})_x + k(\epsilon_{11}+\epsilon_{22})_y)\dot{x}, \tag{4.31}$$

since $y'(x) = k$. The rest of the terms are multiplied by either μ or $\dot{\mu}$, so they can only be integrated exactly if they can be written as the derivative of a product, i.e. it is required that

$$(\epsilon_{11}+k\epsilon_{12})\dot{\mu} + ((\epsilon_{11})_x + (\epsilon_{12})_y + k((\epsilon_{12})_x + (\epsilon_{22})_y))\mu\dot{x} = \frac{d}{dt}((\epsilon_{11}+k\epsilon_{12})\mu). \tag{4.32}$$

The first term on the left-hand-side is clearly of the correct form so this can be reduced to

determining whether or not the following relationship holds for general ϵ:

$$\frac{d}{dt}(\epsilon_{11} + k\epsilon_{12}) = ((\epsilon_{11})_x + (\epsilon_{12})_y + k((\epsilon_{12})_x + (\epsilon_{22})_y))\dot{x}. \tag{4.33}$$

It can be shown that this is not the case by deducing a contradiction. Suppose that equation (4.33) holds. Suppose further that $u(x, y) = 0$, $v(x, y) = xy$ so that

$$\epsilon(x, y) = \begin{pmatrix} 0 & \frac{1}{2}y \\ \frac{1}{2}y & x \end{pmatrix}, \tag{4.34}$$

and

$$k = \frac{x \pm \sqrt{x^2 + y^2}}{y}. \tag{4.35}$$

Then the right-hand-side of equation (4.33) is equal to $\frac{1}{2}\dot{x}$ since all of the derivatives are zero except for $(\epsilon_{12})_y$, which is equal to $\frac{1}{2}$. The left-hand-side can be calculated as follow:

$$\frac{d}{dt}(\epsilon_{11} + k\epsilon_{12}) = \frac{1}{2}\frac{d}{dt}(ky), \tag{4.36}$$

$$= \frac{1}{2}\frac{d}{dt}(x \pm \sqrt{x^2 + y^2}), \tag{4.37}$$

$$= \frac{1}{2}\frac{\partial}{\partial x}(x \pm \sqrt{x^2 + y^2})\dot{x} + \frac{1}{2}\frac{\partial}{\partial y}(x \pm \sqrt{x^2 + y^2})k\dot{x}, \tag{4.38}$$

since $y'(x) = k$,

$$= \frac{1}{2}\left(1 \pm \frac{x}{\sqrt{x^2 + y^2}} \pm \frac{ky}{\sqrt{x^2 + y^2}}\right)\dot{x}, \tag{4.39}$$

$$\frac{d}{dt}(\epsilon_{11} + k\epsilon_{12}) = \dot{x}\left(1 \pm \frac{x}{\sqrt{x^2 + y^2}}\right) \neq \frac{1}{2}\dot{x}. \tag{4.40}$$

This contradicts the assumption that equation (4.33) is valid and hence there are no Riemann invariants for the 2-d plane strain elasticity inverse problem. So, it will be necessary to formulate a numerical method to solve the system of PDEs for a general strain field.

4.3 Multiple Displacement Fields

In the previous section it has been demonstrated that it will be difficult to give appropriate boundary data for any numerical scheme for the solution of the linear elasticity inverse problem. Further, it will, in general, be difficult to determine the values of the Lamé constants on the boundary (how can *a priori* knowledge of the material stiffnesses be obtained from a pair of ultrasound images?). The combination of these two factors makes it desirable to find an alternative to the specification of boundary conditions that leads to a well-posed problem.

Barbone and Gokhale [6] use multiple displacement fields to overcome the lack of *a priori* knowledge of the boundary stiffnesses in the incompressible case. In this section it is shown that this analysis can be extended to the compressible case. This is broken down into four parts. First, λ is eliminated from the inverse problem to give a second order PDE for μ. Then, in Section 4.3.1, it is shown that if two displacement fields are known then Cauchy data specified on any appropriate initial curve with non-zero length (even one that is infinitesimally short) will determine μ throughout the region of interest. It is also shown that the Cauchy data must satisfy a (known) fourth-order ODE on the initial curve. Taken together these two results show that four independent pieces of data (enough to solve the ODE uniquely) are enough to solve the inverse problem (for μ) throughout the region of interest.

The benefits of using more than two displacement fields are considered in Section 4.3.2. The possibility of using three fields is discussed, but the analysis is inconclusive as to how much data is required to uniquely determine μ. In the case of four displacement fields it is shown that μ is determined uniquely up to a multiplicative constant. Further, it is demonstrated that it is not possible to generate more than four independent displacement fields, so no further progress can be made. Finally, in Section 4.3.3 it is demonstrated that if μ is known throughout the region of interest then λ can be calculated uniquely up to an additive constant.

Before beginning the analysis, it is necessary to introduce the following definition for the linear elasticity inverse problem when more than one displacement field is given:

Definition 4.3 (The Multiple Displacement Field Inverse Problem)**.** Given N displacement fields $\mathbf{u}^{(n)}(\mathbf{x}, t)$ and the density ρ defined on $\Omega \times (0, T > 0)$ find *a single* $\lambda(\mathbf{x})$ and *a single* $\mu(\mathbf{x})$ such that equation (4.4) is satisfied for $n = 1, \ldots, N$.

Also, in order to simplify the analysis of the inverse problem the two equations of (4.4) are cross-differentiated and λ is eliminated between them. This gives the following equation for μ:

$$\left(\frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial x^2}\right)(\mu\epsilon_{12}) + \frac{\partial^2}{\partial x \partial y}(\mu\epsilon_{11} - \mu\epsilon_{22}) = \frac{1}{2}\rho\frac{\partial^2}{\partial t^2}\left(\frac{\partial u}{\partial y} - \frac{\partial v}{\partial x}\right). \tag{4.41}$$

Observe that in the inverse problem the right-hand-side of equation (4.41) is known (and is zero in the quasi-static case that we are generally interested in). In Section 4.3.3 it is shown that if equation (4.41) can be solved for μ then λ is determined uniquely up to an additive constant.

4.3.1 Two Displacement Fields

Suppose first that two distinct displacement fields are given (i.e. $N = 2$). For the incompressible case Barbone and Gokhale [6] show that the solution space is (at most) 4-dimensional. It is shown here that this remains the case if the material is now allowed to be compressible.

First, a constraint on the two displacement fields is introduced. This effectively ensures that the same material is being deformed to generate both displacement fields.

Definition 4.4. Two displacement fields $\mathbf{u}^{(1)}$, $\mathbf{u}^{(2)}$ (and their corresponding strain fields $\epsilon^{(1)}$, $\epsilon^{(2)}$) are said to be *compatible* with each other if they can occur in the same material, i.e. if there exists functions λ, μ satisfying equation (4.4) for both $\epsilon^{(1)}$ and $\epsilon^{(2)}$.

Next, it is shown that given two compatible displacement fields and Cauchy data along some initial curve a solution can be constructed throughout the region of interest.

Lemma 4.2. *Suppose that two linearly independent, compatible displacement fields $\epsilon^{(1)}$, $\epsilon^{(2)}$ that are nonzero everywhere in Ω are given. Suppose further that $\bar{\mu}(s) = \mu(x(s), y(s))$ and its normal derivative $\bar{\mu}_n(s) = \frac{\partial \mu}{\partial n}(x(s), y(s))$ are given along some curve $C = (x(s), y(s))$ (i.e. given Cauchy data for the hyperbolic PDE). Then, assuming that the solution $\mu(x, y)$ exists (see lemma 4.3 for a constraint on the Cauchy data necessary for a solution to exist):*

(i) If the eigendirections of $\epsilon^{(1)}$ are distinct from those of $\epsilon^{(2)}$ except at isolated points then μ is uniquely determined throughout Ω.

(ii) If there exists a set of curves B_j, $j = 1, \ldots, Q$ such that each B_j is a characteristic of both $\epsilon^{(1)}$ and $\epsilon^{(2)}$ and C intersects none of the B_j then μ is uniquely determined in the smallest region bounded by $\bigcup(\partial\Omega, B_j)$ which contains C.

Proof. Rewrite equation (4.41) for each displacement field as:

$$\mathcal{L}(\mathbf{u}^{(i)})\mu = f(\mathbf{u}^{(i)}), \quad i = 1, 2. \tag{4.42}$$

Suppose, without loss of generality, that the x- and y-axes are aligned with the characteristics of equation (4.42) with $i = 1$. Then the Cauchy data on C and equation (4.42) are sufficient to determine μ in the rectangular region D_1 bounded by the characteristics of this first equation that pass through the endpoints of C (see Figure 4.2) [53]. (If C is a characteristic of equation (4.42) with $i = 1$ then interchange the roles of $\mathbf{u}^{(1)}$, $\mathbf{u}^{(2)}$.) The boundary of D_1 can then be treated as the curve on which Cauchy data is given for equation (4.42) with $i = 2$, determining μ uniquely in the region D_2 bounded by the characteristics of this second equation that enclose D_1 (again, see Figure 4.2).

By this alternating solution of equation (4.42) with $i = 1, 2$, μ can be determined in increasingly large regions of Ω. It can also be seen that this growth is dependent on the characteristics of the two equations being non-parallel — a curve which is a characteristic of equation (4.42) with $i = 1$ and $i = 2$ limits the expansion of one side of D_i as $i \to \infty$. □

Observe that lemma 4.2 does not impose any limit on the size of the initial curve C — it could be infinitesimal. This suggests that very few data are needed to determine μ. In the next lemma it is made clear precisely what data is required.

Lemma 4.3. *For a solution to equation (4.42) to exist the Cauchy data specified on C must satisfy a given fourth order ODE.*

Proof. Without loss of generality, suppose that the coordinate system is chosen so that the x-axis is aligned with C (which may be a characteristic of at most one of the cases of equation (4.42)). Then the Cauchy data can be written as

$$\mu(x, 0) = \bar{\mu}(x), \tag{4.43}$$

$$\frac{\partial \mu}{\partial y}(x, 0) = \bar{\mu}_n(x). \tag{4.44}$$

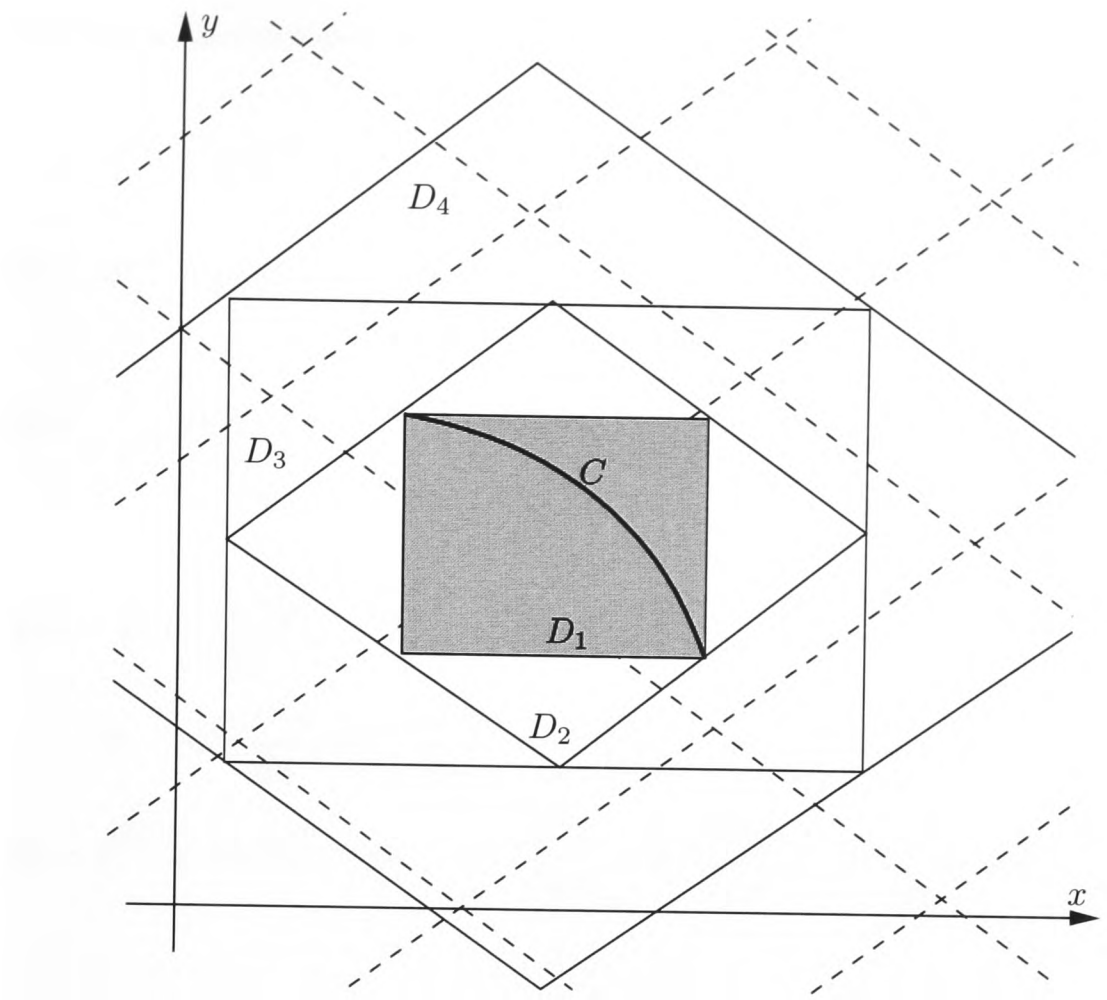


Figure 4.2: From limited initial data to determination of μ throughout the region of interest. Given Cauchy data on C, equation (4.42) with $i = 1$ (characteristics parallel to the x- and y-axes) determines the solution in the (shaded) region D_1. This solution can then be used to provide Cauchy data to solve equation (4.42) with $i = 2$ (characteristics are the dashed lines) in order to determine μ in D_2. The process is then repeated to generate the solution in regions D_3, D_4, ...

Suppose equation (4.42) is rewritten as

$$\mathcal{L}(\mathbf{u}^{(1)})\mu \equiv a^{(1)}\left(\frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial x^2}\right)\mu + b^{(1)}\frac{\partial^2 \mu}{\partial x \partial y} + c^{(1)}\frac{\partial \mu}{\partial x} + d^{(1)}\frac{\partial \mu}{\partial y} + e^{(1)}\mu = f^{(1)}, \qquad (4.45)$$

$$\mathcal{L}(\mathbf{u}^{(2)})\mu \equiv a^{(2)}\left(\frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial x^2}\right)\mu + b^{(2)}\frac{\partial^2 \mu}{\partial x \partial y} + c^{(2)}\frac{\partial \mu}{\partial x} + d^{(2)}\frac{\partial \mu}{\partial y} + e^{(2)}\mu = f^{(2)}, \qquad (4.46)$$

where the coefficients $a^{(j)}, \ldots, f^{(j)}$, $j = 1, 2$ are known in terms of the displacement components and their derivatives. Any function μ that is a solution of the inverse problem must satisfy both of these equations.

Now eliminate the y-derivatives between the two equations to obtain an ODE in x:

$$a^{(1)} \times (4.46) - a^{(2)} \times (4.45) \quad \Rightarrow \quad b^{(3)}\frac{\partial \mu}{\partial x \partial y} + c^{(3)}\frac{\partial \mu}{\partial x} + d^{(3)}\frac{\partial \mu}{\partial y} + e^{(3)}\mu = f^{(3)}. \tag{4.47}$$

$$b^{(3)} \times (4.45) - b^{(1)} \times (4.47) \quad \Rightarrow \quad a^{(4)}\left(\frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial x^2}\right)\mu + c^{(4)}\frac{\partial \mu}{\partial x} + d^{(4)}\frac{\partial \mu}{\partial y} + e^{(4)}\mu = f^{(4)}. \tag{4.48}$$

$$b^{(3)}\frac{\partial (4.48)}{\partial x} - a^{(4)}\frac{\partial (4.47)}{\partial y} \quad \Rightarrow \quad a^{(5)}\frac{\partial^3 \mu}{\partial x^3} + b^{(5)}\frac{\partial^2 \mu}{\partial x^2} + c^{(5)}\frac{\partial \mu}{\partial x} + d^{(5)}\frac{\partial \mu}{\partial y} + e^{(5)}\mu = f^{(5)}. \tag{4.49}$$

$$d^{(5)} \times (4.47) - b^{(3)}\frac{\partial (4.49)}{\partial x} \quad \Rightarrow \quad a^{(6)}\frac{\partial^4 \mu}{\partial x^4} + b^{(6)}\frac{\partial^3 \mu}{\partial x^3} + c^{(6)}\frac{\partial^2 \mu}{\partial x^2} + d^{(6)}\frac{\partial \mu}{\partial x} + e^{(6)}\frac{\partial \mu}{\partial y} + f^{(6)}\mu = g^{(6)}. \tag{4.50}$$

$$e^{(6)} \times (4.49) - d^{(5)} \times (4.50) \quad \Rightarrow \quad a^{(7)}\frac{\partial^4 \mu}{\partial x^4} + b^{(7)}\frac{\partial^3 \mu}{\partial x^3} + c^{(7)}\frac{\partial^2 \mu}{\partial x^2} + d^{(7)}\frac{\partial \mu}{\partial x} + e^{(7)}\mu = f^{(7)}. \tag{4.51}$$

Note that in simplifying equations (4.49) and (4.50), equations (4.47) and (4.48) were used to eliminate occurrences of μ_{xy} and μ_{yy} respectively. Observe also that in the quasi-static case all of the right-hand-sides in the above equations are zero (since $f^{(1)}$ and $f^{(2)}$ are zero for a quasi-static deformation).

Equation (4.51) must be satisfied throughout Ω by any $\mu(x, y)$ that solves both equations (4.45) and (4.46) in that domain. In particular, it must be satisfied by the Cauchy data given on $y = 0$. Hence, $\bar{\mu}(x)$ must be a solution of equation (4.51). Further, equation (4.50) shows that $\bar{\mu}_n(x)$ is determined entirely in terms of $\bar{\mu}(x)$. □

Lemma 4.3 shows that the Cauchy data imposed on our pair of equations (4.42) must satisfy strict constraints — in particular we do not have any freedom at all to choose $\bar{\mu}_n(x)$. This is a marked contrast to the single equation case, for which any Cauchy data may be imposed. This leads to the following theorem about the two displacement field problem:

Theorem 4.1. *Given two linearly independent, compatible displacement fields* $\mathbf{u}^{(1)}$, $\mathbf{u}^{(2)}$ *that are everywhere nonzero and such that their eigendirections are distinct except at isolated points, let* $M^{(i)}$, $i = 1, 2$, *be the set of all functions* μ *such that*

$$\mathcal{L}(\mathbf{u}^{(i)})\mu = 0, \quad i = 1, 2. \tag{4.52}$$

Then,

$$\text{Dim}\left\{M^{(1)} \cap M^{(2)}\right\} \leqslant 4. \tag{4.53}$$

In equation (4.53), Dim *represents the dimension of the set.*

Proof. Let $\bar{\mu}(x)$ be a set of Cauchy data, prescribed along the x-axis, for one of the two equations. Then lemma 4.3 tells us that $\bar{\mu}$ satisfies a 4th order ODE. Hence $\bar{\mu}$ can contain at most four arbitrary constants. By the first part of lemma 4.2, $\bar{\mu}$ determines $\mu(\mathbf{x})$ throughout Ω. So, $\mu(\mathbf{x})$ contains at most four arbitrary constant. □

4.3.2 More Than Two Displacement Fields

In the previous section it was shown that using two displacement fields greatly reduced the amount of *a priori* knowledge required to reconstruct $\lambda(\mathbf{x})$, $\mu(\mathbf{x})$: rather than requiring complete boundary data only four pieces of information about μ are required, i.e. its values at four distinct points in the region of interest. Next, the possibility of using more than two displacement fields to reduce further the necessary amount of *a priori* knowledge of μ is considered.

4.3.2.1 Three Displacement Fields

Suppose first that three displacement fields are available. Then equation (4.41) can be applied to each of the displacement fields so that

$$\mathcal{L}(\mathbf{u}^{(i)})\mu \equiv a^{(i)}\left(\frac{\partial^2}{\partial y^2}-\frac{\partial^2}{\partial x^2}\right)\mu+b^{(i)}\frac{\partial^2\mu}{\partial x\partial y}+c^{(i)}\frac{\partial\mu}{\partial x}+d^{(i)}\frac{\partial\mu}{\partial y}+e^{(i)}\mu = f^{(i)}, \quad i=1,2,3. \tag{4.54}$$

Clearly Lemma 4.2 can be applied to any two of the three displacement fields to show that the solution can be constructed across the whole domain if appropriate Cauchy data is specified on any appropriate curve C (even if it is infinitesimally short). Recall that in the previous section the compatibility of the two displacement fields was used to place restrictions on the form of the Cauchy data (it was shown to satisfy a fourth-order ODE) and reduce the amount of data required to determine the solution. In this section the idea is to use the compatibility of the three given displacement fields to restrict even further the choice of Cauchy data. If it could be shown to satisfy an algebraic equation then the solution would be determined uniquely throughout the domain. Unfortunately, this turns

out not to be the case for reasons that will be discussed below. Indeed, it is not clear exactly how many pieces of data are required if three displacement fields are given.

Suppose that Lemma 4.3 is applied to the first two displacement fields as in Section 4.3.1. Then on the curve C the following relationship holds:

$$a^{(4)}\frac{\partial^4\mu}{\partial x^4} + b^{(4)}\frac{\partial^3\mu}{\partial x^3} + c^{(4)}\frac{\partial^2\mu}{\partial x^2} + d^{(4)}\frac{\partial\mu}{\partial x} + e^{(4)}\mu = f^{(4)}, \tag{4.55}$$

where the coefficients depend on the two displacement fields and in the quasi-static case $f^{(4)} = 0$. Now, applying Lemma 4.3 to the first and third displacement fields gives

$$a^{(5)}\frac{\partial^4\mu}{\partial x^4} + b^{(5)}\frac{\partial^3\mu}{\partial x^3} + c^{(5)}\frac{\partial^2\mu}{\partial x^2} + d^{(5)}\frac{\partial\mu}{\partial x} + e^{(5)}\mu = f^{(5)}, \tag{4.56}$$

on C. Applying the lemma to the second and third displacement gives

$$a^{(6)}\frac{\partial^4\mu}{\partial x^4} + b^{(6)}\frac{\partial^3\mu}{\partial x^3} + c^{(6)}\frac{\partial^2\mu}{\partial x^2} + d^{(6)}\frac{\partial\mu}{\partial x} + e^{(6)}\mu = f^{(6)}, \tag{4.57}$$

Then, the fourth derivative of μ can be eliminated between equations (4.55) and (4.56):

$$a^{(5)} \times (4.55) - a^{(4)} \times (4.56) \quad \Rightarrow \quad b^{(7)}\frac{\partial^3\mu}{\partial x^3} + c^{(7)}\frac{\partial^2\mu}{\partial x^2} + d^{(7)}\frac{\partial\mu}{\partial x} + e^{(7)}\mu = f^{(7)}. \tag{4.58}$$

So, the Cauchy data must satisfy a third order ODE on C. Hence, the solution of the inverse problem with three displacement fields requires at most three pieces of initial data.

However, it appears that it may be possible to reduce the number of data points further since the fourth x-derivative of μ could be eliminated between *any two* of equations (4.55) – (4.57). This gives two further third order PDEs for μ:

$$b^{(i)}\frac{\partial^3\mu}{\partial x^3} + c^{(i)}\frac{\partial^2\mu}{\partial x^2} + d^{(i)}\frac{\partial\mu}{\partial x} + e^{(i)}\mu = f^{(i)}, \quad i = 8, 9. \tag{4.59}$$

Observe that only each of the equations (4.55) – (4.57) have been used *twice* in deriving equations (4.58) and (4.59), so it is not clear that they should be linearly independent. Suppose initially that they are. Then the process of eliminating the highest derivatives can be repeated to produce three second-order ODEs for μ, then three first-order equations and finally three algebraic equations of the form

$$e^{(i)}\mu = f^{(i)}, \quad i = 10, 11, 12. \tag{4.60}$$

At each step it is assumed that the process of eliminating the highest derivative does not break the linear independence of the system. Observe, however, that three algebraic equations for one variable *cannot* be linearly independent. So, the linear independence must be contradicted at some point in the analysis. However, it is not clear at exactly which stage this occurs. If it occurs when the third derivatives are eliminated than three pieces of data will be required. On the other hand, if it occurs when the first derivatives are eliminated then only one piece of data is required. So, all that can be said for the case $N = 3$ is that the number of pieces of boundary data that are required is between one and three, i.e. giving three displacement fields is better than giving two but it is not clear exactly how much better.

4.3.2.2 Four Displacement Fields

If $N = 4$ then the analysis is much simpler — it turns out that for the quasi-static case a unique (up to a multiplicative constant) solution for μ can be written down in terms of integrals of known quantities. This is demonstrated in the following theorem:

Theorem 4.2. *Given four linearly independent, mutually compatible equilibrium displacement fields* $\mathbf{u}^{(i)}$, $i = 1, 2, 3, 4$, *then the solution* $\mu(x, y)$ *of the set of equations*

$$\mathcal{L}(\mathbf{u}^{(i)})\mu = 0, \quad i = 1, 2, 3, 4, \tag{4.61}$$

is unique up to a multiplicative constant. Further, the solution is given by

$$\mu(x, y) = \mu(x_0, y_0) \exp\left[\frac{1}{2}\int_{y_0}^{y}\int_{x_0}^{x}\left(\frac{\partial R}{\partial x'} + \frac{\partial Q}{\partial y'}\right)dx'dy' - \int_{y_0}^{y} R dy' - \int_{x_0}^{x} Q dx'\right], \tag{4.62}$$

for some given functions $Q(x, y)$ *and* $R(x, y)$.

Proof. Start by rewriting equations (4.61) as

$$a^{(i)}\left(\frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial x^2}\right)\mu + b^{(i)}\frac{\partial^2 \mu}{\partial x \partial y} + c^{(i)}\frac{\partial \mu}{\partial x} + d^{(i)}\frac{\partial \mu}{\partial y} + e^{(i)}\mu = 0, \quad i = 1, 2, 3, 4, \tag{4.63}$$

where the ordering is chosen so that $a^{(1)}b^{(2)} - a^{(2)}b^{(1)} \neq 0$. Then equation (4.63) can be solved with $i = 1$ and $i = 2$ to obtain the values of $\frac{\partial^2 \mu}{\partial y^2} - \frac{\partial^2 \mu}{\partial x^2}$ and $\frac{\partial^2 \mu}{\partial x \partial y}$. Substituting these

results into the cases $i = 3, 4$ and rearranging leads to the two equations

$$\frac{\partial \mu}{\partial x} + Q(x, y)\mu = 0, \tag{4.64}$$

$$\frac{\partial \mu}{\partial y} + R(x, y)\mu = 0. \tag{4.65}$$

A solvability condition on equations (4.64), (4.65) is

$$\frac{\partial Q}{\partial y} = \frac{\partial R}{\partial x}. \tag{4.66}$$

This is assumed to be satisfied since the displacement fields are compatible (i.e. at least one μ satisfying the equations *does* exist). Then the joint solution of equations (4.64) and (4.65) is equation (4.62). This can be seen by differentiating equation (4.62) using the Leibniz integral rule. For example, differentiating with respect to x gives:

$$\frac{\partial \mu}{\partial x} = \mu \frac{\partial}{\partial x}\left[\frac{1}{2}\int_{y_0}^{y}\int_{x_0}^{x}\left(\frac{\partial R}{\partial x'} + \frac{\partial Q}{\partial y'}\right)dx'dy' - \int_{y_0}^{y} R dy' - \int_{x_0}^{x} Q dx'\right], \tag{4.67}$$

$$= \left[\frac{1}{2}\int_{y_0}^{y}\left(\frac{\partial R}{\partial x} + \frac{\partial Q}{\partial y'}\right)dy' - \int_{y_0}^{y}\frac{\partial R}{\partial x}dy' - Q\right]\mu, \tag{4.68}$$

$$= \left[\int_{y_0}^{y}\frac{\partial Q}{\partial y'}dy' - \int_{y_0}^{y}\frac{\partial Q}{\partial y'}dy' - Q\right]\mu, \tag{4.69}$$

on applying the solvability condition given in equation (4.66).

$$= (Q - Q - Q)\mu, \tag{4.70}$$

$$= -Q\mu, \tag{4.71}$$

which is equivalent to equation (4.64). A similar result holds for the y-derivative. □

4.3.2.3 More Than Four Displacement Fields

Although it appears that it may be possible to determine the multiplicative constant by the inclusion of a fifth displacement field this does not turn out to be possible. The following lemma shows that a fifth linearly independent displacement field cannot exist.

Lemma 4.4. *The maximum number of linearly independent, mutually compatible equilib-*

rium displacement fields that can exist in an isotropic, linearly elastic material undergoing plane strain in the (x, y)-plane is four.

Proof. To the set of equations (4.63), add

$$\mathcal{L}(\mathbf{u}^{(5)})\mu \equiv a^{(5)}\left(\frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial x^2}\right)\mu + b^{(5)}\frac{\partial^2 \mu}{\partial x \partial y} + c^{(5)}\frac{\partial \mu}{\partial x} + d^{(5)}\frac{\partial \mu}{\partial y} + e^{(5)}\mu = 0. \tag{4.72}$$

Then these form a homogeneous set of linear equations for the five unknowns $\mu_{yy} - \mu_{xx}$, μ_{xy}, μ_x, μ_y and μ. Since the displacement fields are assumed to be mutually compatible a nontrivial solution exists. Hence the coefficient matrix formed by the displacement fields must be singular. This contradicts the assumed linear independence of the displacement fields. Hence, any five mutually compatible displacement fields must be linearly dependent. □

4.3.3 Determination of λ When μ is Known

Solving equation (4.41) will give the value of μ everywhere in the region of interest. However, once this is done it is still necessary to reconstruct λ. Here it is shown that this can always be done, subject to one piece of data.

Theorem 4.3. *Suppose that the value of $\mu(x, y)$ and the displacement field $\mathbf{u}(x, y)$ are known everywhere in Ω. Then $\lambda(x, y)\epsilon_{kk}(x, y)$ is determined uniquely up to an additive constant. Further, $\lambda(x, y)$ is given by*

$$\lambda(x,y) = \frac{1}{\epsilon_{kk}(x,y)}\left[\lambda(x_0,y_0)\epsilon_{kk}(x_0,y_0) + \frac{1}{2}\int_{y_0}^{y}\int_{x_0}^{x}\left(\frac{\partial g}{\partial x'} + \frac{\partial f}{\partial y'}\right)dx'dy' - \int_{x_0}^{x} f\,dx' - \int_{y_0}^{y} g\,dy'\right], \tag{4.73}$$

for some given functions $f(x, y)$, $g(x, y)$.

Proof. If μ is known everywhere in Ω then equations (4.4) can be simplified to

$$\frac{\partial}{\partial x}(\lambda\epsilon_{kk}) + f(x,y) = 0, \tag{4.74}$$

$$\frac{\partial}{\partial y}(\lambda\epsilon_{kk}) + g(x,y) = 0, \tag{4.75}$$

where $f(x, y)$, $g(x, y)$ depend on both μ and $\mathbf{u}$. Cross-differentiating gives the solvability

condition

$$\frac{\partial g}{\partial x} = \frac{\partial f}{\partial y}. \tag{4.76}$$

Since the displacement field is generated by the deformation of a linear elastic material, assume that this relation is automatically satisfied. Then the joint solution of equations (4.74) and (4.75) is given by equation (4.73). This can be verified by differentiation with respect to x and y applying the Leibniz integral rule as in the proof of Theorem 4.2. Observe that as an equation for $\lambda(x,y)\epsilon_{kk}(x,y)$ this is completely known apart from the constant additive term $\lambda(x_0,y_0)\epsilon_{kk}(x_0,y_0)$ (in fact, $\epsilon_{kk}(x_0,y_0)$ is known for a given (x_0,y_0)).

In order to show that the solution is unique (up to an additive constant) consider any two solutions λ_1, λ_2 to equations (4.74), (4.75). Then their difference satisfies the homogeneous system

$$\frac{\partial}{\partial x}[(\lambda_2 - \lambda_1)\epsilon_{kk}] = 0 \quad \Rightarrow \quad (\lambda_2 - \lambda_1)\epsilon_{kk} = p(y), \tag{4.77}$$

$$\frac{\partial}{\partial y}[(\lambda_2 - \lambda_1)\epsilon_{kk}] = 0 \quad \Rightarrow \quad (\lambda_2 - \lambda_1)\epsilon_{kk} = q(x). \tag{4.78}$$

Hence, $(\lambda_2 - \lambda_1)\epsilon_{kk}$ = constant, i.e. any two solutions of equations (4.74), (4.75), when multiplied by $\epsilon_{kk}(x,y)$ differ at most by an additive constant. □

4.3.4 Summary

In this section the results of Sections 4.3.1 – 4.3.3 are summarized. The most important observation of these sections was that the need to specify boundary data can be circumvented by considering more than one deformation of the same material. In Section 4.3.1 it was shown that only four pieces of data are required to determine μ uniquely. If four displacement fields are available then Section 4.3.2.2 demonstrates that μ can be determined up to a multiplicative constant, i.e. one piece of boundary data is required. In Section 4.3.2.1 it was noted that the exact behaviour of the case $N = 3$ is unclear, but at least one and at most four pieces of data will be required to determine μ using three displacement fields. In Section 4.3.2.3, on the other hand, it was shown that no more than four linearly independent displacement fields can ever be available. Finally, in Section 4.3.3 an expression for λ in terms of μ is obtained. This expression contains an additive constant, so if four displacement fields are available then *two* pieces of boundary data are required to determine the Lamé constants — one to determine the multiplicative constant in the

expression for μ and the second to determine the additive constant in the expression for λ.

4.4 Numerical Solution of the Inverse Problem

In this section a finite element solution of the elasticity inverse problem is formulated. Since hyperbolic equations are difficult to solve numerically (due to the difficulty of ensuring that the numerical domain of dependence of any point is a subset of the true domain of dependence) the problem is rewritten in Section 4.4.1 as a minimization problem. In Section 4.4.2 the Euler-Lagrange equations of the minimization problem are derived using the theory described in Section 2.3. These are shown to be parabolic, but the limiting case of a series of elliptic systems, in Section 4.4.3. Elliptic systems are much easier to solve numerically than hyperbolic systems and in Section 4.4.4 a finite element solution method is described. This reduces the minimization problem to a linear system of algebraic equations that can be trivially solved using any standard linear algebra software.

4.4.1 The Minimization Problem

The first step in formulating a numerical solution method for the linear elasticity inverse problem is to convert it from a hyperbolic system of PDEs to a form that is more numerically tractable. In this section this process is begun by writing down a minimization problem that is equivalent to the system of PDEs. Begin, motivated by equation (4.4), by defining residuals

$$R_i = \frac{\partial}{\partial x_i}\left(\lambda \frac{\partial u_k}{\partial x_k}\right) + \frac{\partial}{\partial x_j}\left(\mu \frac{\partial u_i}{\partial x_j}\right) + \frac{\partial}{\partial x_j}\left(\mu \frac{\partial u_j}{\partial x_i}\right), \quad i = 1, 2. \tag{4.79}$$

In terms of the components of strain these residuals can be written as

$$R_1 = \frac{\partial}{\partial x}(\lambda(\epsilon_{11} + \epsilon_{22}) + 2\mu\epsilon_{11}) + 2\frac{\partial}{\partial y}(\mu\epsilon_{12}), \tag{4.80}$$

$$R_2 = 2\frac{\partial}{\partial x}(\mu\epsilon_{12}) + \frac{\partial}{\partial y}(\lambda(\epsilon_{11} + \epsilon_{22}) + 2\mu\epsilon_{22}). \tag{4.81}$$

If λ, μ are such that they solve equation (4.4) then the two residuals will be equal to zero. So, finding such λ, μ can be reformulated as the following minimization problem:

$$\text{Find } \lambda, \mu \text{ to minimize } I = \int_{\Omega} \left(R_1^2 + R_2^2\right) dxdy. \tag{4.82}$$

This minimization problem is to be solved directly using a finite element method, however in order to ensure that it is elliptic it is convenient to formulate the relevant Euler-Lagrange equations.

4.4.2 The Euler-Lagrange Equations

In this section the Euler-Lagrange equations of the minimization problem defined in the previous section are derived. This work is based on the theory of the calculus of variations presented in Section 2.3. In the case of two space dimensions (x, y) and two independent variables (λ, μ) the Euler equations, defined in equations (2.132) – (2.134), are

$$\frac{\partial}{\partial x}\left(\frac{\partial F}{\partial \lambda_x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial \lambda_y}\right) = \frac{\partial F}{\partial \lambda}, \tag{4.83}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial F}{\partial \mu_x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial \mu_y}\right) = \frac{\partial F}{\partial \mu}, \tag{4.84}$$

where

$$F = R_1^2 + R_2^2. \tag{4.85}$$

Since only the classification of the Euler-Lagrange equations is of interest (and not their solution since the minimization problem is to be solved directly) the boundary conditions arising from the derivation of the Euler equations are not considered here.

The derivatives of F with respect to λ, μ and their partial derivatives are

$$\frac{\partial F}{\partial \lambda} = 2(\epsilon_{11} + \epsilon_{22})_x R_1 + 2(\epsilon_{11} + \epsilon_{22})_y R_2, \tag{4.86}$$

$$\frac{\partial F}{\partial \lambda_x} = 2(\epsilon_{11} + \epsilon_{22}) R_1, \tag{4.87}$$

$$\frac{\partial F}{\partial \lambda_y} = 2(\epsilon_{11} + \epsilon_{22}) R_2, \tag{4.88}$$

$$\frac{\partial F}{\partial \mu} = 4((\epsilon_{11})_x + (\epsilon_{12})_y) R_1 + 4((\epsilon_{12})_x + (\epsilon_{22})_y) R_2, \tag{4.89}$$

$$\frac{\partial F}{\partial \mu_x} = 4(\epsilon_{11} R_1 + \epsilon_{12} R_2), \tag{4.90}$$

$$\frac{\partial F}{\partial \mu_y} = 4(\epsilon_{12} R_1 + \epsilon_{22} R_2). \tag{4.91}$$

On substituting for R_1 and R_2 and simplifying, then the Euler equations become

$$0 = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)(\lambda(\epsilon_{11} + \epsilon_{22})) + 2\left(\frac{\partial^2}{\partial x^2}(\mu\epsilon_{11}) + 2\frac{\partial^2}{\partial x \partial y}(\mu\epsilon_{12}) + \frac{\partial^2}{\partial y^2}(\mu\epsilon_{22})\right), \tag{4.92}$$

$$\begin{aligned} 0 = {} & \epsilon_{11}\left(\frac{\partial^2}{\partial x^2}(\lambda(\epsilon_{11} + \epsilon_{22})) + 2\left(\frac{\partial^2}{\partial x^2}(\mu\epsilon_{11}) + \frac{\partial^2}{\partial x \partial y}(\mu\epsilon_{12})\right)\right) \\ & + 2\epsilon_{12}\left(\frac{\partial^2}{\partial x \partial y}((\lambda + \mu)(\epsilon_{11} + \epsilon_{22})) + \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)(\mu\epsilon_{12})\right) \\ & + \epsilon_{22}\left(\frac{\partial^2}{\partial y^2}(\lambda(\epsilon_{11} + \epsilon_{22})) + 2\left(\frac{\partial^2}{\partial x \partial y}(\mu\epsilon_{12}) + \frac{\partial^2}{\partial y^2}(\mu\epsilon_{22})\right)\right). \end{aligned} \tag{4.93}$$

Observe that the trivial case $\lambda \equiv 0$, $\mu \equiv 0$ is a possible solution of the Euler equations, and further that if one non-zero solution can be determined then any multiple of this solution will also satisfy equations (4.92) and (4.93). In order to ensure that there is a unique solution the constraint

$$J(\lambda) = \int_\Omega \lambda dx dy = 1, \tag{4.94}$$

is imposed. Recall that a method for dealing with integral constraints by introducing a Lagrange multiplier L and replacing the Euler equations with the Euler-Lagrange equations was discussed in Section 2.3.2. The general form of the Euler-Lagrange equation in 1-d was given in equation (2.142) and the generalization to multiple variable and dimensions works in exactly the same way as for the Euler equations. Since J is not a function of μ (or its derivatives) equation (4.93) remains unchanged. Equation (4.92) is replaced by the

Euler-Lagrange equation

$$0 = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)(\lambda(\epsilon_{11} + \epsilon_{22})) + 2\left(\frac{\partial^2}{\partial x^2}(\mu\epsilon_{11}) + 2\frac{\partial^2}{\partial x \partial y}(\mu\epsilon_{12}) + \frac{\partial^2}{\partial y^2}(\mu\epsilon_{22})\right) + \frac{L}{\epsilon_{11} + \epsilon_{22}}, \tag{4.95}$$

where L is the Lagrange multiplier.

4.4.3 Classification of the Euler-Lagrange Equations

In this section the classification of the Euler-Lagrange equations is investigated. This is based on the theory of the classification of second-order systems described in Section 2.2.1.4. Recall that the motivation behind reformulating the inverse problem as a minimization problem was to avoid the necessity of solving a hyperbolic system of PDEs numerically. In Section 4.4.3.1 the Euler-Lagrange equations are shown to be parabolic. Whilst this is preferable to a hyperbolic system it is easiest to solve an elliptic system numerically (since there are no real characteristics). In Section 4.4.3.2 it is shown that the Euler-Lagrange equations can be modified by the addition of an arbitrarily small term that guarantees that they become elliptic. Finally, in Section 4.4.3.3, appropriate boundary conditions for the Euler-Lagrange equations are discussed.

4.4.3.1 Exact Euler-Lagrange Equations

Since no derivatives of the Lagrange multiplier L are present in the Euler-Lagrange equations their classification is equivalent to that of the Euler equations, which can be rewritten in matrix form as

$$A\frac{\partial^2 \mathbf{u}}{\partial x^2} + 2B\frac{\partial^2 \mathbf{u}}{\partial x \partial y} + C\frac{\partial^2 \mathbf{u}}{\partial y^2} = \mathbf{d}, \tag{4.96}$$

where

$$A = \begin{pmatrix} \epsilon_{11} + \epsilon_{22} & 2\epsilon_{11} \\ \epsilon_{11}(\epsilon_{11} + \epsilon_{22}) & 2\left(\epsilon_{11}^2 + \epsilon_{12}^2\right) \end{pmatrix}, \tag{4.97}$$

$$2B = \begin{pmatrix} 0 & 4\epsilon_{12} \\ 2\epsilon_{12}(\epsilon_{11} + \epsilon_{22}) & 4\epsilon_{12}(\epsilon_{11} + \epsilon_{22}) \end{pmatrix}, \tag{4.98}$$

$$C = \begin{pmatrix} \epsilon_{11} + \epsilon_{22} & 2\epsilon_{22} \\ \epsilon_{22}(\epsilon_{11} + \epsilon_{22}) & 2\left(\epsilon_{12}^2 + \epsilon_{22}^2\right) \end{pmatrix}. \tag{4.99}$$

So, the characteristic equation (given by equation (2.103) in Section 2.2.1.4) is

$$\begin{vmatrix} (\epsilon_{11}+\epsilon_{22})(1+k^2) & 2\big(\epsilon_{22}-2k\epsilon_{12}+k^2\epsilon_{11}\big) \\ (\epsilon_{11}+\epsilon_{22})\big(\epsilon_{22}-2k\epsilon_{12}+k^2\epsilon_{11}\big) & 2\big(\epsilon_{12}^2+\epsilon_{22}^2-2k\epsilon_{12}(\epsilon_{11}+\epsilon_{22})+k^2\big(\epsilon_{11}^2+\epsilon_{12}^2\big)\big) \end{vmatrix} = 0, \tag{4.100}$$

where $|M| \equiv \det M$ for a matrix M. This is equivalent to the following quartic equation in k:

$$0 = \big(1+k^2\big)\big(\epsilon_{12}^2+\epsilon_{22}^2-2k\epsilon_{12}(\epsilon_{11}+\epsilon_{22})+k^2\big(\epsilon_{11}^2+\epsilon_{12}^2\big)\big)-\big(\epsilon_{22}-2k\epsilon_{12}+k^2\epsilon_{11}\big)^2. \tag{4.101}$$

$$\Rightarrow \quad 0 = \epsilon_{12}^2k^4+2\epsilon_{12}(\epsilon_{11}-\epsilon_{22})k^3+\big((\epsilon_{11}-\epsilon_{22})^2-2\epsilon_{12}^2\big)k^2-2\epsilon_{12}(\epsilon_{11}-\epsilon_{22})k+\epsilon_{12}^2. \tag{4.102}$$

This can be reduced to a biquadratic equation (a quadratic in k^2) by making the substitution

$$k = z-\frac{\epsilon_{11}-\epsilon_{22}}{2\epsilon_{12}}. \tag{4.103}$$

Then,

$$z^4-\left(2+\frac{(\epsilon_{11}-\epsilon_{22})^2}{2\epsilon_{12}^2}\right)z^2+\frac{(\epsilon_{11}-\epsilon_{22})^4}{16\epsilon_{12}^4}+\frac{(\epsilon_{11}-\epsilon_{22})^2}{2\epsilon_{12}^2}+1=0. \tag{4.104}$$

This equation has two distinct solutions,

$$z = \pm\sqrt{1+\frac{(\epsilon_{11}-\epsilon_{22})^2}{4\epsilon_{12}^2}}, \tag{4.105}$$

and so the Euler-Lagrange equations are neither elliptic nor hyperbolic.

4.4.3.2 Modified Euler-Lagrange Equations

In order to make the system elliptic (and, hence, make it easier to find a smooth solution numerically) 'smoothing' terms can be added to the Euler-Lagrange equations (4.93), (4.95)

so that they become

$$\begin{aligned}0 = &\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)(\lambda(\epsilon_{11}+\epsilon_{22}))+2\left(\frac{\partial^2}{\partial x^2}(\mu\epsilon_{11})+2\frac{\partial^2}{\partial x\partial y}(\mu\epsilon_{12})+\frac{\partial^2}{\partial y^2}(\mu\epsilon_{22})\right)\\&+\frac{\delta}{\epsilon_{11}+\epsilon_{22}}\left(\frac{\partial^2\lambda}{\partial x^2}+\frac{\partial^2\lambda}{\partial y^2}\right)+\frac{L}{\epsilon_{11}+\epsilon_{22}},\end{aligned} \tag{4.106}$$

$$\begin{aligned}0 = &\epsilon_{11}\left(\frac{\partial^2}{\partial x^2}(\lambda(\epsilon_{11}+\epsilon_{22}))+2\left(\frac{\partial^2}{\partial x^2}(\mu\epsilon_{11})+\frac{\partial^2}{\partial x\partial y}(\mu\epsilon_{12})\right)\right)\\&+2\epsilon_{12}\left(\frac{\partial^2}{\partial x\partial y}((\lambda+\mu)(\epsilon_{11}+\epsilon_{22}))+\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)(\mu\epsilon_{12})\right)\\&+\epsilon_{22}\left(\frac{\partial^2}{\partial y^2}(\lambda(\epsilon_{11}+\epsilon_{22}))+2\left(\frac{\partial^2}{\partial x\partial y}(\mu\epsilon_{12})+\frac{\partial^2}{\partial y^2}(\mu\epsilon_{22})\right)\right)+\tfrac{1}{2}\delta\left(\frac{\partial^2\mu}{\partial x^2}+\frac{\partial^2\mu}{\partial y^2}\right),\end{aligned} \tag{4.107}$$

for some $\delta > 0$. This is known as Tikhonov regularization (see Tikhonov and Arsenin [83]) and is equivalent to adding $\delta(|\nabla\lambda|^2 + |\nabla\mu|^2)$ to $F(\lambda, \mu)$ in equation (4.85). The modified system (again ignoring the Lagrange multiplier L) is defined by the matrix equation (4.96) where the matrices are now given by

$$A = \begin{pmatrix}(\epsilon_{11}+\epsilon_{22})^2+\delta & 2\epsilon_{11}(\epsilon_{11}+\epsilon_{22})\\ \epsilon_{11}(\epsilon_{11}+\epsilon_{22}) & 2\big(\epsilon_{11}^2+\epsilon_{12}^2\big)+\frac{1}{2}\delta\end{pmatrix}, \tag{4.108}$$

$$2B = \begin{pmatrix}0 & 4\epsilon_{12}(\epsilon_{11}+\epsilon_{22})\\ 2\epsilon_{12}(\epsilon_{11}+\epsilon_{22}) & 4\epsilon_{12}(\epsilon_{11}+\epsilon_{22})\end{pmatrix}, \tag{4.109}$$

$$C = \begin{pmatrix}(\epsilon_{11}+\epsilon_{22})^2+\delta & 2\epsilon_{22}(\epsilon_{11}+\epsilon_{22})\\ \epsilon_{22}(\epsilon_{11}+\epsilon_{22}) & 2\big(\epsilon_{12}^2+\epsilon_{22}^2\big)+\frac{1}{2}\delta\end{pmatrix}. \tag{4.110}$$

So, on a characteristic the determinant of the matrix

$$\begin{vmatrix}\big((\epsilon_{11}+\epsilon_{22})^2+\delta\big)(1+k^2) & 2(\epsilon_{11}+\epsilon_{22})\big(\epsilon_{22}-2k\epsilon_{12}+k^2\epsilon_{11}\big)\\ (\epsilon_{11}+\epsilon_{22})\big(\epsilon_{22}-2k\epsilon_{12}+k^2\epsilon_{11}\big) & 2\big(\epsilon_{12}^2+\epsilon_{22}^2+\frac{1}{4}\delta-2k\epsilon_{12}(\epsilon_{11}+\epsilon_{22})+k^2\big(\epsilon_{11}^2+\epsilon_{12}^2+\frac{1}{4}\delta\big)\big)\end{vmatrix} \tag{4.111}$$

is equal to zero, i.e. the characteristic equation is

$$\begin{aligned}0 = &\big((\epsilon_{11}+\epsilon_{22})^2+\delta\big)(1+k^2)\big(\epsilon_{12}^2+\epsilon_{22}^2+\tfrac{1}{4}\delta-2k\epsilon_{12}(\epsilon_{11}+\epsilon_{22})+k^2\big(\epsilon_{11}^2+\epsilon_{12}^2+\tfrac{1}{4}\delta\big)\big)\\&-(\epsilon_{11}+\epsilon_{22})^2\big(\epsilon_{22}^2-4k\epsilon_{12}\epsilon_{22}+2k^2\epsilon_{11}\epsilon_{22}+4k^2\epsilon_{12}^2-4k^3\epsilon_{11}\epsilon_{12}+k^4\epsilon_{11}^2\big).\end{aligned} \tag{4.112}$$

In order to solve this equation for $0 < \delta \ll 1$ consider the perturbation theory approximation:

$$k = k_0 + \delta^{1/2} k_1. \tag{4.113}$$

Then the characteristic equation can be rewritten as

$$\begin{aligned}
0 = {} & \big((\epsilon_{11}+\epsilon_{22})^2 + \delta\big)(1 + k_0^2 + 2\delta^{1/2}k_0k_1 + \delta k_1^2) \\
& \times \big(\epsilon_{12}^2 + \epsilon_{22}^2 + \tfrac{1}{4}\delta - 2\epsilon_{12}(k_0 + \delta^{1/2}k_1)(\epsilon_{11}+\epsilon_{22}) \\
& \quad + (k_0^2 + 2\delta^{1/2}k_0k_1 + \delta k_1^2)\big(\epsilon_{11}^2 + \epsilon_{12}^2 + \tfrac{1}{4}\delta\big)\big) \\
& - (\epsilon_{11}+\epsilon_{22})^2\big(\epsilon_{22}^2 - 4(k_0 + \delta^{1/2}k_1)\epsilon_{12}\epsilon_{22} + 2(k_0^2 + 2\delta^{1/2}k_0k_1 + \delta k_1^2)\epsilon_{11}\epsilon_{22} \\
& \qquad + 4(k_0^2 + 2\delta^{1/2}k_0k_1 + \delta k_1^2)\epsilon_{12}^2 \\
& \qquad - 4(k_0^3 + 3\delta^{1/2}k_0^2k_1 + 3\delta k_0k_1^2 + \delta^{3/2}k_1^3)\epsilon_{11}\epsilon_{12} \\
& \qquad + (k_0^4 + 4\delta^{1/2}k_0^3k_1 + 6\delta k_0^2k_1^2 + 4\delta^{3/2}k_0k_1^3 + \delta^2 k_1^4)\epsilon_{11}^2\big),
\end{aligned} \tag{4.114}$$

The $O(1)$ equation is simply equation (4.102), so k_0 is given by equations (4.103) and (4.105). Hence, the classification of the system of equations (4.106), (4.107) depends on the values of k_1. If there are four distinct values of $k_0 + \delta^{1/2}k_1$, all with no imaginary part then the system will become hyperbolic (note that since k_0 and δ have no imaginary part this depends entirely on the imaginary part of k_1). If all the possible values of $k_0 + \delta k_1$ (no matter how many of them there are) have a non-zero imaginary part then the system will become elliptic. If there are still only two possible values of $k_0 + \delta k_1$ and k_1 is entirely real then the system will remain neither hyperbolic nor parabolic.

Considering the $O(\delta^{1/2})$ terms in characteristic equation (4.114) gives

$$\begin{aligned}
0 = {} & (1 + k_0^2)\big(k_0(\epsilon_{11}^2 + \epsilon_{12}^2) - \epsilon_{12}(\epsilon_{11}+\epsilon_{22})\big) \\
& + k_0\big(\epsilon_{12}^2 + \epsilon_{22}^2 - 2\epsilon_{12}k_0(\epsilon_{11}+\epsilon_{22}) + k_0^2(\epsilon_{11}^2 + \epsilon_{12}^2)\big) \\
& + 2\big(\epsilon_{12}\epsilon_{22} - k_0\epsilon_{11}\epsilon_{22} - 2k_0\epsilon_{12}^2 + 3k_0^2\epsilon_{11}\epsilon_{12} - k_0^3\epsilon_{11}^2\big),
\end{aligned} \tag{4.115}$$

$$\Rightarrow \quad 0 = 2\epsilon_{12}k_0^3 + 3\epsilon_{12}(\epsilon_{11} - \epsilon_{22})k_0^2 + \big((\epsilon_{11}-\epsilon_{22})^2 - 2\epsilon_{12}^2\big)k_0 - \epsilon_{12}(\epsilon_{11}-\epsilon_{22}). \tag{4.116}$$

Observe that this expression has no dependence on k_1. Observe further that differentiating the right-hand-side of equation (4.102) with respect to k and setting $k = k_0$ gives a multiple of the right-hand-side of equation (4.116). Now, since equation (4.102) has a double root at $k = k_0$ its derivative must be zero there. Hence equation (4.116) is automatically satisfied by k_0.

So, in order to determine k_1 the $O(\delta)$ terms in equation (4.114) must be considered, giving

$$\begin{aligned} 0 = {} & \big(1 + k_0^2 + (\epsilon_{11} + \epsilon_{22})^2 k_1^2\big)\big((\epsilon_{12} - k_0\epsilon_{11})^2 + (\epsilon_{22} - k_0\epsilon_{12})^2\big) \\ & + 4k_0k_1^2(\epsilon_{11} + \epsilon_{22})^2\big(k_0\big(\epsilon_{11}^2 + \epsilon_{12}^2\big) - \epsilon_{12}(\epsilon_{11} + \epsilon_{22})\big) \\ & + (\epsilon_{11} + \epsilon_{22})^2(1 + k_0^2)\big(\tfrac{1}{4}(1 + k_0^2) + \big(\epsilon_{11}^2 + \epsilon_{12}^2\big)k_1^2\big) \\ & - 2(\epsilon_{11} + \epsilon_{22})^2\big(\epsilon_{11}\epsilon_{22} + 2\epsilon_{12}^2 - 6k_0\epsilon_{11}\epsilon_{12} + 3k_0^2\epsilon_{11}^2\big)k_1^2. \end{aligned} \tag{4.117}$$

Solving this equation for k_1^2 gives

$$k_1^2 = -\frac{\big(1 + k_0^2\big)\alpha}{(\epsilon_{11} + \epsilon_{22})^2\beta}, \tag{4.118}$$

where

$$\alpha = (\epsilon_{12} - k_0\epsilon_{11})^2 + (\epsilon_{22} - k_0\epsilon_{12})^2 + \tfrac{1}{4}(\epsilon_{11} + \epsilon_{22})^2\big(1 + k_0^2\big), \tag{4.119}$$

$$\begin{aligned} \beta = {} & (\epsilon_{12} - k_0\epsilon_{11})^2 + (\epsilon_{22} - k_0\epsilon_{12})^2 + 4k_0\big(k_0\big(\epsilon_{11}^2 + \epsilon_{12}^2\big) - \epsilon_{12}(\epsilon_{11} + \epsilon_{22})\big) \\ & + \big(1 + k_0^2\big)\big(\epsilon_{11}^2 + \epsilon_{12}^2\big) - 2\big(\epsilon_{11}\epsilon_{22} + 2\epsilon_{12}^2 - 6k_0\epsilon_{11}\epsilon_{12} + 3k_0^2\epsilon_{11}^2\big). \end{aligned} \tag{4.120}$$

Recall that the system is elliptic if and only if $k_1^2 < 0$. This is equivalent to requiring that α and β have the same sign. Clearly $\alpha > 0$ since all of the terms in it are squares. Further, the expression for β can be rearranged to give

$$\beta = 6\epsilon_{12}k_0^2 + 6\epsilon_{12}(\epsilon_{11} - \epsilon_{22})k_0 + (\epsilon_{11} - \epsilon_{22})^2 - 2\epsilon_{12}^2. \tag{4.121}$$

Now, recall that equation (4.102) has a double root at $k = k_0$. Observe that since $\epsilon_{12}^2 > 0$, the right-hand-side of equation (4.102) tends to infinity as $k \to \pm\infty$. So, $k = k_0$ is a *minimum* of the right-hand-side of equation (4.102). Hence, the second derivative of the right-hand-side is *positive*, i.e.

$$2\beta = 12\epsilon_{12}k_0^2 + 12\epsilon_{12}(\epsilon_{11} - \epsilon_{22})k_0 + 2(\epsilon_{11} - \epsilon_{22})^2 - 4\epsilon_{12}^2 > 0. \tag{4.122}$$

Thus, $\beta > 0$ and, hence, $k_1^2 < 0$ which means that the modified Euler-Lagrange equations are *elliptic* for any $\delta > 0$.

So, the Euler-Lagrange equations can be seen to be the limit of a series of elliptic equations. Hence, it is hoped that the numerical solution of them will be well-behaved. However, if it is not there is a simple strategy to ensure that it is: an additional term of the form

$\delta(|\nabla\lambda|^2+|\nabla\mu|^2)$, where $0<\delta$, can be added to the integrand in the minimization problem. In deriving a numerical solution to the minimization problem in the remainder of this chapter it is assumed that $\delta \neq 0$. However, the simulations of this numerical problem that are presented in Chapter 5 do in fact turn out to be well-behaved for $\delta = 0$.

4.4.3.3 Boundary Conditions

Like any system of PDEs the Euler-Lagrange equations require boundary conditions in order for a unique solution to exist. Recall from Section 2.2.2 that one type of appropriate boundary condition for an elliptic system of second-order PDEs is to specify the values of the variables everywhere on the boundary (the Dirichlet problem). Since the modified Euler-Lagrange equations are elliptic then clearly specifying λ and μ on the boundary will ensure that the problem is well-posed. Since the exact Euler-Lagrange equations are the limit of a series of elliptic equations (as $\delta \to 0$) it is assumed that Dircihlet conditions are also appropriate for this system. Again, if the numerical simulations show that this is not the case then a small term of the form $\delta(|\nabla\lambda|^2+|\nabla\mu|^2)$, where $0<\delta\ll 1$, can be added and the modified Euler-Lagrange equations solved instead. In fact, the simulations presented in Chapter 5 turn out to give the correct solution to the Euler-Lagrange equations with $\delta = 0$ (this is verified by comparing the solution to one with small but non-zero δ). Since the system with $\delta = 0$ is not elliptic this is not what we would expect to see — it is possible that a small amount of numerical diffusion is making the finite element formulation hyperbolic.

4.4.4 Finite Element Solution

The final step in solving the linear elasticity inverse problem is to write it in a form that can be solved by standard techniques of linear algebra, i.e. as a system of algebraic equations. In this section this is done using the finite element methods introduced in Section 2.4. In order to do this the minimization problem is first reformulated in finite element form in Section 4.4.4.1. Then, in Section 4.4.4.2 this is further converted to a matrix form. Recall that it is necessary to either specify boundary conditions or to use multiple displacement fields in order to determine a unique solution for the Lamé constants. In Section 4.4.4.3 the finite element form of the boundary conditions is discussed and added to the matrix form derived in Section 4.4.4.2. Finally, a way of including the information from multiple displacement fields in the finite element method is described in Section 4.4.4.4.

4.4.4.1 Finite Element Form

In order to solve the minimization problem using finite element methods it is more convenient to consider it in integral form than it is to formulate the Euler-Lagrange equations and then re-write them in weak form, i.e. the problem to be solved is to find λ, μ, L, to minimize

$$\int_\Omega F(x,y,\lambda,\lambda_x,\lambda_y,\mu,\mu_x,\mu_y)dxdy = \int_\Omega \big(R_1^2+R_2^2+\delta\big(|\nabla\lambda|^2+|\nabla\mu|^2\big)-L\lambda\big)dxdy. \quad (4.123)$$

Recall from the theory of the calculus of variations given in Section 2.3 (in particular from Lemma 2.1) that at a minimum of the functional given in equation (4.123) its Gâteaux derivatives (defined in Definition 2.5) must be zero. Then on calculating these Gâteaux derivatives, it can be seen that at the minimum the following integral relationships hold for any pair of test functions η_1 and η_2:

$$0 = \int_\Omega \Bigg(R_1\frac{\partial}{\partial x}(\eta_1(\epsilon_{11}+\epsilon_{22})) + R_2\frac{\partial}{\partial y}(\eta_1(\epsilon_{11}+\epsilon_{22})) + \delta\left(\frac{\partial\eta_1}{\partial x}\frac{\partial\lambda}{\partial x}+\frac{\partial\eta_1}{\partial y}\frac{\partial\lambda}{\partial y}\right)-\tfrac{1}{2}L\eta_1\Bigg)dxdy, \quad (4.124)$$

$$0 = \int_\Omega \Bigg(R_1\left(\frac{\partial}{\partial x}(\eta_2\epsilon_{11})+\frac{\partial}{\partial y}(\eta_2\epsilon_{12})\right)+R_2\left(\frac{\partial}{\partial x}(\eta_2\epsilon_{12})+\frac{\partial}{\partial y}(\eta_2\epsilon_{22})\right) + \tfrac{1}{2}\delta\left(\frac{\partial\eta_2}{\partial x}\frac{\partial\mu}{\partial x}+\frac{\partial\eta_2}{\partial y}\frac{\partial\mu}{\partial y}\right)\Bigg)dxdy. \quad (4.125)$$

These two equations can be used to give a finite element representation of the minimization problem.

In order to do this, consider a finite element mesh (with m elements and N nodes) on which functions are considered to be piecewise linear. Then write

$$\lambda(x,y) = \Lambda_j\phi_j(x,y), \quad \mu(x,y) = M_j\phi_j(x,y), \quad (4.126)$$

where $\phi_j(x,y)$ are the piecewise linear basis functions defined to be equal to 1 at node j, to decay linearly to zero on the elements containing node j and to be equal to zero elsewhere

in Ω. Then equations (4.124) and (4.125) can be rewritten as

$$0 = \sum_{p=1}^{m}\sum_{j=1}^{N}\int_{\Delta_p}\left(R_1\frac{\partial}{\partial x}(\phi_i(\epsilon_{11}+\epsilon_{22})) + R_2\frac{\partial}{\partial y}(\phi_i(\epsilon_{11}+\epsilon_{22})) + \delta\Lambda_j\left(\frac{\partial\phi_i}{\partial x}\frac{\partial\phi_j}{\partial x}+\frac{\partial\phi_i}{\partial y}\frac{\partial\phi_j}{\partial y}\right) - \tfrac{1}{2}L\phi_i\right)dxdy, \tag{4.127}$$

$$0 = \sum_{p=1}^{m}\sum_{j=1}^{N}\int_{\Delta_p}\left(R_1\left(\frac{\partial}{\partial x}(\phi_i\epsilon_{11})+\frac{\partial}{\partial y}(\phi_i\epsilon_{12})\right) + R_2\left(\frac{\partial}{\partial x}(\phi_i\epsilon_{12})+\frac{\partial}{\partial y}(\phi_i\epsilon_{22})\right) + \tfrac{1}{2}\delta M_j\left(\frac{\partial\phi_i}{\partial x}\frac{\partial\phi_j}{\partial x}+\frac{\partial\phi_i}{\partial y}\frac{\partial\phi_j}{\partial y}\right)\right)dxdy, \tag{4.128}$$

both for $i = 1, \ldots, N$, where Δ_p is the pth element of the mesh. In addition the integral constraint can be written in discrete form as

$$\sum_{p=1}^{m}\sum_{j=1}^{N}\Lambda_j\int_{\Delta_p}\phi_j dxdy = 1, \tag{4.129}$$

to give a total of $2N+1$ linear equations (since R_1 and R_2 depend linearly on Λ_j and M_j). These can be solved directly for the $2N+1$ unknowns (the Lagrange multiplier k, Λ_j, M_j, $j = 1, \ldots, N$).

4.4.4.2 Matrix Form

Next, the system is rewritten in matrix form in preparation for solving it using linear algebra software. It is clear that this can be achieved by writing

$$A\mathbf{x} = \mathbf{b}, \tag{4.130}$$

where $\mathbf{x} = (\Lambda_1, \ldots, \Lambda_N, M_1, \ldots, M_N, L)^T$, $\mathbf{b}$ is a $(2N+1)$-vector were all of the entries are zero except for the $(2N+1)$-th which is 1 and

$$A_{ij} = \int_\Omega \left(\frac{\partial}{\partial x}(\phi_i(\epsilon_{11}+\epsilon_{22}))\frac{\partial}{\partial x}(\phi_j(\epsilon_{11}+\epsilon_{22})) + \frac{\partial}{\partial y}(\phi_i(\epsilon_{11}+\epsilon_{22}))\frac{\partial}{\partial y}(\phi_j(\epsilon_{11}+\epsilon_{22})) \right.$$
$$\left. + \delta\left(\frac{\partial \phi_i}{\partial x}\frac{\partial \phi_j}{\partial x} + \frac{\partial \phi_i}{\partial y}\frac{\partial \phi_j}{\partial y} \right) \right) dxdy, \tag{4.131}$$

if $1 \leqslant i \leqslant N$, $1 \leqslant j \leqslant N$,

$$A_{ij} = 2\int_\Omega \left(\frac{\partial}{\partial x}(\phi_i(\epsilon_{11}+\epsilon_{22}))\left(\frac{\partial}{\partial x}(\phi_j\epsilon_{11}) + \frac{\partial}{\partial y}(\phi_j\epsilon_{12}) \right) \right.$$
$$\left. + \frac{\partial}{\partial y}(\phi_i(\epsilon_{11}+\epsilon_{22}))\left(\frac{\partial}{\partial x}(\phi_j\epsilon_{12}) + \frac{\partial}{\partial y}(\phi_j\epsilon_{22}) \right) \right) dxdy, \tag{4.132}$$

if $1 \leqslant i \leqslant N$, $N+1 \leqslant j \leqslant 2N$,

$$A_{ij} = -\tfrac{1}{2}\phi_j, \tag{4.133}$$

if $1 \leqslant i \leqslant N$, $j = 2N+1$,

$$A_{ij} = \int_\Omega \left(\left(\frac{\partial}{\partial x}(\phi_i\epsilon_{11}) + \frac{\partial}{\partial y}(\phi_i\epsilon_{12}) \right) \frac{\partial}{\partial x}(\phi_j(\epsilon_{11}+\epsilon_{22})) \right.$$
$$\left. + \left(\frac{\partial}{\partial x}(\phi_i\epsilon_{12}) + \frac{\partial}{\partial y}(\phi_i\epsilon_{22}) \right) \frac{\partial}{\partial y}(\phi_j(\epsilon_{11}+\epsilon_{22})) \right) dxdy, \tag{4.134}$$

if $N+1 \leqslant i \leqslant 2N$, $1 \leqslant j \leqslant N$,

$$A_{ij} = 2\int_\Omega \left(\left(\frac{\partial}{\partial x}(\phi_i\epsilon_{11}) + \frac{\partial}{\partial y}(\phi_i\epsilon_{12}) \right) \left(\frac{\partial}{\partial x}(\phi_j\epsilon_{11}) + \frac{\partial}{\partial y}(\phi_j\epsilon_{12}) \right) \right.$$
$$+ \left(\frac{\partial}{\partial x}(\phi_i\epsilon_{12}) + \frac{\partial}{\partial y}(\phi_i\epsilon_{22}) \right) \left(\frac{\partial}{\partial x}(\phi_j\epsilon_{12}) + \frac{\partial}{\partial y}(\phi_j\epsilon_{22}) \right)$$
$$\left. + \tfrac{1}{4}\delta\left(\frac{\partial \phi_i}{\partial x}\frac{\partial \phi_j}{\partial x} + \frac{\partial \phi_i}{\partial y}\frac{\partial \phi_j}{\partial y} \right) \right) dxdy, \tag{4.135}$$

if $N+1 \leqslant i \leqslant 2N$, $N+1 \leqslant j \leqslant 2N$,

$$A_{ij} = \int_\Omega \phi_j dxdy, \tag{4.136}$$

if $i = 2N + 1$, $1 \leqslant j \leqslant N$,

$$A_{ij} = 0, \tag{4.137}$$

otherwise.

4.4.4.3 Boundary Conditions

Recall from Section 4.4.3.2 that the modified Euler-Lagrange equations (the equations with non-zero δ) for the minimization problem are elliptic, so appropriate boundary data is to prescribe both λ and μ everywhere around the boundary $\partial\Omega$, i.e.

$$\lambda(\mathbf{x}) = \lambda_0(\mathbf{x}), \tag{4.138}$$

$$\mu(\mathbf{x}) = \mu_0(\mathbf{x}), \tag{4.139}$$

for $\mathbf{x} \in \partial\Omega$. In this section the finite element scheme is modified to allow for the specification of this boundary data.

In order to do this, suppose that node i is on the boundary (at $\mathbf{x}_i$, say). Then the entries of the matrix A in the previous section are altered so

$$A_{ij} = \delta_{ij}, \tag{4.140}$$

$$A_{i+N,j} = \delta_{i+N,j}, \tag{4.141}$$

where δ is Kronecker delta. The vector $\mathbf{b}$ is updated by specifying

$$b_i = \lambda_0(\mathbf{x}_i), \tag{4.142}$$

$$b_{i+N} = \mu_0(\mathbf{x}_i). \tag{4.143}$$

If boundary data is applied the solution is determined uniquely without the need for the Lagrange multiplier, so the $(2N+1)$th row and column of A are ignored, as is the $(2N+1)$th entry of $\mathbf{b}$.

4.4.4.4 Multiple Displacement Fields

One major advantage of the minimization method is that it makes the consideration of multiple displacement fields much easier. In this section the finite element method is

adapted to allow for the use of more than one displacement field. This is done by defining appropriate residuals for each available displacement field and then minimizing the sum of all of the residuals.

Consider a series of compatible displacement fields $\mathbf{u}^{(i)}$, $i = 1, \ldots, N$ with corresponding strain fields $\epsilon^{(i)}$, $i = 1, \ldots, N$. Define the residuals

$$R_1^{(i)} = \frac{\partial}{\partial x}\left(\lambda\left(\epsilon_{11}^{(i)} + \epsilon_{22}^{(i)}\right) + 2\mu\epsilon_{11}^{(i)}\right) + 2\frac{\partial}{\partial y}\left(\mu\epsilon_{12}^{(i)}\right), \tag{4.144}$$

$$R_2^{(i)} = 2\frac{\partial}{\partial x}\left(\mu\epsilon_{12}^{(i)}\right) + \frac{\partial}{\partial y}\left(\lambda\left(\epsilon_{11}^{(i)} + \epsilon_{22}^{(i)}\right) + 2\mu\epsilon_{22}^{(i)}\right), \tag{4.145}$$

$i = 1, \ldots, N$. Then, each of the residuals should be equal to zero if λ and μ solve the linear elasticity inverse problem for all of the displacement fields. Since the displacement fields are compatible such a λ and μ do exist, and the minimization problem becomes:

$$\text{Find } \lambda, \mu \text{ to minimize } I = \int_\Omega \left(\sum_{i=1}^{N} \left[\left(R_1^{(i)}\right)^2 + \left(R_2^{(i)}\right)^2\right] + \delta\left(|\nabla\lambda|^2 + |\nabla\mu|^2\right) - L\lambda \right) dxdy. \tag{4.146}$$

Since the integrand here is a linear combination of integrands of the form given in equation (4.123) all of the analysis of the previous sections holds for this minimization problem and it reduces to solving a matrix equation of the form

$$\left(\sum_{i=1}^{N} A^{(i)}\right)\mathbf{x} = \mathbf{b}, \tag{4.147}$$

where $\mathbf{b}$ is a $(2N+1)$-vector with all entries zero except for a 1 in the $(2N+1)$th position. The entries of each $A^{(i)}$ are given by replacing ϵ with $\epsilon^{(i)}$ in equations (4.131) – (4.137).

4.4.5 Summary

In this section the linear elasticity inverse problem has been reformulated into an algebraic system of equations that are trivial to solve numerically. First, in Section 4.4.1 the inverse problem introduced in Section 4.1 was rewritten as a minimization problem by introducing residuals that are zero when the equations of linear elasticity are satisfied and strictly positive otherwise (i.e. a solution minimizes the sum of the residuals). The Euler-Lagrange equations of the minimization were calculated in Section 4.4.2, using the theory of Section

2.3. In Section 4.4.3 the Euler-Lagrange equations were shown to be neither hyperbolic nor elliptic and modified by the addition of an additional, arbitrarily small term to make them elliptic and, hence, ensure that they are well-behaved numerically. Finally, in Section 4.4.4 a method for solving the minimization problem using finite element methods was described. This method solves the inverse problem both if boundary data is specified or if multiple displacement fields are used. In summary, this section provides a numerical method for recovering the tissue properties of a region of linearly elastic soft tissue from a knowledge of one or more displacement fields.

Chapter 5

Testing the Parameter Recovery Method

In this chapter some of the initial results of the finite element solution of the linear elasticity inverse problem are presented. This is an original contribution of the thesis. The parameter recovery method is shown to work well, except in the case of discontinuities in λ and μ. In Section 5.1 the method is tested on exact solutions of the forward problem. In this section the method can only be tested using one displacement field and boundary conditions since four independent exact solutions cannot be obtained explicitly for any Lamé constants. In Section 5.1.2 the failure of the solution in the case of discontinuous Lamé constants is seen.

Since very few explicit exact solutions are known for continuous (and non-constant) λ and μ, in Section 5.2 the parameter recovery method is tested using a variety of numerical solutions of the forward problem. It is seen to work well if the distributions of λ and μ are continuous, even if they have discontinuous first derivatives. However, when four displacement fields are used the values of λ and μ always approach zero near the corners. This behaviour is explained in Section 5.3 in terms of stress singularities caused by the change in applied boundary conditions at the corners.

Finally, in Section 5.4 the robustness of the method in response to errors in the data that it is given is tested. Firstly, the method is tested when errors are added to the displacement field given. This is done in the form of normally distributed 'noise'. The method turns out to be able to deal with errors in displacement estimation of below 1% of a typical displacement, but not with larger errors. Secondly, the effect of specifying

incorrect boundary data for the Lamé constants is investigated. In this respect the method turns out to be much more robust — even relatively large errors in the boundary data do not propagate into the centre of the domain.

All of the solutions in this chapter are calculated by solving the finite element formulation of either the forward or inverse linear elasticity problems on a single finite element mesh, consisting of 5248 elements and 2705 nodes.[1] For the forward problem this formulation was described in Section 2.4, for the inverse problem in Section 4.4.4. The solutions to the forward problem are generated using the quadratic basis functions defined in equations (2.191) – (2.196) (to ensure that strain is continuous) and those to the inverse problem using the linear basis functions defined in equations (2.184) – (2.186). In both cases the resulting systems of linear algebraic equations are solved using the 'backslash' command in Matlab 7. This command solves linear systems by Gaussian elimination (see, e.g., Süli and Mayers [81]).

5.1 Explicit Solutions of the Forward Problem

In order to validate the solution method, explicit solutions of the linear elasticity forward problem are calculated and the algorithm described in Section 4.4 is used to recover the parameters present in the solution. Very few explicit solutions of the equations of linear elasticity have been determined when the Lamé constants vary over the domain. Further, in the cases where solutions do exist they are usually only for one specific type of displacement so there is little chance of obtaining four compatible displacement fields, so only the method using boundary conditions can be tested.

Recall from Section 4.1.2 that the parameter recovery method assumes that the region of interest is undergoing *plane strain.* Fortunately, this is one of the few types of geometry for which explicit solutions of the elasticity equations do exist. Less fortunately, virtually all of the solutions are based on the further assumption that λ and μ are constant throughout the domain (or that they are infinite in some part of the domain). This makes the solutions particularly poor candidates for testing a method of determining how the values of λ and μ vary. Nevertheless, in the first part of this section the method is tested on a displacement

[1]There is one exception: when the existence of stress singularities at the corners of the domain is considered in Section 5.3 a coarser mesh is used to show that the apparent singularities do not vanish as the mesh is refined (and the solution converges to the exact solution).

field resulting from the uniaxial compression of a material with constant λ and μ. In the second part the method is tested on the displacement field generated by the uniaxial compression of a material with a cylindrical inclusion having different values of λ and μ from the surrounding matrix. Recall from Section 2.1.3.1 that no solutions of the linear elasticity forward problem with continuously varying λ and μ can be found in the literature.

5.1.1 Uniaxial Tension

One of the simplest explicit solutions of the plane strain linear elasticity equations is the displacement of a material undergoing uniaxial tension (or compression). Consider material occupying the square $(x, y) \in [-1, 1] \times [-1, 1] \subset \mathbb{R}^2$. Suppose that the material is linearly elastic, i.e. it satisfies equation (4.4), and is subject to boundary conditions

$$\sigma_{11} = T, \quad \sigma_{12} = 0, \quad \text{on } x = \pm 1, \tag{5.1}$$

$$\sigma_{12} = \sigma_{22} = 0, \quad \text{on } y = \pm 1, \tag{5.2}$$

i.e. the ends $x = \pm 1$ are subject to a uniform tensile force of magnitude T per unit length acting in the x-direction.[2] Suppose further that λ and μ are constant functions on the domain of interest. Then a solution for the strain is

$$\epsilon_{11} = \frac{(\lambda + 2\mu)T}{4\mu(\lambda + \mu)} = \frac{T}{E}, \tag{5.3}$$

$$\epsilon_{22} = -\frac{\lambda T}{4\mu(\lambda + \mu)} = -\frac{T\nu}{E}, \tag{5.4}$$

$$\epsilon_{12} = 0, \tag{5.5}$$

where

$$E = \frac{4\mu(\lambda + \mu)}{\lambda + 2\mu}, \tag{5.6}$$

$$\nu = \frac{\lambda}{\lambda + 2\mu}. \tag{5.7}$$

[2]Note that this defines uniaxial *tension* if $T > 0$ and uniaxial *compression* if $T < 0$.

Then solving for $\mathbf{u}$ using the definition of Cauchy's infinitesimal strain tensor ϵ given in equation (2.24) gives

$$\mathbf{u} = \left(\frac{Tx}{E}, -\frac{T\nu y}{E}\right) = (\alpha x, -\alpha\nu y), \tag{5.8}$$

where α depends on the material properties and the applied tension T.

In order to test the parameter recovery method a displacement field of the form given in equation (5.8) is used as input data on the unit square, with $\alpha = 0.1$, $\nu = 3/7$. This is equivalent to imposing a uniaxial compression so that $v = -0.1$ on the line $y = 1$ in a material where $\lambda = 1$, $\mu = \frac{2}{3}$. The resulting solution for λ and μ is shown in Figure 5.1. It can be seen that there is a good qualitative agreement between the actual solutions and the input data — λ and μ both appear to be constant functions with the values specified by the data. In order to quantify the agreement with the true solution the residuals of the method are plotted in Figure 5.2. It can be seen that away from the boundaries $x = 0, 1$ the residuals are of the order of 10^{-10}. This indicates that the solution in the interior of the domain is in very good agreement with the true solution.

5.1.2 Goodier's Solution for a Cylindrical Inclusion

In this section consider a material with Lamé constants λ_1, μ_1 containing a cylindrical inclusion with Lamé constants λ_2, μ_2 of radius a centred on $\mathbf{x} = (0, 0)^T$. Suppose that the material is assumed to be undergoing plane strain and is subject to uniaxial tension in the y-direction, so that $\sigma_{11} \to T$, $\sigma_{12} \to 0$ as $x \to \pm\infty$ and $\sigma_{12}, \sigma_{22} \to 0$ as $y \to \pm\infty$ (c.f. the boundary conditions for simple uniaxial compression in the previous section). An explicit solution of the equations of linear elasticity in this case has been determined by Goodier [32] using plane harmonic functions. This method is described in detail in Appendix C.1.

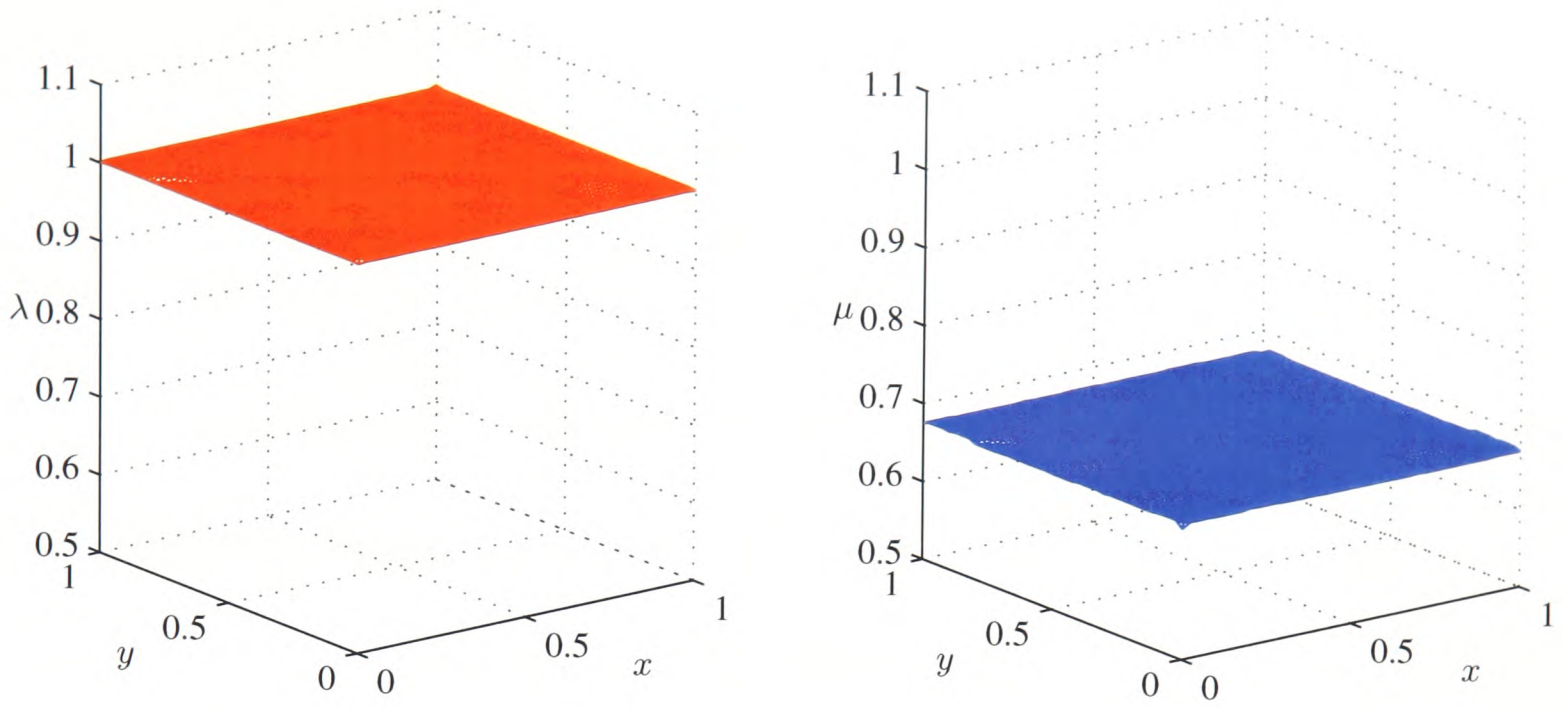


Figure 5.1: Numerical solution of the inverse problem using a displacement field corresponding to uniaxial tension with $\lambda = 1$, $\mu = \frac{2}{3}$ everywhere on the numerical domain. Dirichlet boundary conditions $\lambda = 1$, $\mu = \frac{2}{3}$ were applied all round the boundary.

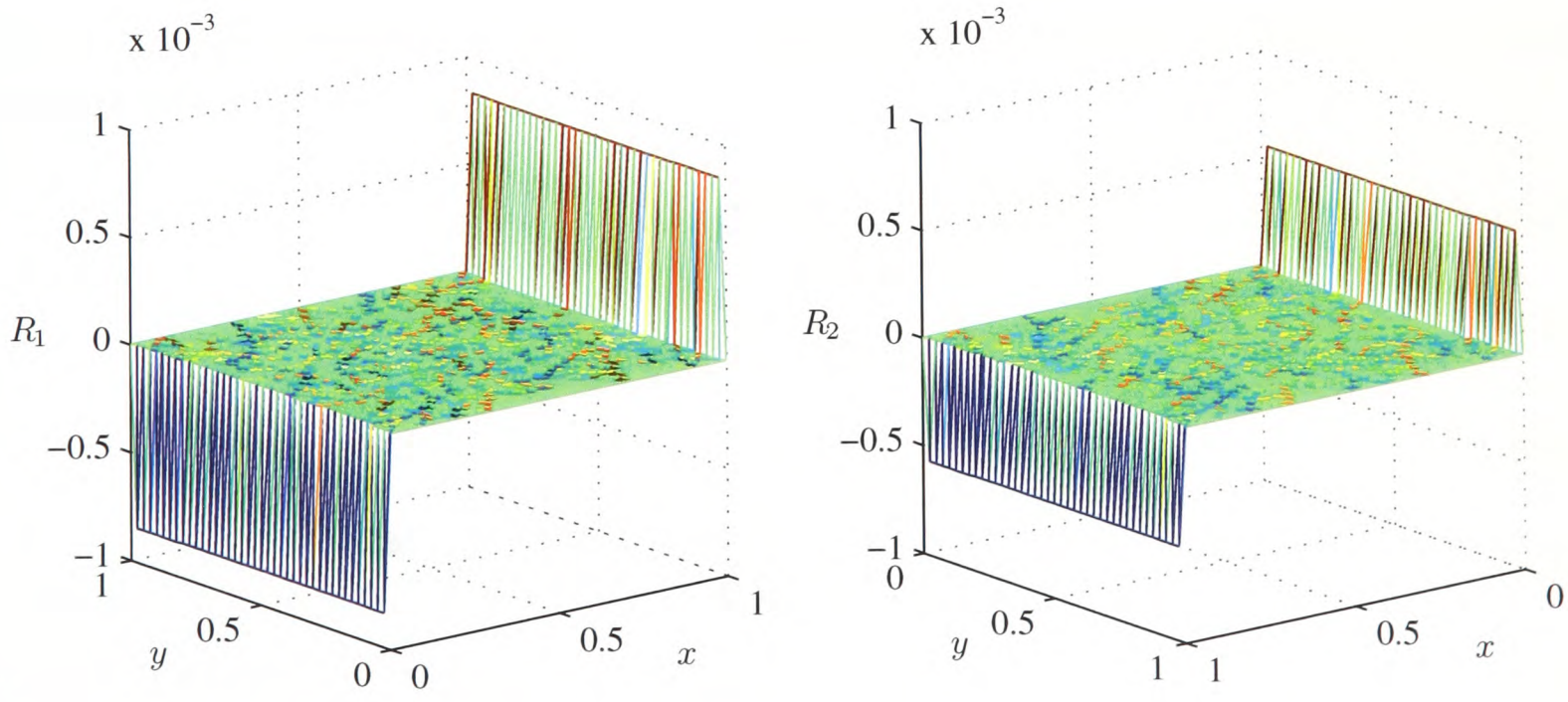


Figure 5.2: Residuals resulting from the numerically determined Lamé constants shown in Figure 5.1. The colour scale runs from -5×10^{-10} to -5×10^{-10}, so for any region which is not coloured dark red or dark blue $|R_i| < 5 \times 10^{-10}$.

The solution takes the form

$$u_r = \frac{A}{r} + \left[-\frac{B}{r^3} + \frac{2C}{r}(1-\nu_1) \right] \cos 2\theta + \frac{Tr}{4\mu_1}(1 - 2\nu_1 + \cos 2\theta), \tag{5.9}$$

$$u_\theta = -\left[\frac{B}{r^3} + \frac{C}{r}(1-2\nu_1) \right] \sin 2\theta - \frac{Tr}{4\mu_1} \sin 2\theta, \tag{5.10}$$

$$\sigma_{rr} = 2\mu_1 \left[-\frac{A}{r^2} + \left(\frac{3B}{r^4} - \frac{2C}{r^2} \right) \cos 2\theta \right] + \frac{T}{2}(1 + \cos 2\theta), \tag{5.11}$$

$$\sigma_{r\theta} = 2\mu_1 \left[\frac{3B}{r^4} - \frac{C}{r^2} \right] \sin 2\theta - \frac{T}{2} \sin 2\theta. \tag{5.12}$$

in the region $r \geqslant a$, where $r^2 = x^2 + a^2$, and

$$u_r = Fr + \left[Gr + 2\nu_2 H r^3 \right] \cos 2\theta, \tag{5.13}$$

$$u_\theta = -\left[Gr + (3 - 2\nu_2) H r^3 \right] \sin 2\theta, \tag{5.14}$$

$$\sigma_{rr} = 2\mu_2 \left[\frac{F}{1-2\nu_2} + G \cos 2\theta \right], \tag{5.15}$$

$$\sigma_{\theta\theta} = 2\mu_2 \left[\frac{F}{1-2\nu_2} - \left(G + 6Hr^2 \right) \cos 2\theta \right], \tag{5.16}$$

$$\sigma_{r\theta} = -2\mu \left[G + 3Hr^2 \right] \sin 2\theta, \tag{5.17}$$

in $r < a$. The constants A, B, C, F, G and H are determined in terms of the Lamé constants by

$$A = \frac{Ta^2}{4\mu_1} \left(\frac{(1-2\nu_2)\mu_1 - (1-2\nu_1)\mu_2}{(1-2\nu_2)\mu_1 + \mu_2} \right), \tag{5.18}$$

$$B = \frac{Ta^4}{4\mu_1} \left(\frac{\mu_1 - \mu_2}{\mu_1 + (3-4\nu_1)\mu_2} \right), \tag{5.19}$$

$$C = \frac{Ta^2}{2\mu_1} \left(\frac{\mu_1 - \mu_2}{\mu_1 + (3-4\nu_1)\mu_2} \right), \tag{5.20}$$

$$F = \frac{T}{2} \left(\frac{(1-2\nu_2)(1-\nu_1)}{(1-2\nu_2)\mu_1 + \mu_2} \right), \tag{5.21}$$

$$G = T \left(\frac{(1-\nu_1)}{\mu_1 + (3-4\nu_1)\mu_2} \right), \tag{5.22}$$

$$H = 0, \tag{5.23}$$

where ν_1, ν_2 are the Poisson's ratios in $r > a$ and $r < a$ respectively and are determined in terms of the appropriate Lamé constants by equation (5.7).

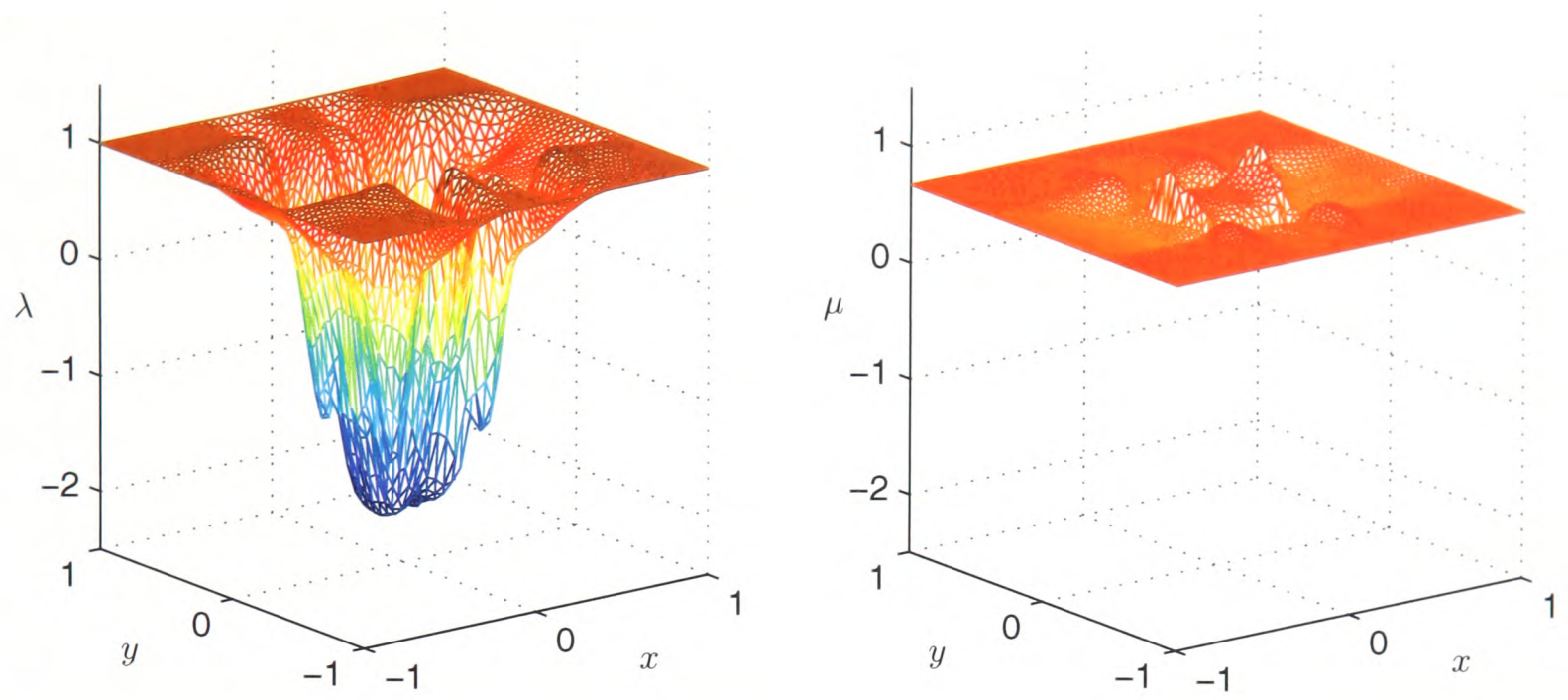


Figure 5.3: Numerical solution of the inverse problem using a displacement field corresponding to Goodier's solution for $\lambda_1 = 1$, $\mu_1 = \frac{2}{3}$ (in $r \geqslant 0.3$, where $r^2 = x^2 + y^2$), $\lambda_2 = 3$, $\mu_2 = 1$ (in $r < 0.3$). Dirichlet boundary conditions $\lambda = 1$, $\mu = \frac{2}{3}$ were applied all round the boundary.

The results of testing the parameter recovery method on Goodier's solution with $a = 0.3$, $\lambda_1 = 1$, $\mu_1 = \frac{2}{3}$, $\lambda_2 = 3$, $\mu_2 = 1$ are shown in Figure 5.3. Clearly the solutions are not a good match for the input parameters. This is not entirely unexpected — there are discontinuities in the values of λ and μ whereas the finite element basis functions used to approximate them are continuous. Notice, however, that there is clearly a feature in the correct place in the distribution of λ (although the value of λ within the feature is completely wrong) whereas μ is almost constant across the whole domain, compared with the actual input μ which is piecewise constant.

5.2 Numerical Solutions of the Forward Problem

Next, the method is tested by specifying $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$ throughout the domain of interest, solving a series of forward problems to obtain various displacement fields for these values of the Lamé constants and then attempting to reconstruct λ and μ using the parameter recovery method. This makes it easy to visually compare the computed solution with the true distribution of λ and μ. Additionally, the residuals at each point of the finite element method are calculated to provide a quantitative measure of the accuracy of the method. In

all cases the results obtained using one displacement field and applying boundary conditions are compared to those resulting from using four displacement fields and no boundary data.

When only one displacement field is used it is generated by applying zero stress boundary conditions on the lines $x = 0, 1$ and by setting $u, v = 0$ on $y = 0$ and $u = 0$, $v = -0.1$ on $y = 1$. This is equivalent to compressing the domain uniformly in the y-direction. If four displacement fields are used then the additional fields are generated by applying the following boundary conditions:

(i) Zero stress on $y = 0, 1$, $u, v = 0$ on $x = 0$, $u = -0.1$, $v = 0$ on $x = 1$, i.e. a compression in the x-direction.

(ii) Zero stress on $x = 0, 1$, $u, v = 0$ on $y = 0$, $u = 0.1$, $v = 0$ on $y = 1$, i.e. a shear is applied in the x-direction.

(iii) Zero stress on $y = 0, 1$, $u, v = 0$ on $y = 0$, $u = 0$, $v = 0.1$ on $y = 1$, i.e. a shear is applied in the y-direction.

These boundary conditions are shown schematically in Figure 5.4.

5.2.1 Smoothly Varying Lamé Constants

First, extremely smooth (in the sense of being infinitely continuously differentiable) functions are considered as the Lamé constants. The functions

$$\lambda_{\text{in}}(x, y) = 1 + 9\exp[-20((x - 0.5)^2 + (y - 0.5)^2)], \tag{5.24}$$

$$\mu_{\text{in}}(x, y) = \tfrac{2}{3} + 6\exp[-20((x - 0.5)^2 + (y - 0.5)^2)], \tag{5.25}$$

are used to generate displacement fields. The spatial variation of these functions over the unit square is shown in Figure 5.5. It can be seen that both λ and μ take larger values in the centre of the domain than near the edges. The functions are intended to be a continuous first approximation to the distribution of the Lamé constants that might be seen in a real tumour.

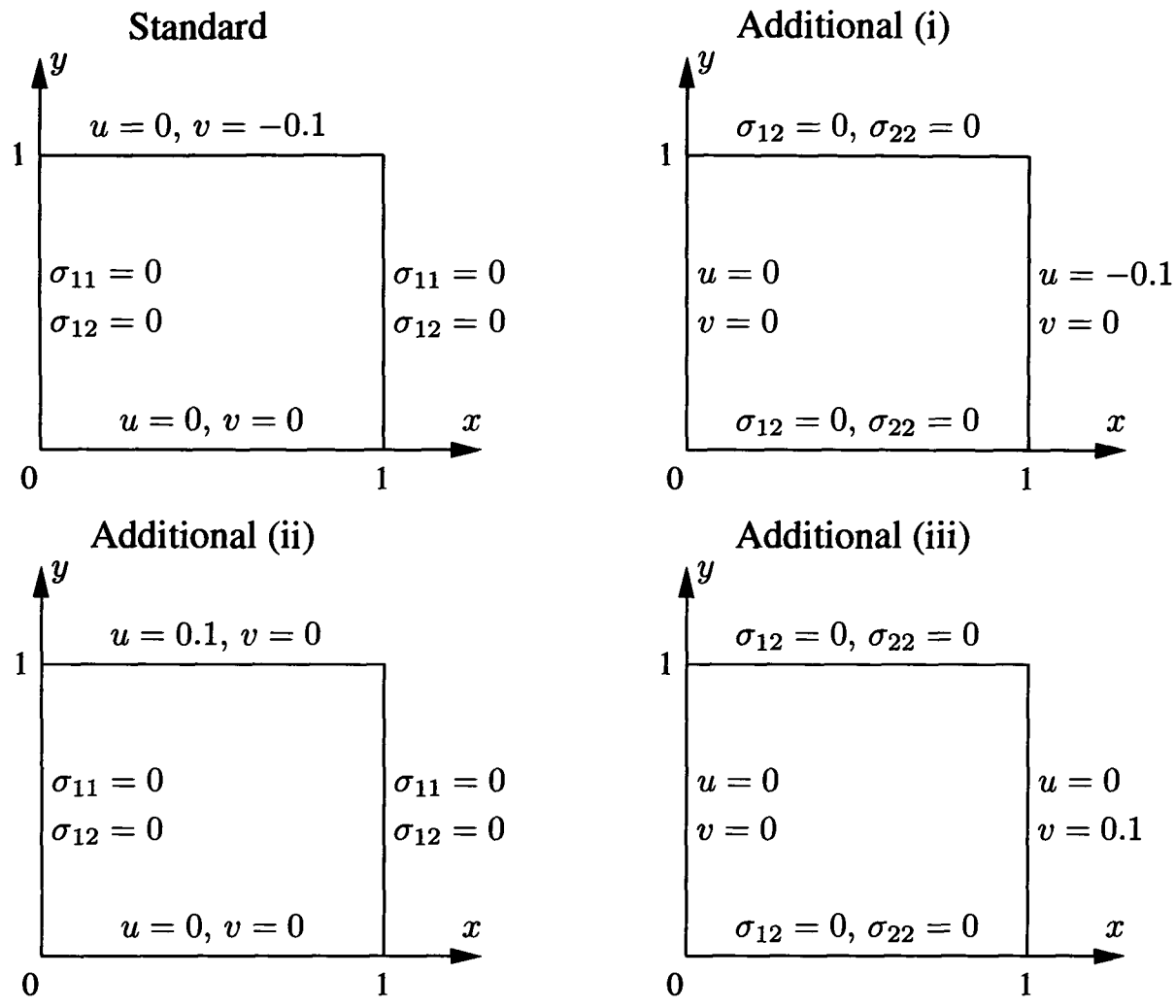


Figure 5.4: Schematic representation of the boundary conditions used in the generation of displacement fields by numerical solution of the linear elasticity forward problem. The boundary conditions used to generate the standard displacement field are shown on the top left. The other three sets of boundary conditions are used to generate linearly independent displacement fields to allow the inverse problem to be solved without boundary conditions on λ and μ.

5.2.1.1 Boundary Conditions

Solving the minimization problem posed in equation (4.123) using a single displacement field calculated from the forms of λ_{in} and μ_{in} and by specifying the values of the Lamé constants on the boundary gives rise to the solutions shown in Figure 5.6. It can be seen the solution is qualitatively similar to the input shown in Figure 5.5. The maximum values of both $\lambda(x, y)$ and $\mu(x, y)$ are both slightly larger for the numerical solution than for the input and there is a noticeable difference in the corners of the domain.

Figure 5.7 shows the spatial variation of the residuals R_1 and R_2 defined in equations (4.80) and (4.81). These plots quantify the qualitative observations made above. It is clear that throughout most of the domain the size of both of the residuals is less than 5×10^{-7}. There

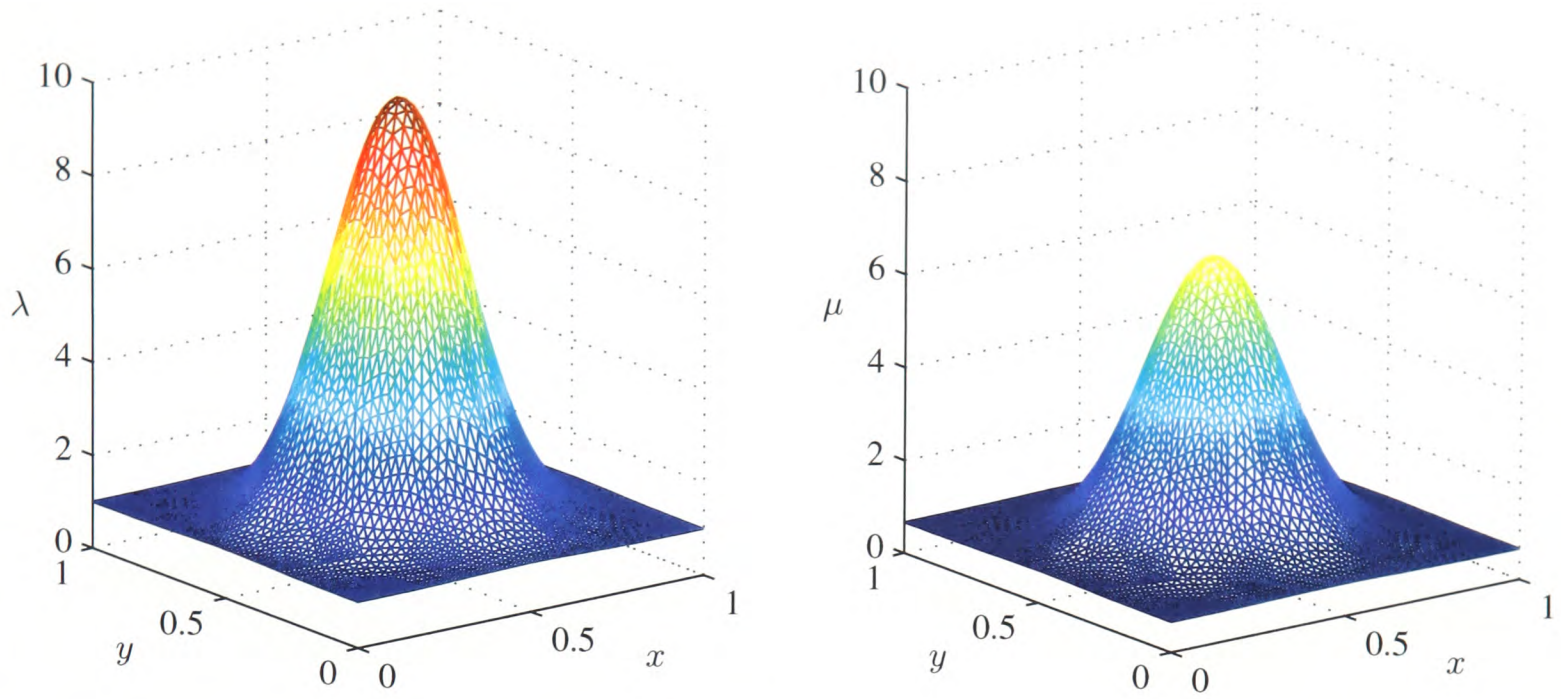


Figure 5.5: Spatial variation of the specified Lamé constants for testing the numerical solution of the linear elasticity inverse problem with smooth λ, μ. Here $\lambda_{\text{in}}(x, y) = 1 + 9\exp[-20((x-0.5)^2 + (y-0.5)^2)]$, $\mu_{\text{in}}(x, y) = 2\lambda_{\text{in}}(x, y)/3$.

are much larger errors at the corners and also, to a lesser extent, in the case of R_1 along the two edges $x = 0, 1$ (the errors along the edges extend only one element into the domain). In general R_2 is larger than R_1 and there are some areas where $|R_2| > 5 \times 10^{-7}$. However, away from the corners $|R_2| > 10^{-6}$ at only a single mesh point.

5.2.1.2 Multiple Displacement Fields

In Figure 5.8 the results of solving the inverse problem using four displacement fields generated using λ_{in}, μ_{in} as defined in equations (5.24) and (5.25) are shown. Recall that the solutions are constrained so that $\int_0^1 \int_0^1 \lambda dxdy = 1$, so the scale is not the same as that for λ_{in}, μ_{in}. The most obvious property of the solutions is that the surface is curved near to the edges of the domain, especially close to the corners. Away from the edges, however, the solutions appear to have the same shape as λ_{in}, μ_{in}, although the ratio of their heights is less than the $\frac{2}{3}$ used for the input parameters.

These observations are supported by the size of the residuals, which are plotted in Figure 5.9. Again there are relatively large errors near the boundaries and, especially, the corners — and these errors are larger than their equivalents in the solution using boundary data. This is to be expected as the errors from the numerical integration used in each of the

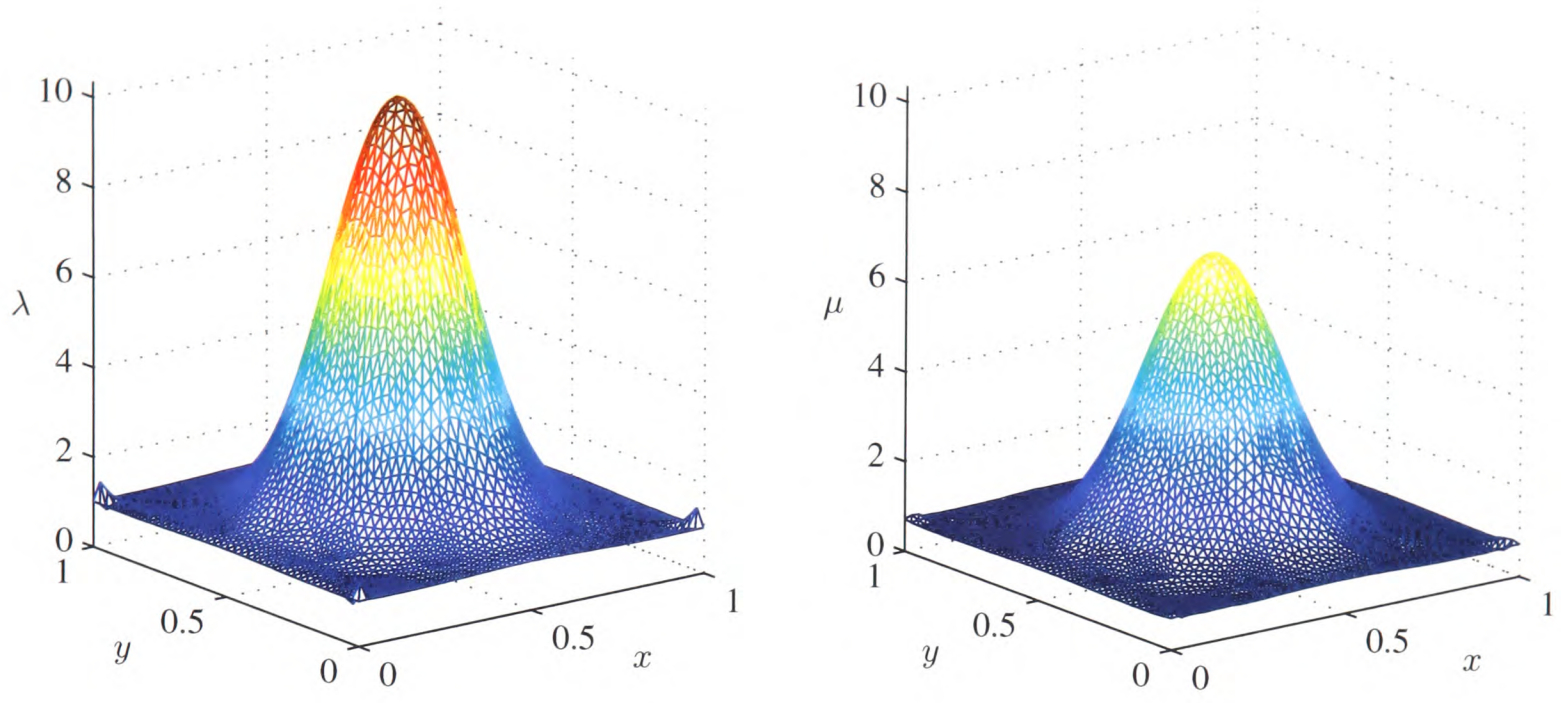


Figure 5.6: Numerical solution of the inverse problem using a displacement field generated from a single finite element solution of the forward problem with Lamé constants as shown in Figure 5.5. Dirichlet boundary conditions $\lambda = \lambda_{\text{in}}$, $\mu = \mu_{\text{in}}$ are applied on the boundary, where λ_{in} and μ_{in} are defined by equations (5.24) and (5.25) respectively.

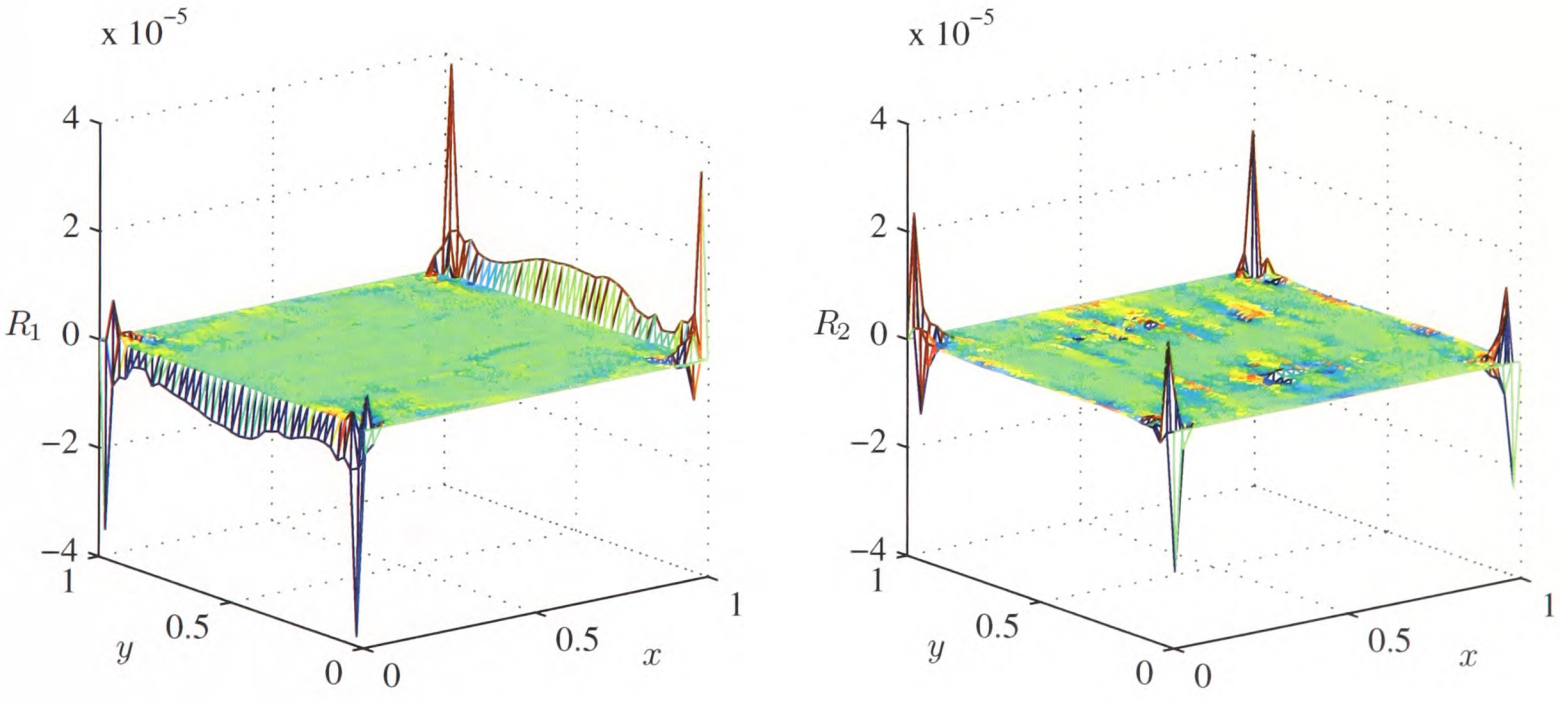


Figure 5.7: Residuals resulting from the numerically determined Lamé constants shown in Figure 5.6. The colour scale runs from -5×10^{-7} to -5×10^{-7}, so for any region which is not coloured dark red or dark blue $|R_i| < 5 \times 10^{-7}$.

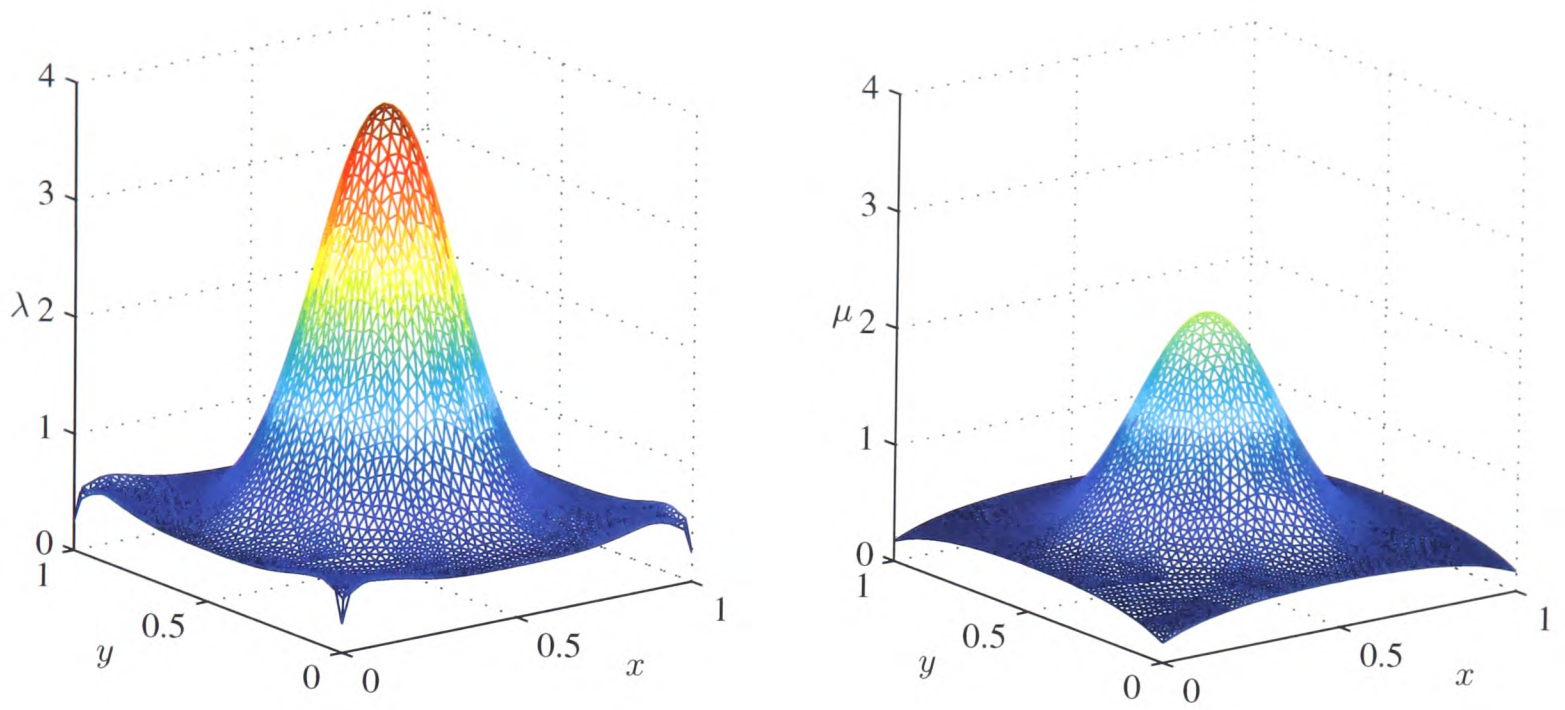


Figure 5.8: Numerical solution of the inverse problem using displacement fields generated from four finite element solutions of the forward problem with Lamé constants as shown in Figure 5.5.

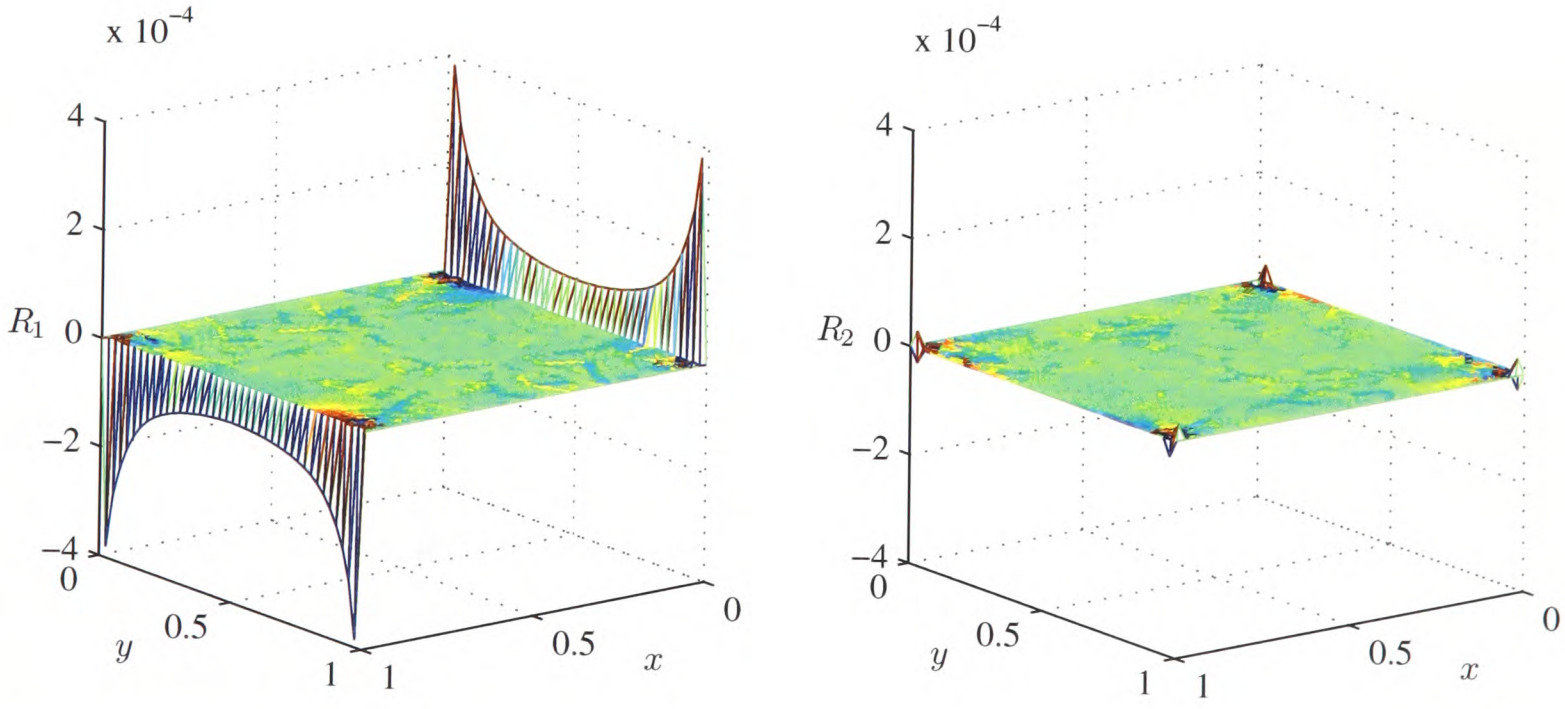


Figure 5.9: Residuals resulting from the numerically determined Lamé constants shown in Figure 5.6. The colour scale runs from -5×10^{-7} to -5×10^{-7}, so for any region which is not coloured dark red or dark blue $|R_i| < 5 \times 10^{-7}$.

forward problems may add up, whereas there is only one set of errors to consider when boundary data is used (plus the correct values of λ and μ are specified on the boundary). In the interior of the region the size of the residuals is comparable to those seen when boundary data was used in the previous section. In this case, there are no points away from the corner where either residual has size greater than 5×10^{-7}.

5.2.2 Lamé Constants With Discontinuous Derivative

Next, the Lamé constants used in the solutions of the forward problems are allowed to have discontinuous derivatives. The displacement fields are generated with

$$\lambda_{\text{in}}(x,y) = \begin{cases} 3, & \text{if } r < 0.15, \\ 1 + \dfrac{2(r-0.25)}{0.15-0.25}, & \text{if } 0.15 \leqslant r < 0.25, \\ 1, & \text{if } r \geqslant 0.25, \end{cases} \tag{5.26}$$

$$\mu_{\text{in}}(x,y) = \begin{cases} 2, & \text{if } r < 0.15, \\ \dfrac{2}{3} + \dfrac{4(r-0.25)}{3(0.15-0.25)}, & \text{if } 0.15 \leqslant r < 0.25, \\ \frac{2}{3}, & \text{if } r \geqslant 0.25, \end{cases} \tag{5.27}$$

where $r^2 = (x-0.5)^2 + (y-0.5)^2$. The spatial variation of these functions is shown in Figure 5.10.

5.2.2.1 Boundary Conditions

Figure 5.11 shows the results of solving the inverse problem with applied boundary conditions and a single displacement field generated with Lamé constants defined by equations (5.26) and (5.27). The corresponding residuals are shown in Figure 5.12. As in the case of the infinitely differentiable Lamé constants, qualitatively the solution looks close to the input parameters shown in Figure 5.10. The ratio of the values of λ and μ at any given point is $\frac{2}{3}$ as expected. Again there are errors in the corners and along the edges of the domain. Also, the flat surfaces in the input are not totally flat in the computed solution. Analysis of the residuals also suggests that the solution is less accurate in the sloped region than in the rest of the domain.

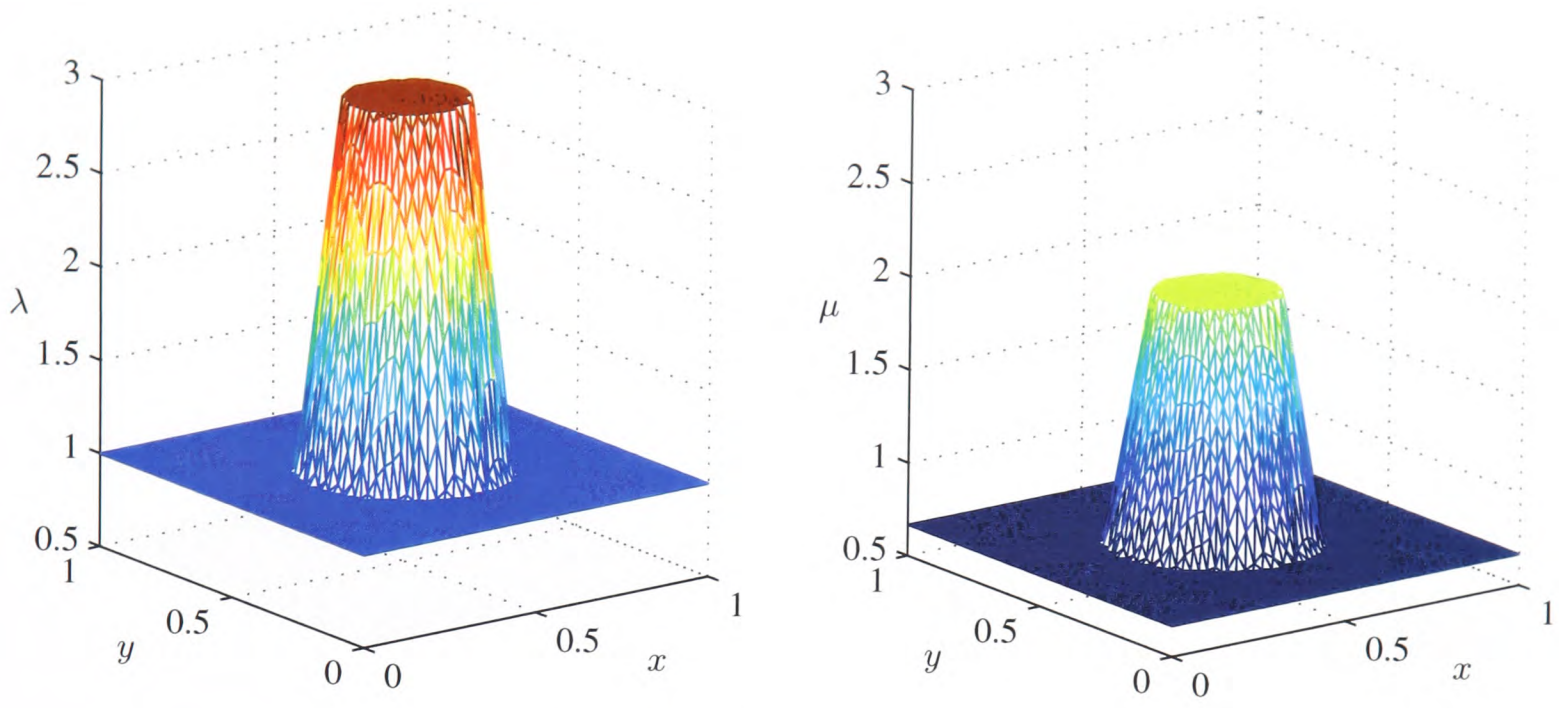


Figure 5.10: Spatial variation of the specified Lamé constants for testing the numerical solution of the linear elasticity inverse problem with λ, μ having a discontinuous first derivative. Here $\lambda(x, y) = 1$ if $r > 0.25$, $\lambda(x, y) = 3$ if $r < 0.15$, $\lambda(x, y) = 1 + 2(r - 0.25)/(0.15 - 0.25)$ otherwise, where $r^2 = (x - 0.5)^2 - (y - 0.5)^2$. $\mu(x, y) = 2\lambda(x, y)/3$ everywhere.

5.2.2.2 Multiple Displacement Fields

The results of solving the inverse problem using four displacement fields generated with λ_{in} and μ_{in} as defined in equations (5.26) and (5.27) are shown in Figure 5.13, with the corresponding residuals shown in Figure 5.14. Again the results are broadly similar to those in the case of the smoothly varying Lamé constants, with a large amount of curvature close to the corners of the domain, but good agreement with the input data otherwise. The flat surfaces are much more accurately represented in this case than they were by the solution with boundary conditions. The size of the residuals is less than 5×10^{-7} away from the boundary of the domain and the discontinuities in the derivatives of λ_{in} and μ_{in}. Again, the size of the residuals in these areas is larger than it was in the case where boundary data was given, probably due to the accumulation of numerical errors in the solution of four forward problems.

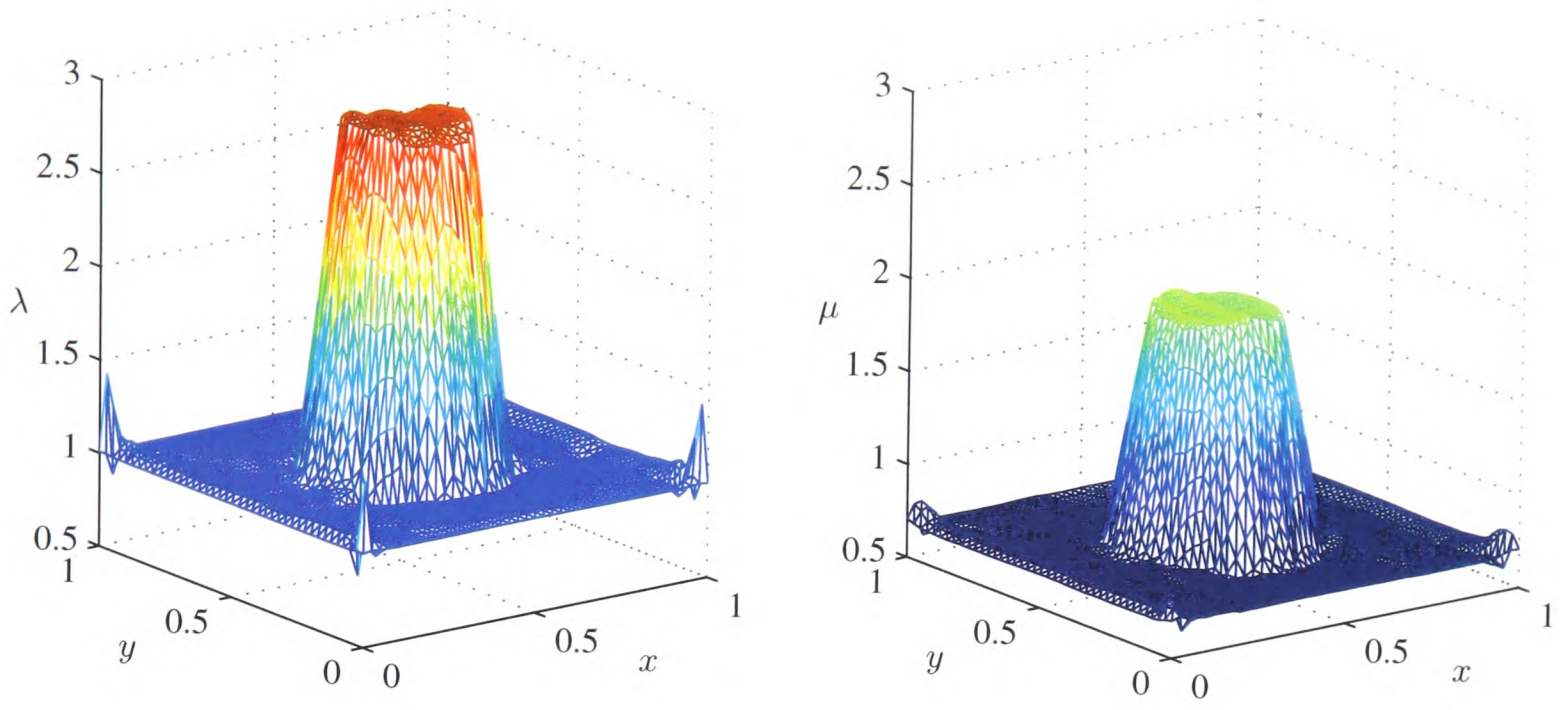


Figure 5.11: Numerical solution of the inverse problem using a displacement field generated from a single finite element solution of the forward problem with Lamé constants as shown in Figure 5.10. Dirichlet boundary conditions $\lambda = 1$, $\mu = \frac{2}{3}$ are applied on the boundary.

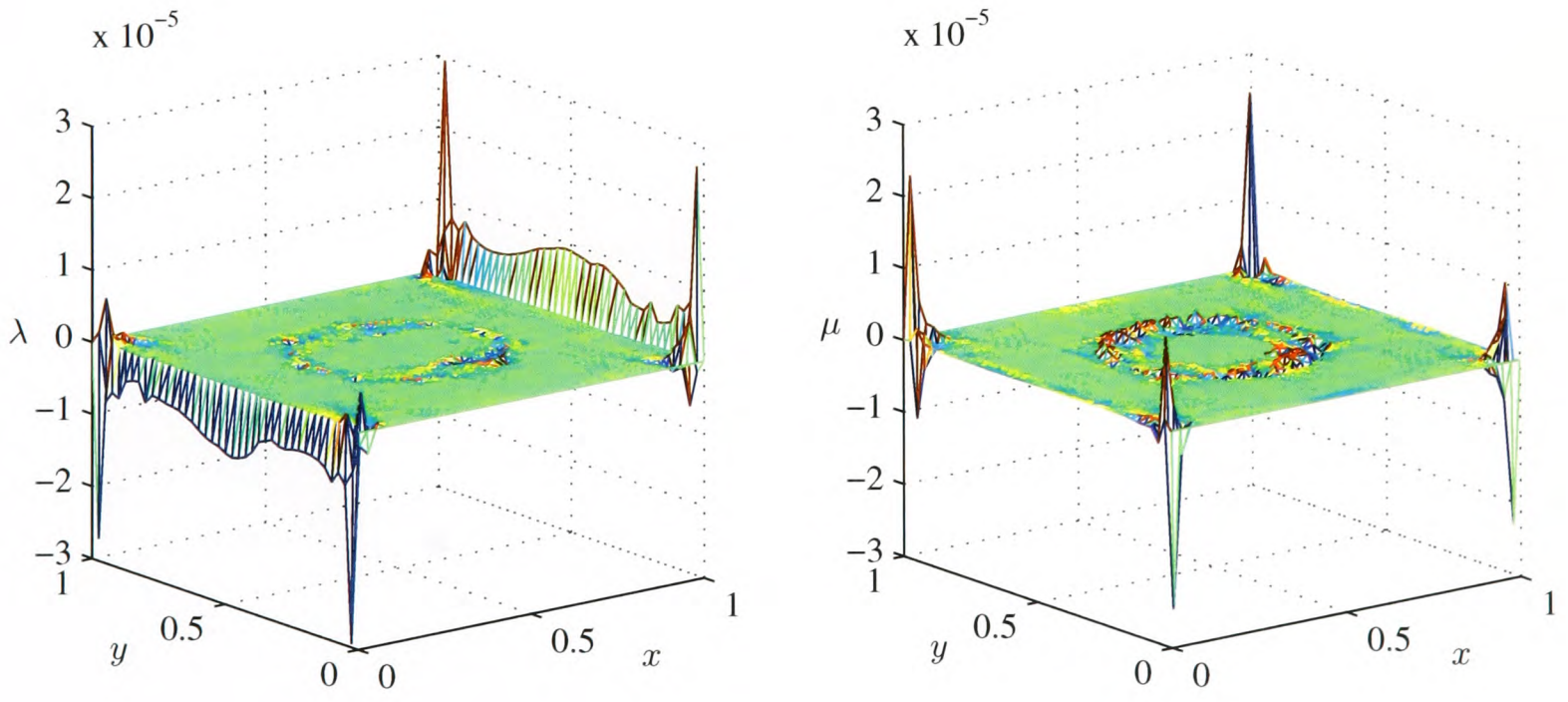


Figure 5.12: Residuals resulting from the numerically determined Lamé constants shown in Figure 5.11. The colour scale runs from -5×10^{-7} to -5×10^{-7}, so for any region which is not coloured dark red or dark blue $|R_i| < 5 \times 10^{-7}$.

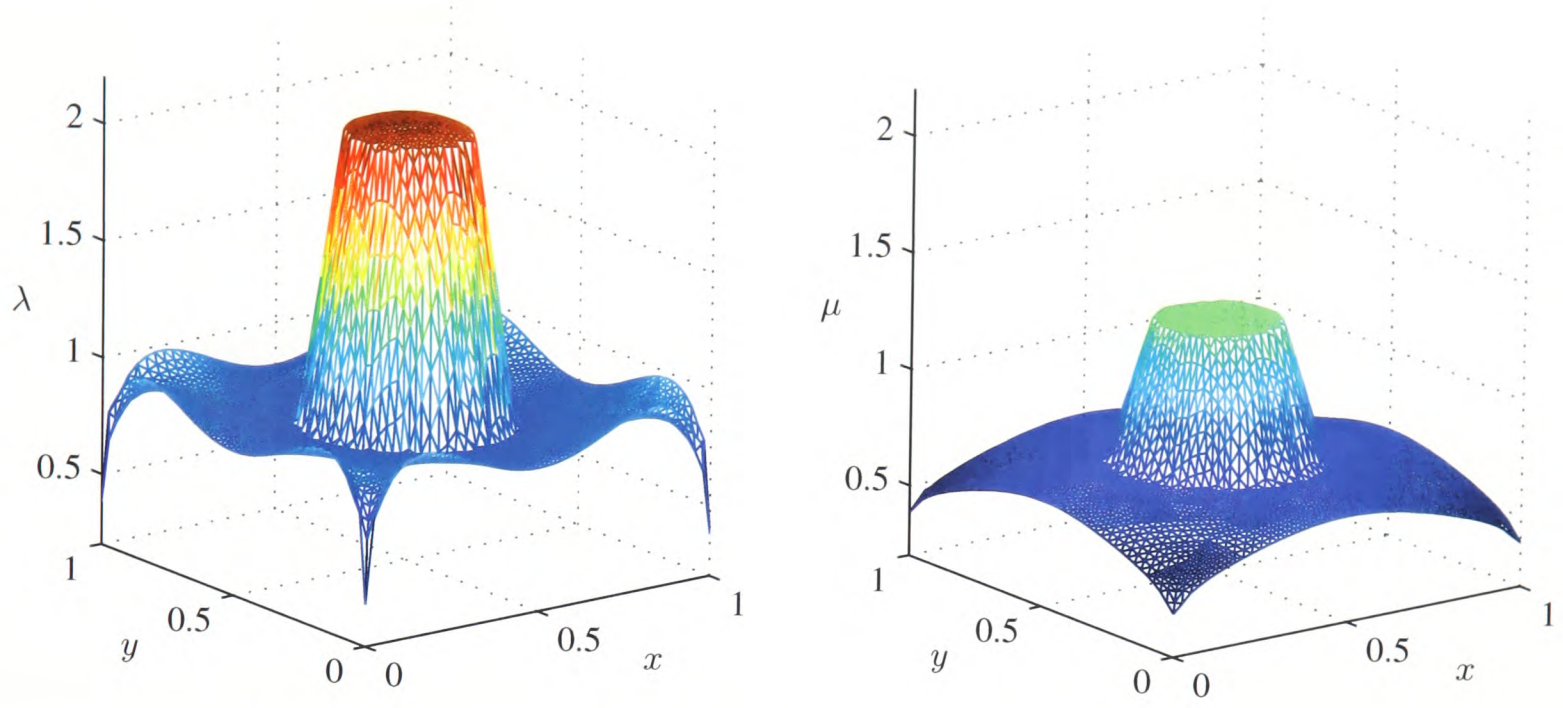


Figure 5.13: Numerical solution of the inverse problem using displacement fields generated from four finite element solutions of the forward problem with Lamé constants as shown in Figure 5.10.

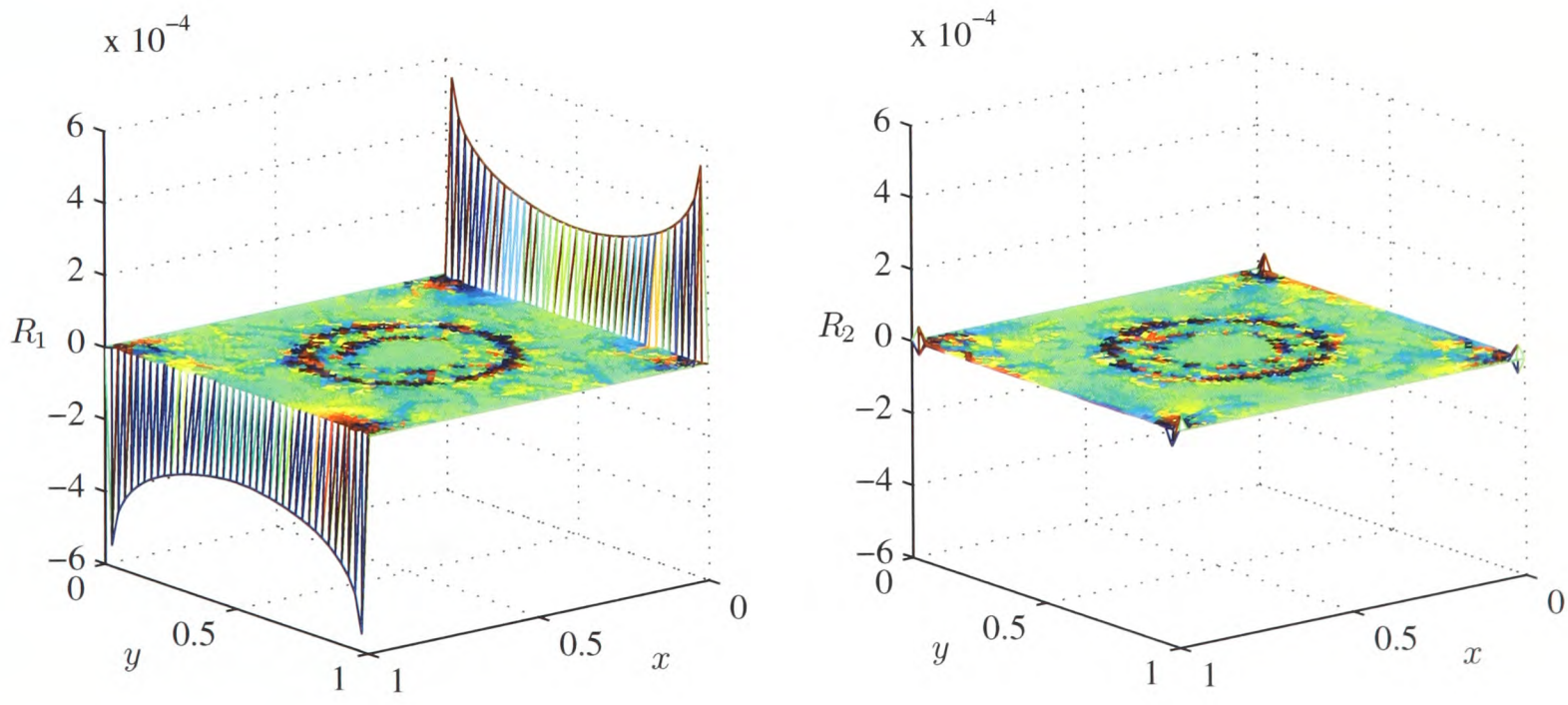


Figure 5.14: Residuals resulting from the numerically determined Lamé constants shown in Figure 5.11. The colour scale runs from -5×10^{-7} to -5×10^{-7}, so for any region which is not coloured dark red or dark blue $|R_i| < 5 \times 10^{-7}$.

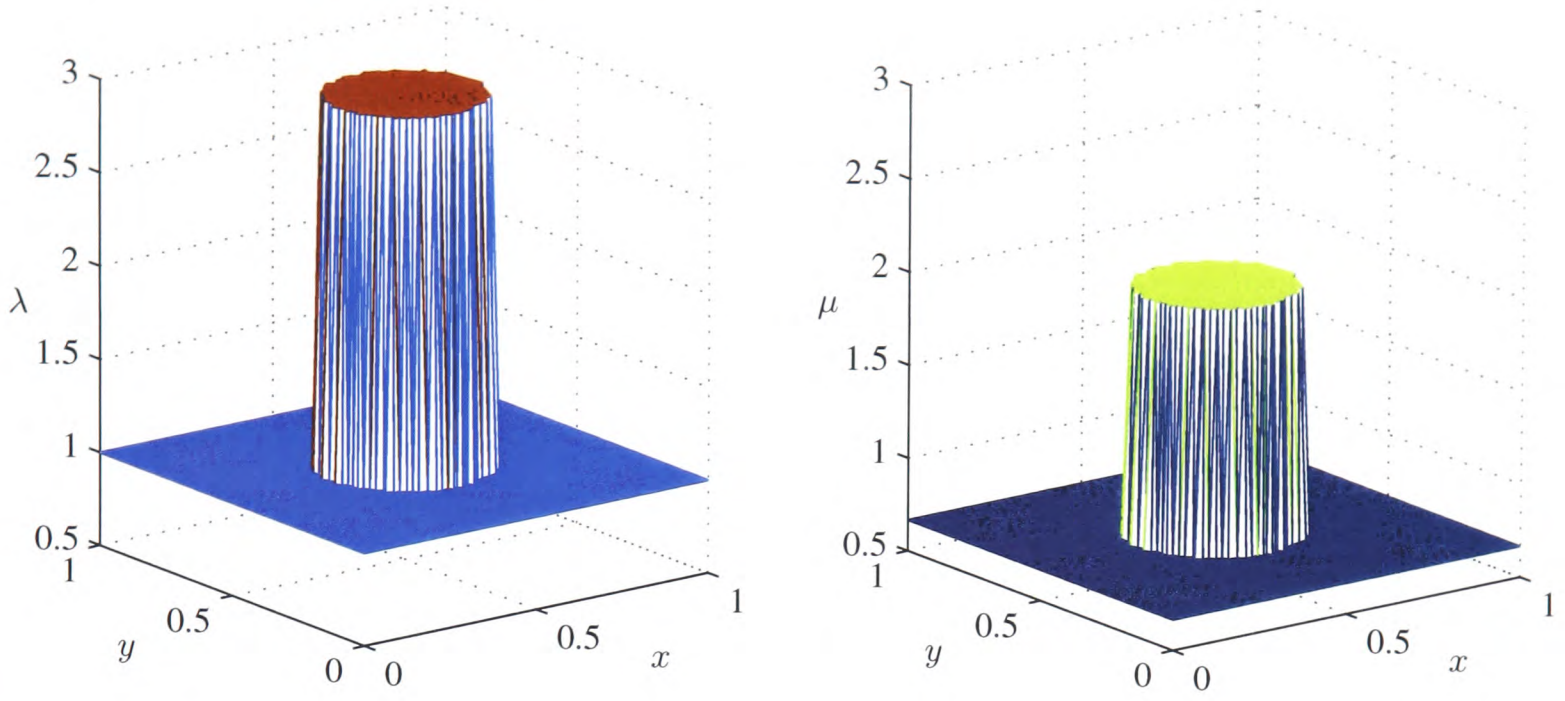


Figure 5.15: Spatial variation of the specified Lamé constants for testing the numerical solution of the linear elasticity inverse problem with discontinuous λ, μ. Here $\lambda(x, y) = 1$ if $r > 0.2$, $\lambda(x, y) = 3$ if $r < 0.2$, where $r^2 = (x - 0.5)^2 - (y - 0.5)^2$. $\mu(x, y) = 2\lambda(x, y)/3$ everywhere.

5.2.3 Discontinuous Lamé Constants

In this section one drawback of the parameter recovery method is shown: it is unable to deal with discontinuous Lamé constants. Displacement fields are generated with

$$\lambda_{\text{in}}(x, y) = \begin{cases} 3, & \text{if } r < 0.2, \\ 1, & \text{if } r \geqslant 0.2, \end{cases} \tag{5.28}$$

$$\mu_{\text{in}}(x, y) = \begin{cases} 2, & \text{if } r < 0.2, \\ \frac{2}{3}, & \text{if } r \geqslant 0.2, \end{cases} \tag{5.29}$$

where $r^2 = (x - 0.5)^2 + (y - 0.5)^2$. Figure 5.15 shows the spatial variation of these forms for λ_{in}, μ_{in}.

The results of trying to recover these input parameters using one displacement field and boundary conditions and using four displacement fields are shown in Figures 5.16 and 5.17 respectively. It is immediately obvious that these solutions are not a good match to the input data. Observe that the solutions are similar to those for Goodier's solution shown in Figure 5.3. This is encouraging since the distribution of Lamé constants given in

equations (5.28) and (5.29) corresponds to a stiff cylindrical inclusion,[3] so it suggests that the method fails in a way that it may be possible to predict and, hence, correct for. Also, the use of discontinuous finite element basis functions may provide a way of dealing with this problem. This is discussed in more depth in Section 8.2.1.

Recall that in the previous section similar input data was used, but there was a sloped region of width 0.1 (i.e. approximately five elements across) between the two regions with constant Lamé constants and the input parameters were successfully recovered. So, combining these two results, it is concluded that the method successfully recovers the input parameters so long as variations in the parameters occur on length scales larger than the element width. Practically this is not likely to be a limitation to the method since cancers do not usually have distinct, sharp boundaries[4] and, in any case, reconstruction could be performed over multiple scales. However, further investigation of this point is needed.

5.3 Stress Singularities at the Corners

Observe that in the results of the previous section the largest errors in the recovered solution occur at the corners of the domain. These errors are particularly apparent in the solutions where multiple displacement fields are used: in these cases the solutions are curved so that the values of both λ and μ approach zero at the corners. In this section this behaviour is explained by the presence of singularities in the values of the stress tensor at the corners. This is done first by observing the presence of singularities in the numerical solutions of the forward problem and then by referring to existing theory that accounts for their presence. The effects that the stress singularities have on the solution of the inverse problem are then discussed.

Figure 5.18 shows a typical stress field for a solution of the forward problem. It can be seen that all of the components of stress have singularities at each of the corners of the domain, i.e. the absolute values of the stress components tend to infinity as the distance from the corner approaches zero. Further, there appear to be oscillations in the stress components close to the corners. Note, by comparing with Figure 5.19 which shows the

[3]Note, however, that the displacement field used here does is not Goodier's solution since the boundary conditions applied do not match the displacements seen in Goodier's solution — in particular $u = 0$ on $y = \pm 1$.

[4]Note that the boundaries of cysts and between regions in gelatin phantoms used for testing have sharper boundaries so these may be harder to reconstruct accurately.

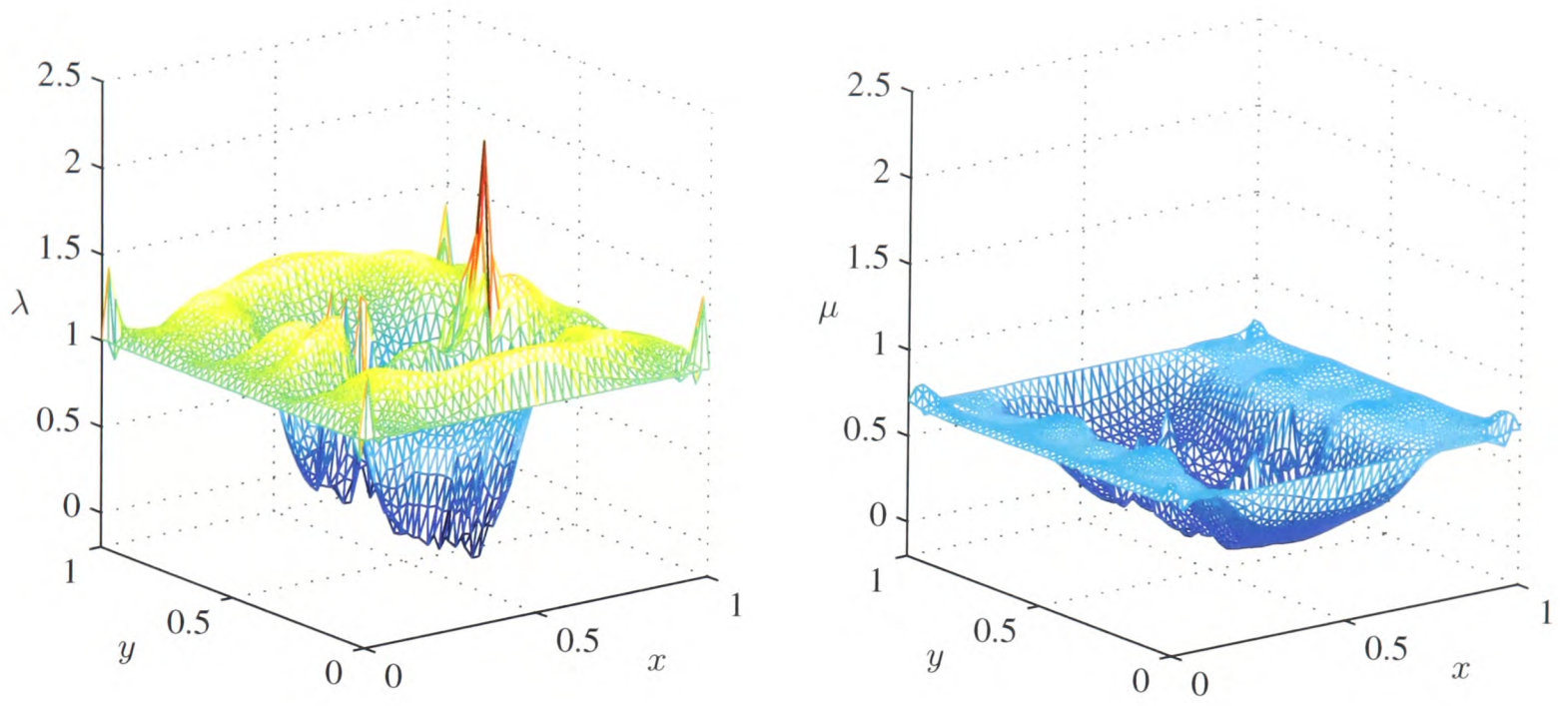


Figure 5.16: Numerical solution of the inverse problem using a displacement field generated from a single finite element solution of the forward problem with Lamé constants as shown in Figure 5.15.

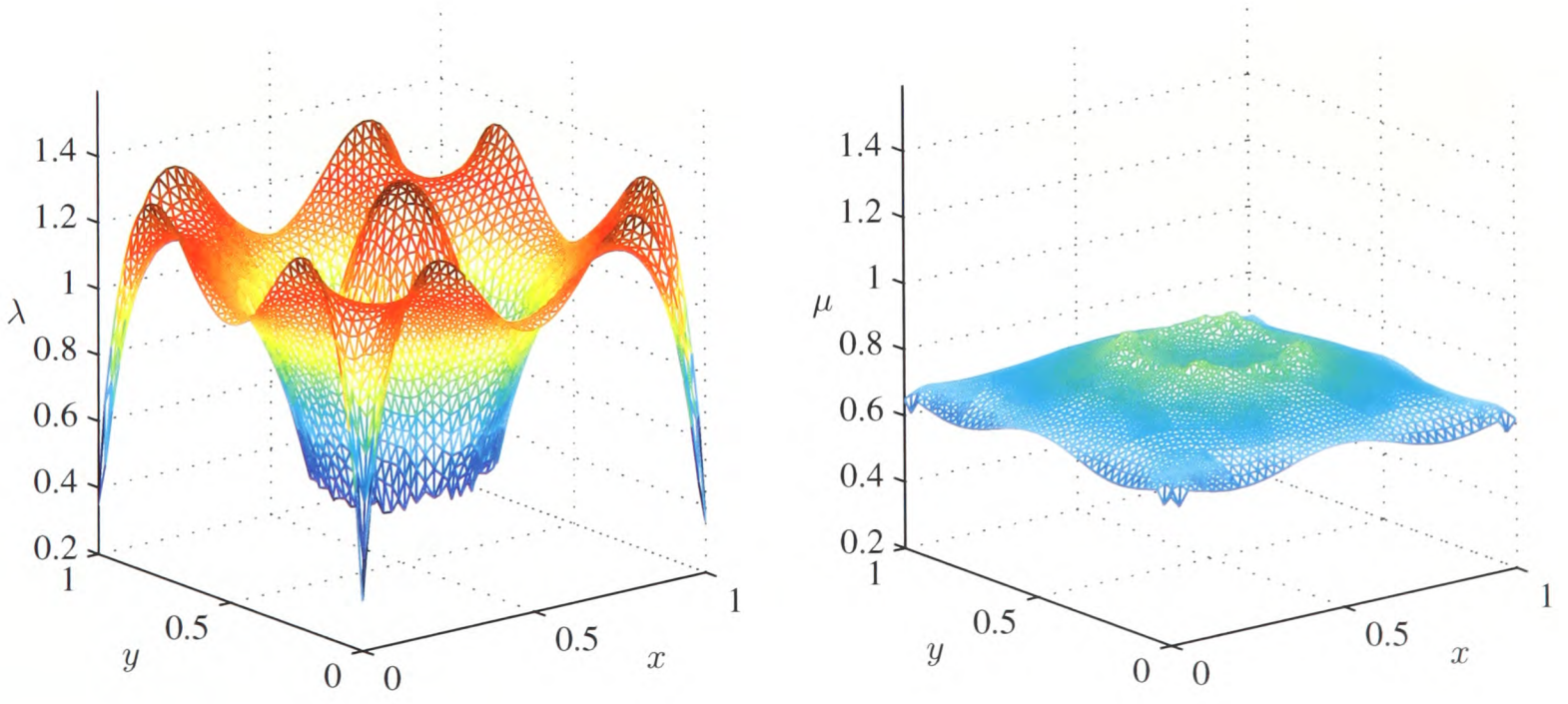


Figure 5.17: Numerical solution of the inverse problem using displacement fields generated from four finite element solutions of the forward problem with Lamé constants as shown in Figure 5.15.

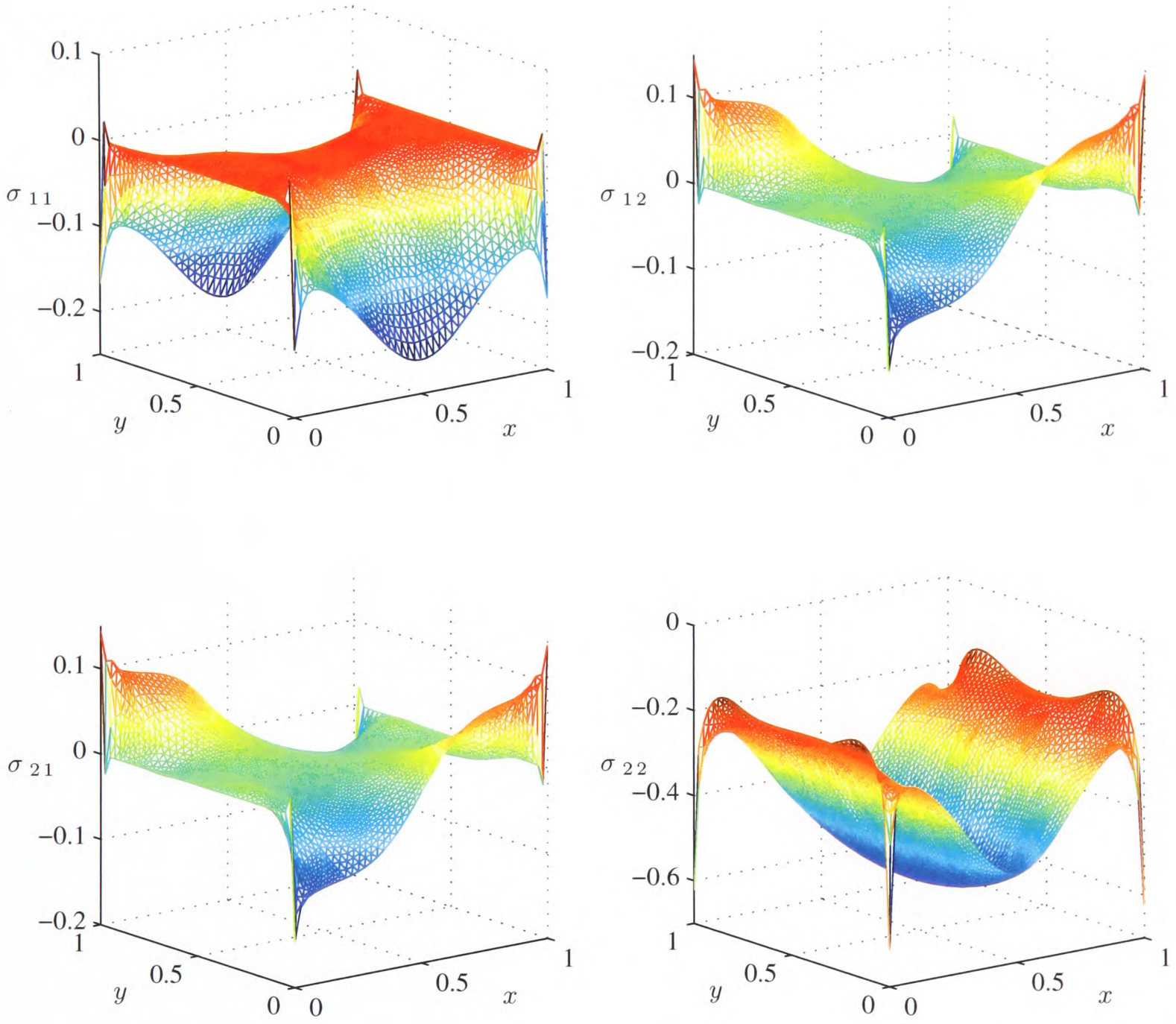


Figure 5.18: Stress fields resulting from solution of the linear elasticity forward problem on a mesh containing 5248 elements, with Lamé constants as shown in Figure 5.5, with zero stress boundary conditions on $x = 0, 1$, $u, v = 0$ on $y = 0$ and $u = 0$, $v = -0.1$ on $y = 0$.

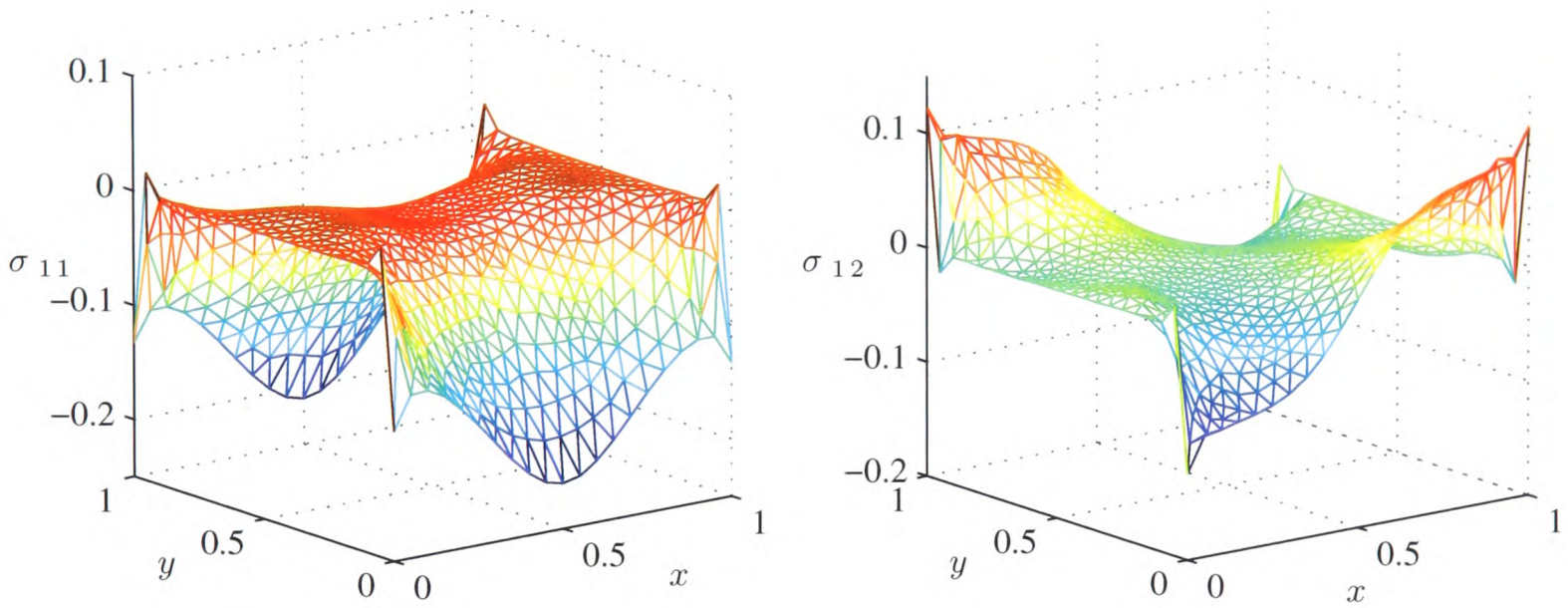


Figure 5.19: Stress fields resulting from solution of the linear elasticity forward problem on a mesh containing 1312 elements, with Lamé constants as shown in Figure 5.5, with zero stress boundary conditions on $x = 0, 1$, $u, v = 0$ on $y = 0$ and $u = 0$, $v = -0.1$ on $y = 0$.

same stress field on a coarser mesh, that the actual values of the stress components at the corners *increase* in magnitude as the mesh is refined (this is most obvious in the σ_{22} field, where the values at the corners are approximately -0.5 on the coarse mesh, but increase in magnitude to -0.6 on the finer mesh). This shows that the singularities are a feature of the exact solution and not a numerical artifact that will vanish as the mesh is progressively refined.

The existence of stress singularities in mixed boundary values problems for linear elasticity (i.e. those in which displacements are specified on parts of the boundary and tractions on other parts) is well known and described in detail in, e.g., England [23]. A summary of the theory can be found in Appendix C.2. The key result for this work is the observation that as the distance r from the corner of the domain tends to zero the limiting behaviour of the stress components is

$$\sigma_{12} = O\big(r^{-1/2}\sin(\beta \log r)\big), \tag{5.30}$$

$$\sigma_{22} = O\big(r^{-1/2}\cos(\beta \log r)\big). \tag{5.31}$$

So, the stress components oscillate as $r \to 0$ and their magnitudes tend to infinity. This agrees with the behaviour that was observed in Figure 5.18. So, the solutions of the forward problem used here are harder than those that would result from the real imaging problem (since the real displacement field would not contain infinite stresses).

If the stress components have singularities at the corners of the domain then, since strain is a linear function of stress, the components of strain must also have singularities at the corners. Hence, in order for the Euler-Lagrange equations (4.92) and (4.95) to be satisfied, the Lamé constants must vanish at the corners. Further, since $|\sigma| = O(r^{-1/2})$ as $r \to 0$, where r is the distance from a corner, the Lamé constants are $O(r^{1/2})$ as $r \to 0$. The shape of the curved surfaces in Figures 5.8 and 5.13 are consistent with this.

5.4 Robustness of the Method

Having shown that the parameter recovery method can be used on perfect input data the next step is to test how it deals with errors in the input. This is done first by adding small (normally distributed) errors to the displacement fields and investigating how this effects the recovered Lamé constants. This is intended to give an idea of how the method will

deal with errors in displacement estimation between ultrasound images. Second, the effect of prescribing incorrect boundary data is investigated. This reflects the difficulty there is likely to be in determining what boundary conditions to apply in real tissue.

5.4.1 Errors in Displacement Estimation

In this section the effect of errors in the ultrasound displacement estimation technique is investigated. This is done by adding normally distributed 'noise' to the displacement field generated by solving the forward problem of linear elasticity.[5] Suppose that a displacement field $(\mathbf{U}, \mathbf{V})$ is given where U_i, V_i, $i = 1, \ldots, N$, are the x- and y-displacements respectively of the i-th node of the finite element mesh. Suppose further that $R^{(i)}$, $i = 1, \ldots, 4$ are N-by-2 matrices of numbers that are randomly generated from a normal distribution with mean 0 and standard deviation 1. Then a 'noisy' displacement field $(\tilde{\mathbf{U}}, \tilde{\mathbf{V}})$ is defined by

$$\tilde{U}_i = U_i + \delta R^{(i)}_{i1}, \tag{5.32}$$

$$\tilde{V}_i = V_i + \delta R^{(i)}_{i2}, \tag{5.33}$$

where δ is a constant. Observe that the perturbation in the displacement is not proportional to the size of the displacement itself and is the same in both the x- and y-directions. These assumptions are unlikely to hold for a real ultrasound image, but represent the simplest type of additive noise. As before, the cases where one displacement field with boundary conditions and multiple displacement fields without boundary data are considered in turn.

5.4.1.1 Boundary Conditions

Figures 5.20 – 5.24 show the results of adding errors of varying magnitudes to a numerically calculated solution of the forward problem of linear elasticity, with (smooth) Lamé constants given by equations (5.24) and (5.25) and then solving the inverse problem numerically with the values of λ and μ on the boundary prescribed. As in Section 5.2, the solution of the forward problem is generated by applying boundary conditions $u, v = 0$ on $y = 0$, $u = 0$, $v = -0.1$ on $y = 1$ and zero stress on $x = 0, 1$. For this displacement field

[5]Note that this distribution of errors is not what would necessarily be expected from a displacement estimation algorithm — the exact distribution of errors depends on the algorithm used. One reason why a normal distribution of errors might not be expected is that attenuation of the ultrasound signal makes larger errors more likely further away from the probe.

the mean absolute value of the displacement the x-direction is approximately 6×10^{-3}. For the displacements in the y-direction the corresponding value is 5×10^{-2}. The magnitude of the noise will be considered in relation to these typical displacements (although it must be remembered that in some regions of the domain the actual displacements are much smaller than those given here).

The solution of the inverse problem for a noisy displacement with $\delta = 5 \times 10^{-6}$ (i.e. the typical size of the noise is 0.1% that of the typical displacements) is shown in Figure 5.20. Figure 5.21 shows the residuals calculated from equations (4.80) and (4.81) using the solution of the inverse problem and the displacement fields with no noise added. Comparison of Figures 5.5 and 5.20 shows a good qualitative agreement with the actual values of the Lamé constants throughout the numerical domain. The residuals can then be used to quantify the errors in the solution. It can be seen in Figure 5.21 that the largest residuals are $O\left(10^{-6}\right)$. These are larger by a factor of 10 than the corresponding residuals with no noise (shown in Figure 5.7). However, near to the boundary (where the value of the Lamé constants is given) the residuals are much smaller — approximately the same size as those for the solution with no noise.

The effect of increasing the value of δ is shown in Figures 5.22 – 5.24. The solution when $\delta = 5 \times 10^{-5}$ (10 times that used above and 1% of a typical displacement) is shown in Figure 5.24. It can be seen that the parameter recovery method has completely broken down with noise of this magnitude — the values of the Lamé constants are effectively random fluctuations about some mean value. Figure 5.22 shows the solution of the inverse problem for $\delta = 10^{-5}$. This is an intermediate step between $\delta = 5 \times 10^{-6}$, where the method works well, and $\delta = 5 \times 10^{-5}$, where it does not work at all. Here the solutions for the Lamé constants appear to be the same shape as those used to generate the displacement field, but the size of μ in the centre of the domain is much smaller in comparison to that of λ than it should be. This may be a result of the noise having a larger effect on the x-displacements than the y-displacements, since they are smaller by a factor of 10 (see above): μ is the shear modulus, so it will be more drastically affected by large relative changes in the x-displacements than λ will be (recall that the material is being compressed in the y-direction). The residuals associated with this solution are shown in Figure 5.23. They show a similar pattern to those in Figure 5.21 — being comparable to those of the exact solution near to the boundary but much larger in the interior of the domain. The size of the residuals increases for larger values of δ as would be expected.

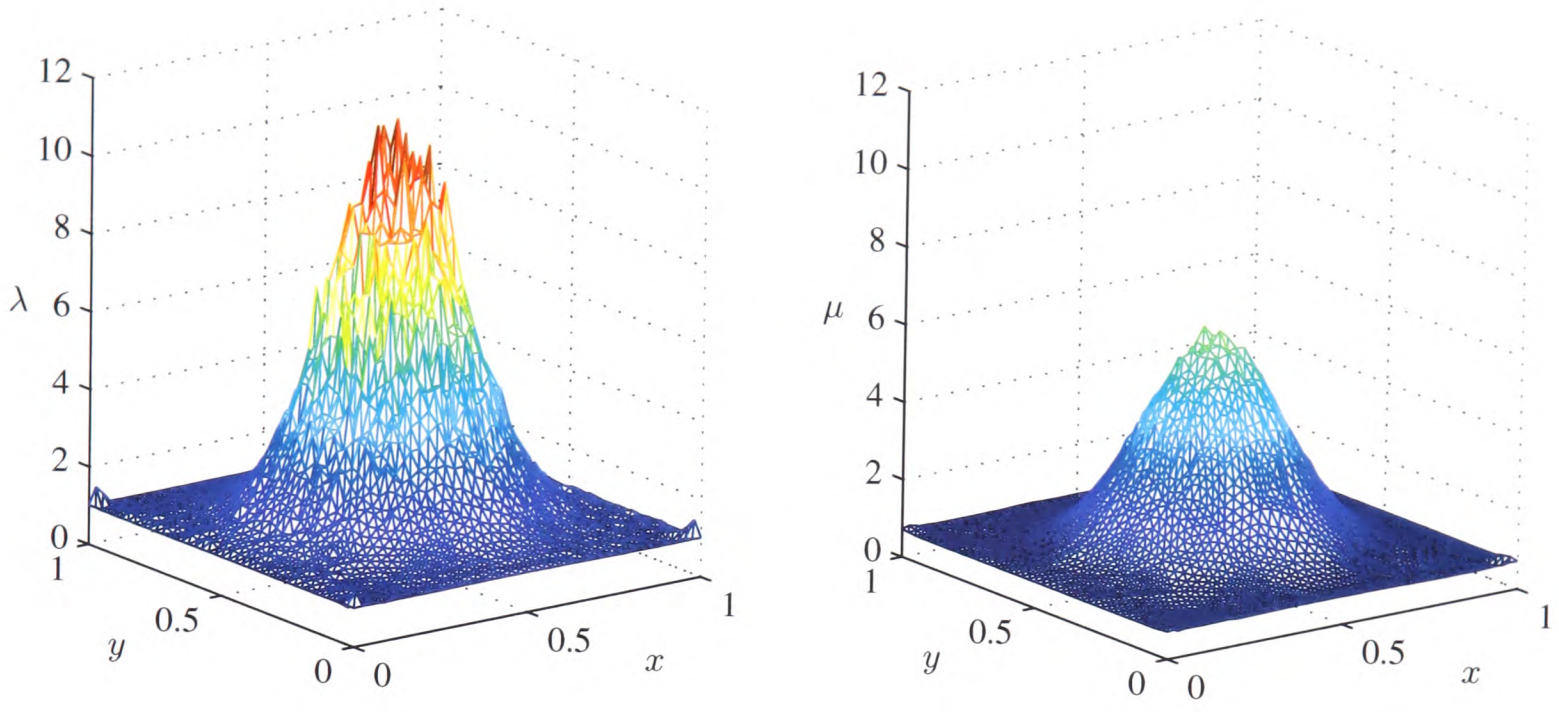


Figure 5.20: Numerical solution of the inverse problem using a displacement field generated from a single finite element solution of the forward problem with Lamé constants as shown in Figure 5.5 and with normally distributed noise with mean zero and variance 5×10^{-6}. Dirichlet boundary conditions $\lambda = \lambda_{\text{in}}$, $\mu = \mu_{\text{in}}$ are applied on the boundary, where λ_{in} and μ_{in} are defined by equations (5.24) and (5.25) respectively.

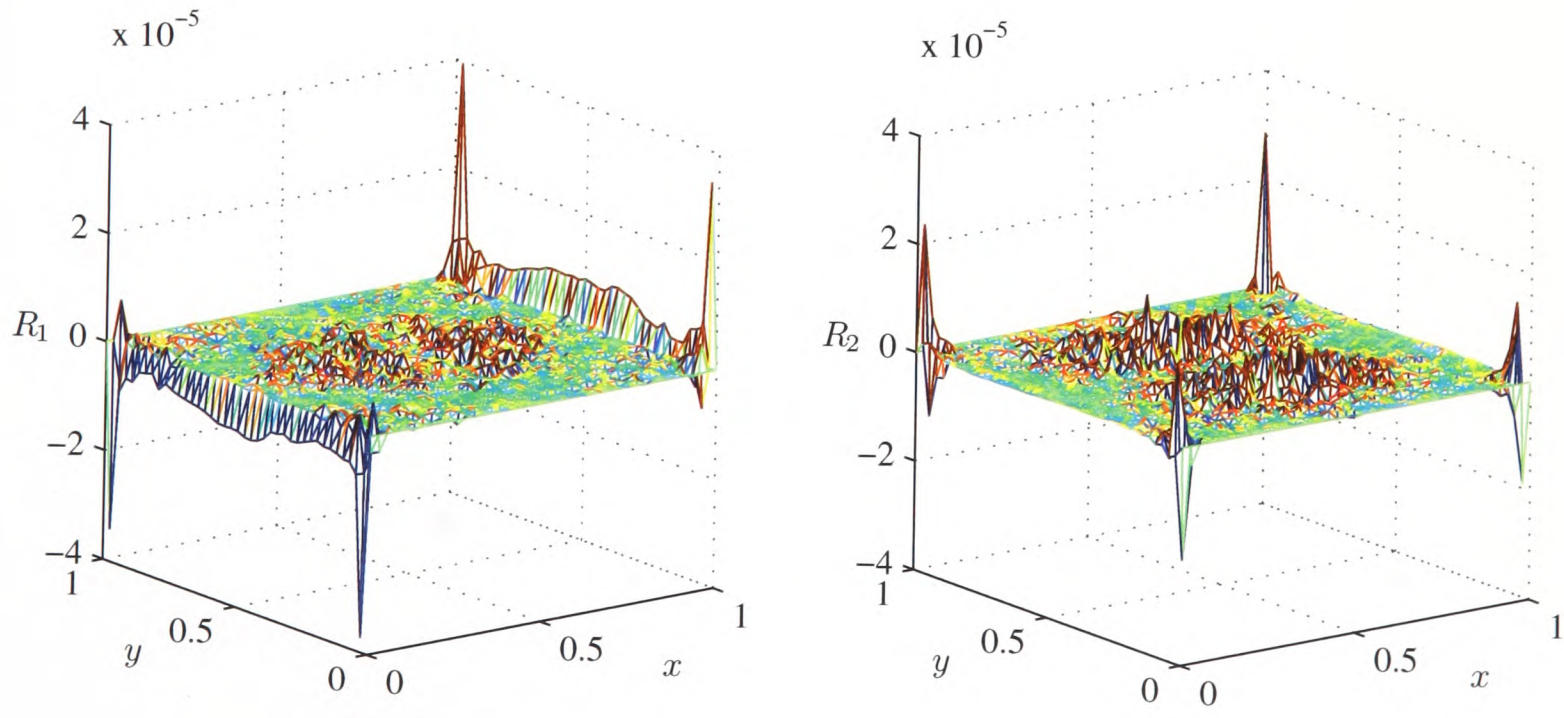


Figure 5.21: Residuals resulting from the numerically determined Lamé constants shown in Figure 5.20. The colour scale runs from -10^{-6} to 10^{-6}, so for any region which is not coloured dark red or dark blue $|R_i| < 10^{-6}$.

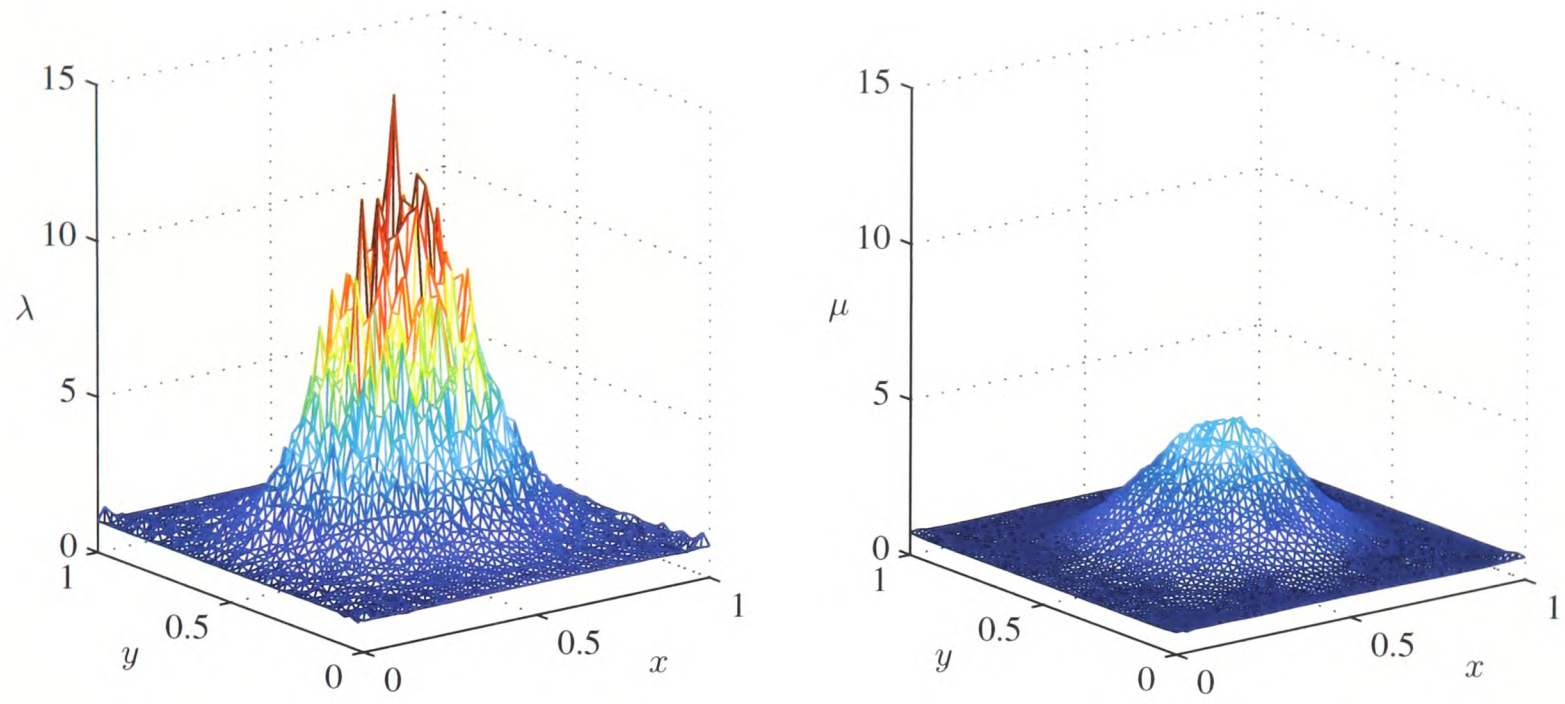


Figure 5.22: Numerical solution of the inverse problem using a displacement field generated from a single finite element solution of the forward problem with Lamé constants as shown in Figure 5.5 and with normally distributed noise with mean zero and variance 10^{-5}. Dirichlet boundary conditions $\lambda = \lambda_{\text{in}}$, $\mu = \mu_{\text{in}}$ are applied on the boundary, where λ_{in} and μ_{in} are defined by equations (5.24) and (5.25) respectively.

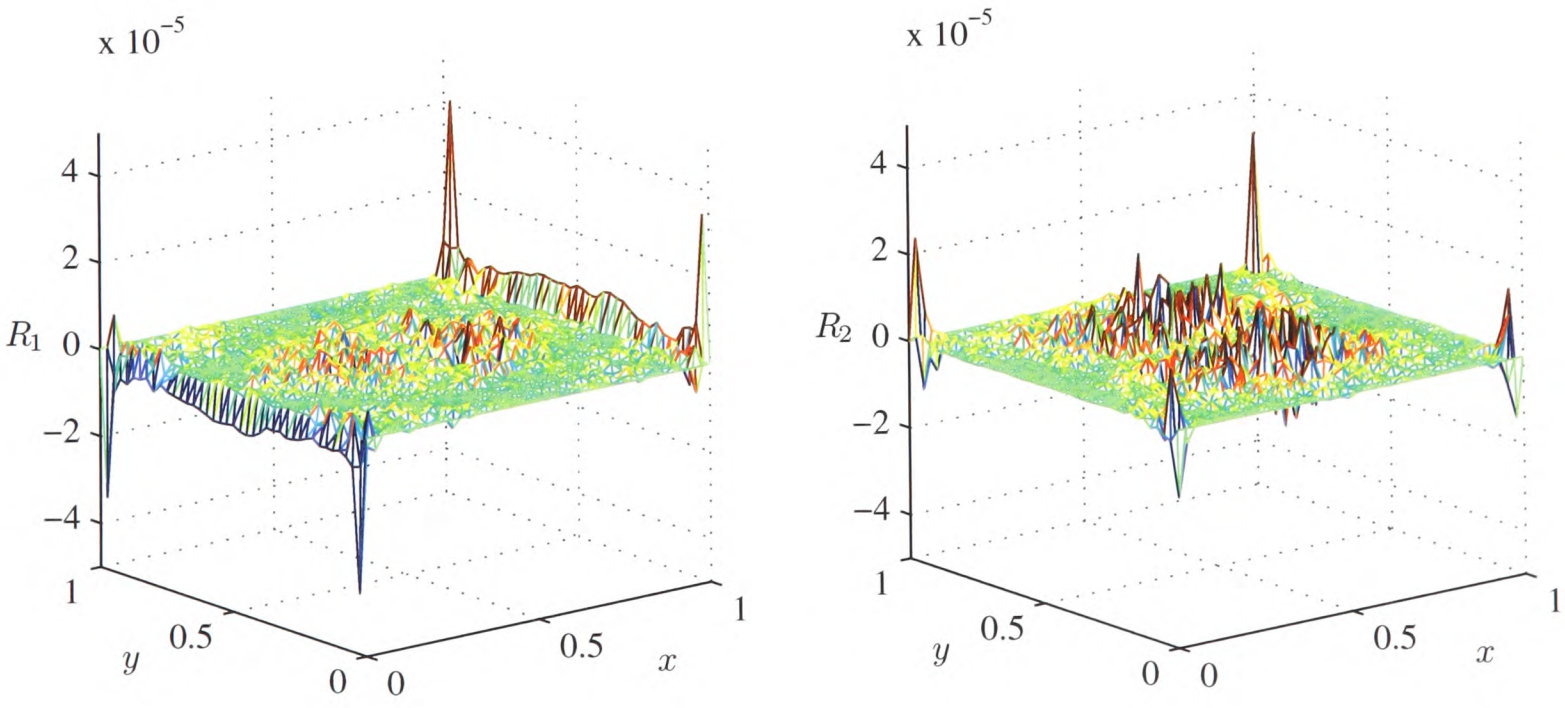


Figure 5.23: Residuals resulting from the numerically determined Lamé constants shown in Figure 5.22. The colour scale runs from -5×10^{-6} to 5×10^{-6}, so for any region which is not coloured dark red or dark blue $|R_i| < 5 \times 10^{-6}$.

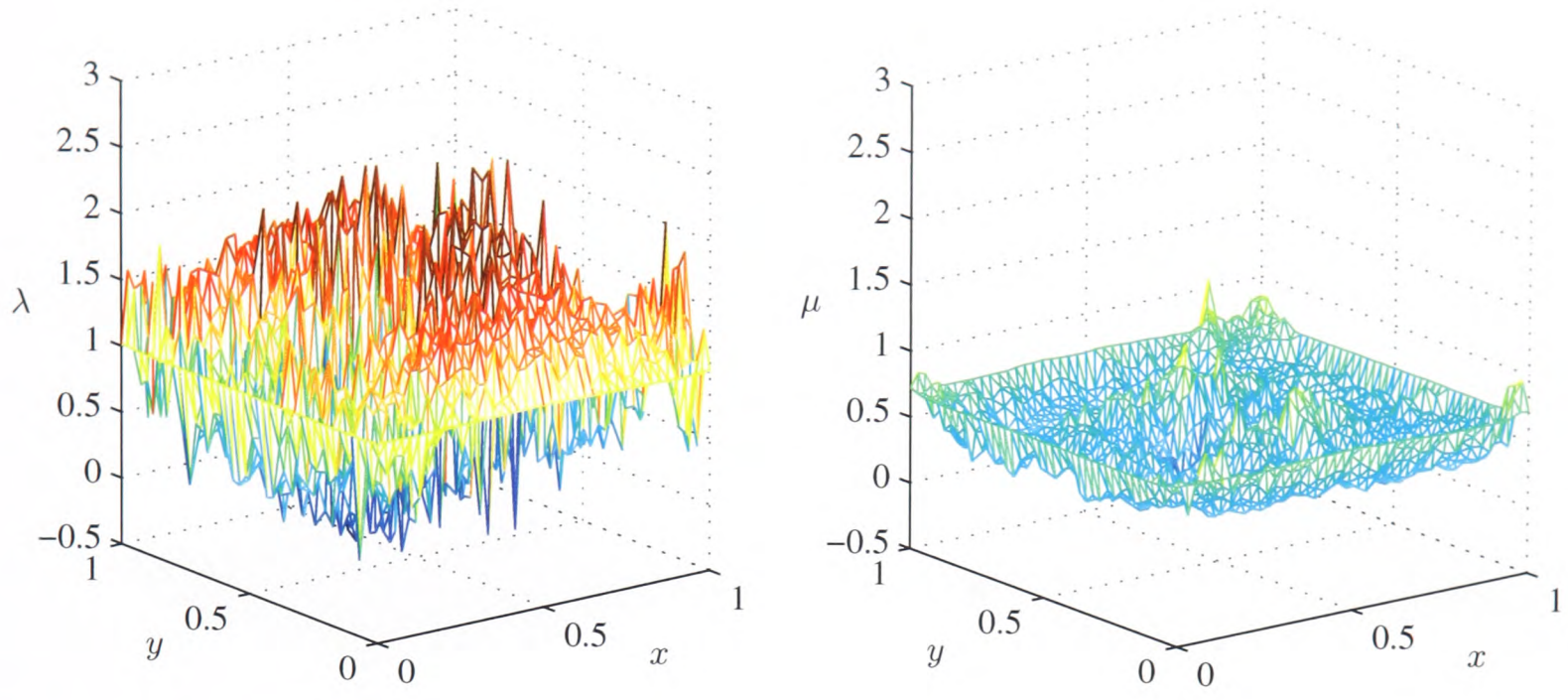


Figure 5.24: Numerical solution of the inverse problem using a displacement field generated from a single finite element solution of the forward problem with Lamé constants as shown in Figure 5.5 and with normally distributed noise with mean zero and variance 5×10^{-5}. Dirichlet boundary conditions $\lambda = \lambda_{\text{in}}$, $\mu = \mu_{\text{in}}$ are applied on the boundary, where λ_{in} and μ_{in} are defined by equations (5.24) and (5.25) respectively.

5.4.1.2 Multiple Displacement Fields

The results of solving the inverse problem with four displacement fields, each with normally distributed noise (independent, but from the same distribution) added, but no boundary conditions are shown in Figures 5.25 – 5.31. The four displacement fields used are the same as those described in Section 5.2 and the Lamé constants of the material are described by equations (5.24) and (5.25), i.e. they vary smoothly over the domain. Recall from Section 5.4.1.1 that for the fields generated by compression the typical size of the displacements are 5×10^{-2} in the direction of the compression and 6×10^{-3} perpendicular to it. For the final two displacement fields, generated by shearing the domain in either the x- or y-direction, the mean displacements are 5×10^{-2} in the direction of the shear and 8×10^{-3} perpendicular to it. So, the ratio of δ to the typical displacement size is approximately the same for all four fields.

The behaviour of the parameter recovery method as the amount of noise increases is very similar to that which was observed for a single displacement field. Figure 5.25 shows the solution of the inverse problem when $\delta = 10^{-6}$. These solutions show good agreement with the input data shown in Figure 5.5 (recall that for multiple displacement fields the

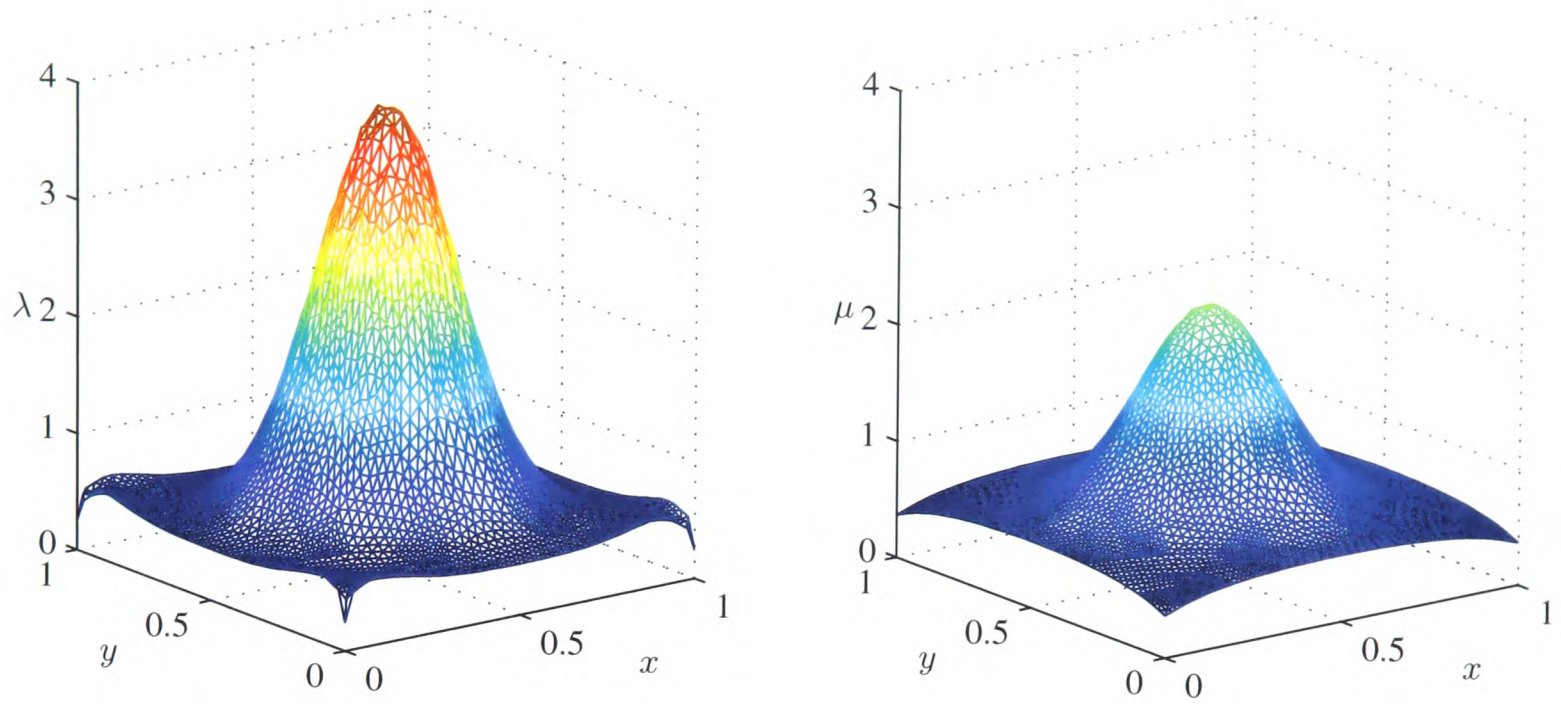


Figure 5.25: Numerical solution of the inverse problem using a displacement field generated from a single finite element solution of the forward problem with Lamé constants as shown in Figure 5.5 and with normally distributed noise with mean zero and variance 10^{-6}.

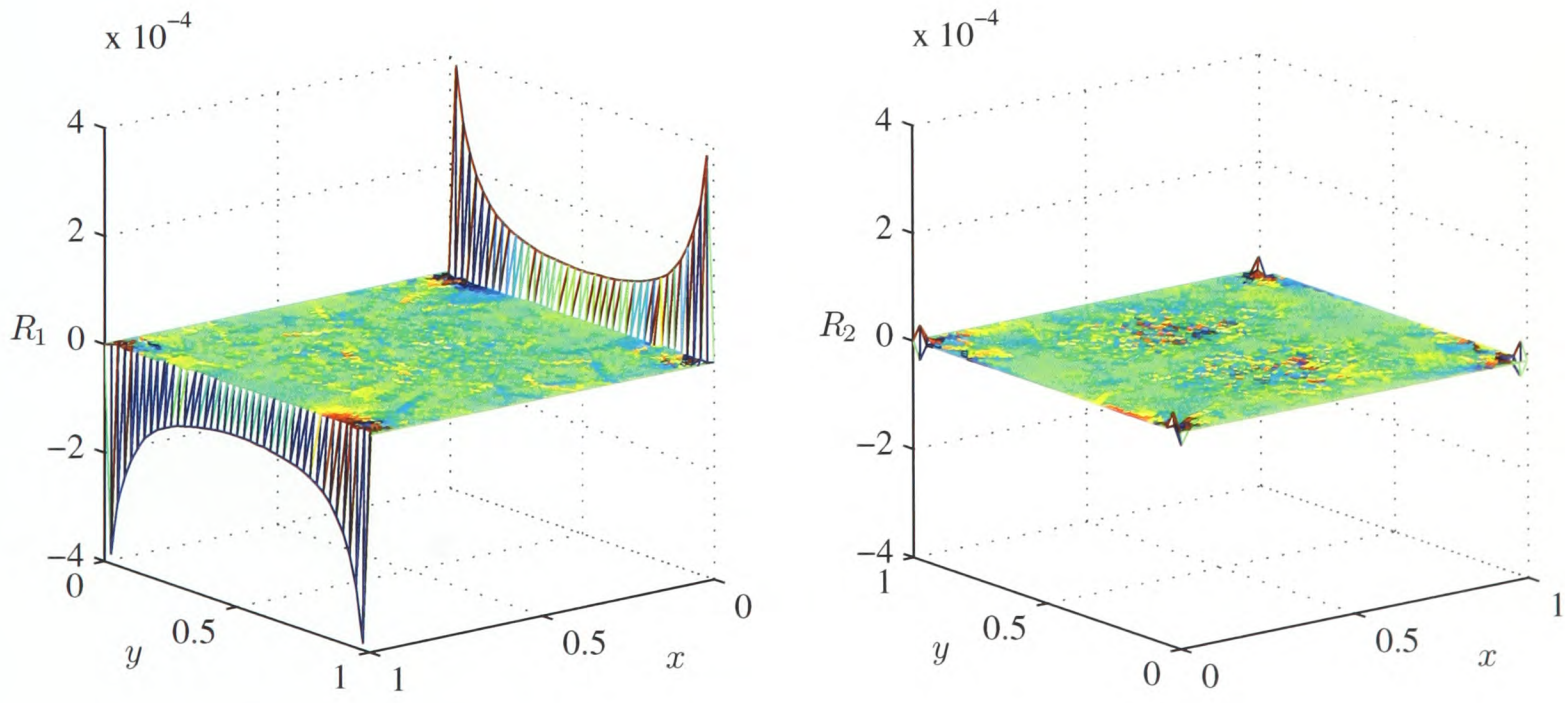


Figure 5.26: Residuals resulting from the numerically determined Lamé constants shown in Figure 5.25. The colour scale runs from -5×10^{-7} to 5×10^{-7}, so for any region which is not coloured dark red or dark blue $|R_i| < 5 \times 10^{-7}$.

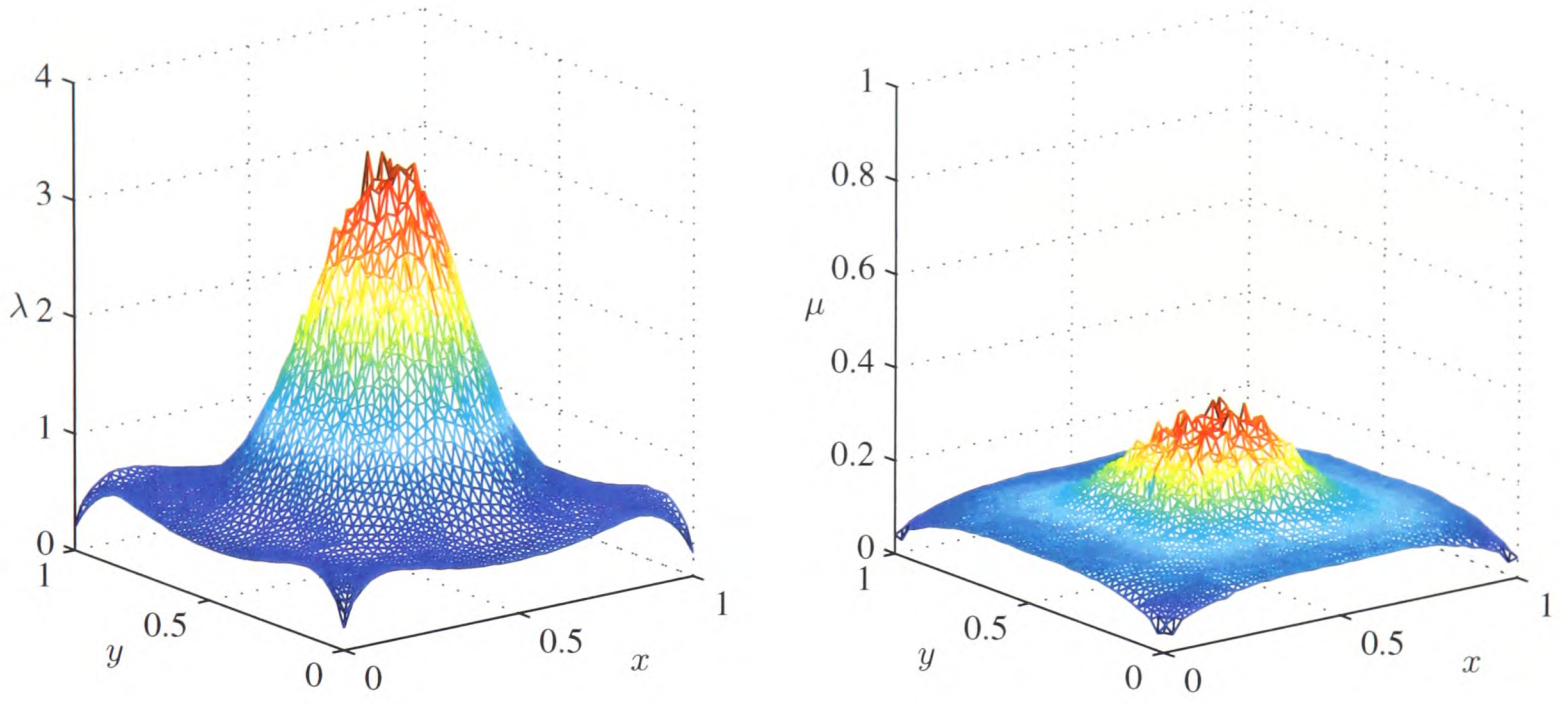


Figure 5.27: Numerical solution of the inverse problem using a displacement field generated from a single finite element solution of the forward problem with Lamé constants as shown in Figure 5.5 and with normally distributed noise with mean zero and variance 5×10^{-6}.

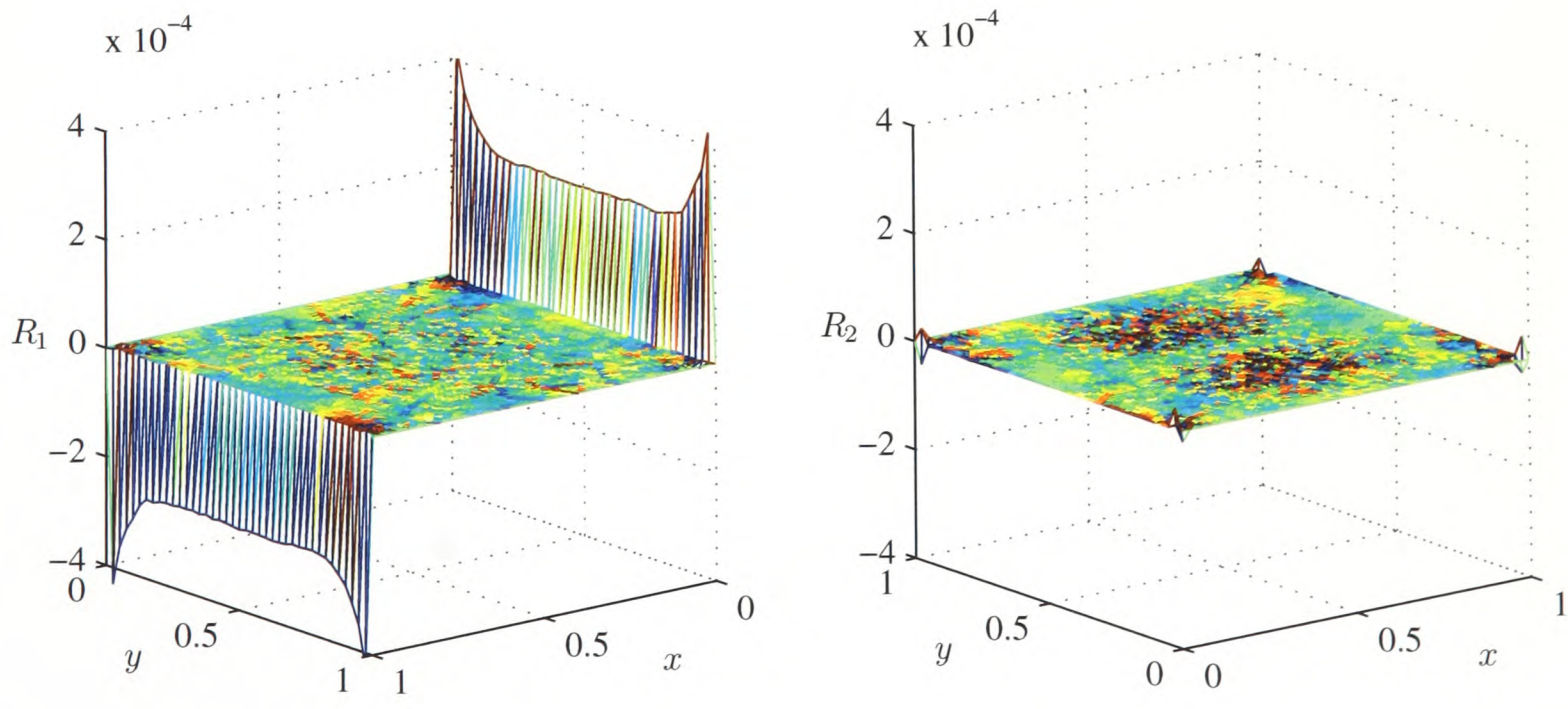


Figure 5.28: Residuals resulting from the numerically determined Lamé constants shown in Figure 5.27. The colour scale runs from -5×10^{-7} to 5×10^{-7}, so for any region which is not coloured dark red or dark blue $|R_i| < 5 \times 10^{-7}$.

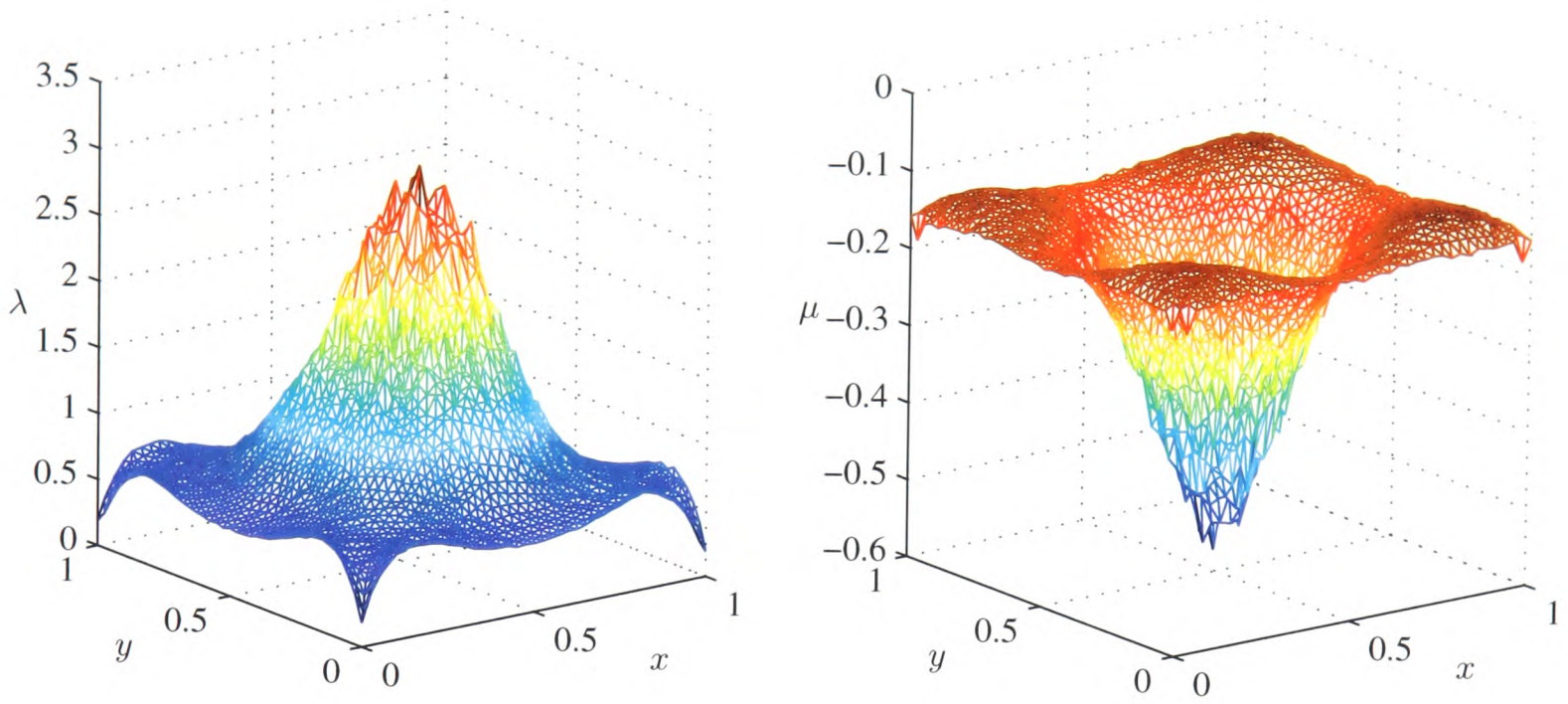


Figure 5.29: Numerical solution of the inverse problem using a displacement field generated from a single finite element solution of the forward problem with Lamé constants as shown in Figure 5.5 and with normally distributed noise with mean zero and variance 10^{-5}.

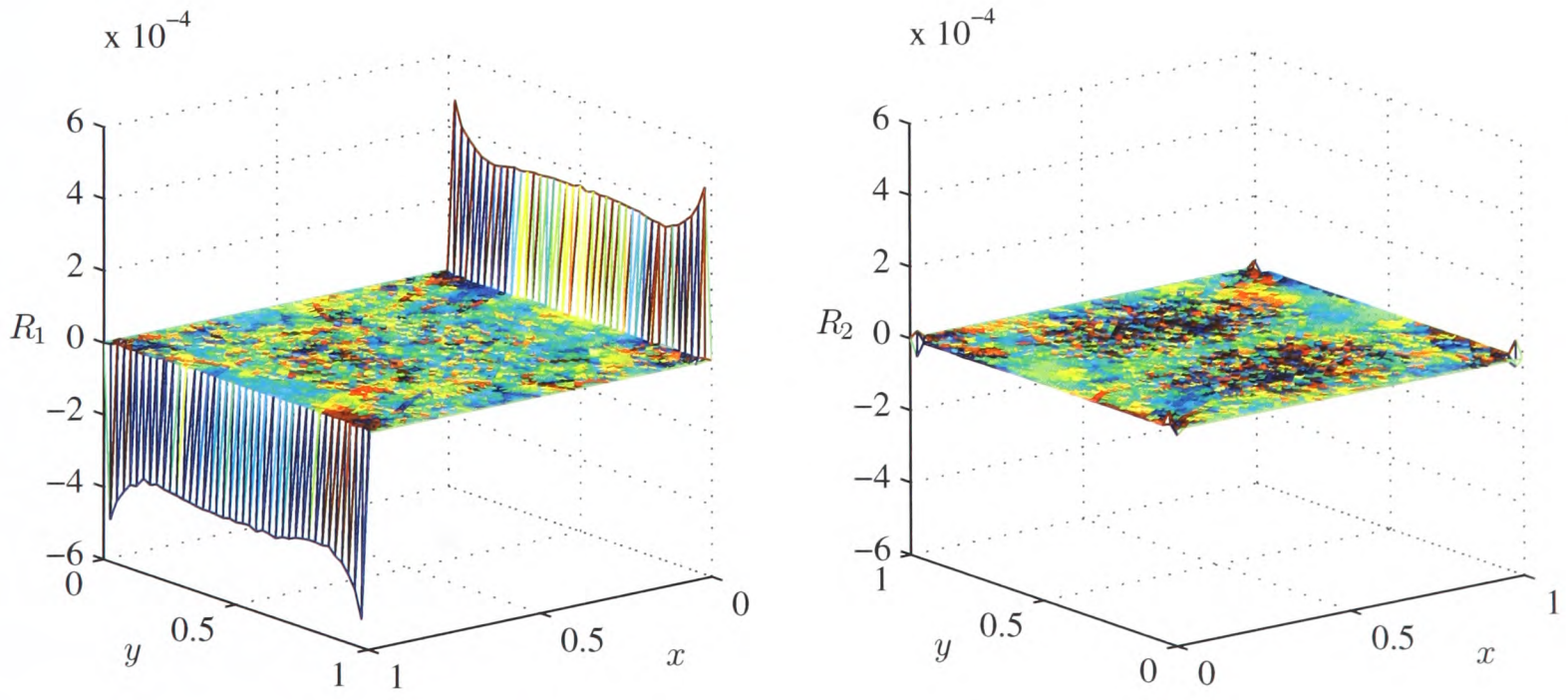


Figure 5.30: Residuals resulting from the numerically determined Lamé constants shown in Figure 5.29. The colour scale runs from -5×10^{-7} to 5×10^{-7}, so for any region which is not coloured dark red or dark blue $|R_i| < 5 \times 10^{-7}$.

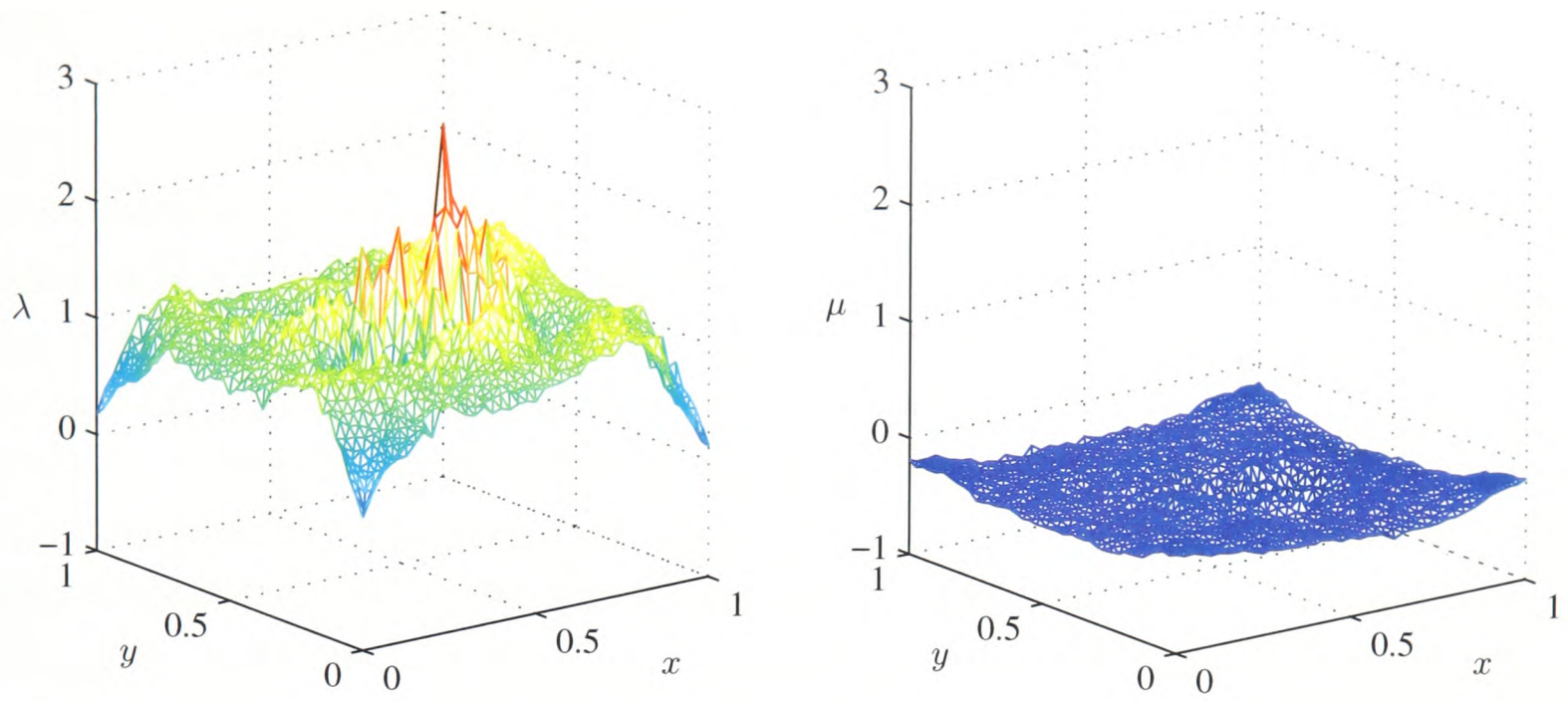


Figure 5.31: Numerical solution of the inverse problem using a displacement field generated from a single finite element solution of the forward problem with Lamé constants as shown in Figure 5.5 and with normally distributed noise with mean zero and variance 5×10^{-5}.

solution is unique only up to a multiplicative constant so it is the ratio of λ to μ that is important rather than their actual values). This agreement is supported by the size of the residuals illustrated in Figure 5.26, which throughout the domain are of comparable size to the residuals for the solution with no noise added. Figure 5.31 shows that for $\delta = 5 \times 10^{-5}$ the method is not successful in recovering the Lamé constants. This is the same value of δ for which the method was shown to break down when using one displacement field and boundary conditions.

Note, however, that for the single displacement field the method worked well when $\delta = 5 \times 10^{-5}$ whereas Figure 5.27 shows that the method has begun to be less reliable in the case of multiple displacement fields for this value of δ — the values of μ are much too small compared to those of λ. This is exactly the type of behaviour that was seen as the method began to fail with one displacement field. Further, Figure 5.29 shows a second type of intermediate behaviour for values of δ between those that the method works very well for and those for which it stops working completely. Here the values of μ become negative (this is thermodynamically impossible for a real material since applying a force in one direction cannot result in work being done in the opposite direction — see, e.g. Love [45]). The spatial variation of $-\mu$, however, is similar to that of the input data (although $-\mu$ is too small in comparison to λ) and it is clear where the feature is in the domain.

5.4.2 Sensitivity to Boundary Data

In this section the effect of prescribing incorrect boundary data for the Lamé constants is investigated. In a real imaging problem it may not be possible to produce images of a region of tissue under four linearly independent displacement fields. It will also, in general, not be possible to determine the values of λ and μ around the entire boundary of the image (in fact only on the top boundary — the skin — is this likely to be a possibility). Hence, it would be desirable for the effects of the prescribed boundary conditions to be seen only in a thin boundary adjacent to the boundary. Clearly this can only be tested for a single displacement field, since when four displacement fields are used there is no need to specify boundary conditions. The displacement field used in this section is the same as the one used in Sections 5.2 and 5.4.1, i.e. it results from a compression of a material with Lamé constants given by equations (5.24) and (5.25) in the y-direction, so the correct boundary conditions are

$$\lambda_0 = 1 + 9\exp[-20(0.25 + (y - 0.5)^2)], \quad \text{on } x = 0, 1, \tag{5.34}$$

$$\lambda_0 = 1 + 9\exp[-20(0.25 + (x - 0.5)^2)], \quad \text{on } y = 0, 1, \tag{5.35}$$

$$\mu_0 = \tfrac{2}{3}\lambda_0, \quad \text{everywhere on the boundary.} \tag{5.36}$$

Recall that if λ and μ constitute a solution of the inverse linear elasticity problem then so do $k\lambda$ and $k\mu$, where k is a constant. So, in order to determine the effect of altering the boundary conditions it is necessary to do more than simply multiply each of them by a constant.

Figure 5.32 shows the solution when the boundary values are prescribed to be $\lambda_0 = 1$, $\mu_0 = \frac{2}{3}$. These values are within 6% of the true boundary values everywhere on $\partial\Omega$ (and much closer near to the corners of the domain). The solution looks very similar to the input values of λ and μ as shown in Figure 5.5 and this is confirmed by calculating the residuals defined by equations (4.80) and (4.81). These residuals are shown in Figure 5.33, where it can be seen that they are of the same order of magnitude as those that arise when the specified boundary conditions are correct (shown in Figure 5.7), except for in some small regions close to the boundary (e.g. close to $x = 0.5$, $y = 0$ — one of the areas where the errors in the boundary values are at their largest). Clearly the parameter recovery method is robust with respect to small errors in specified boundary data.

The results of large errors in the boundary data are shown in Figure 5.34. These solutions

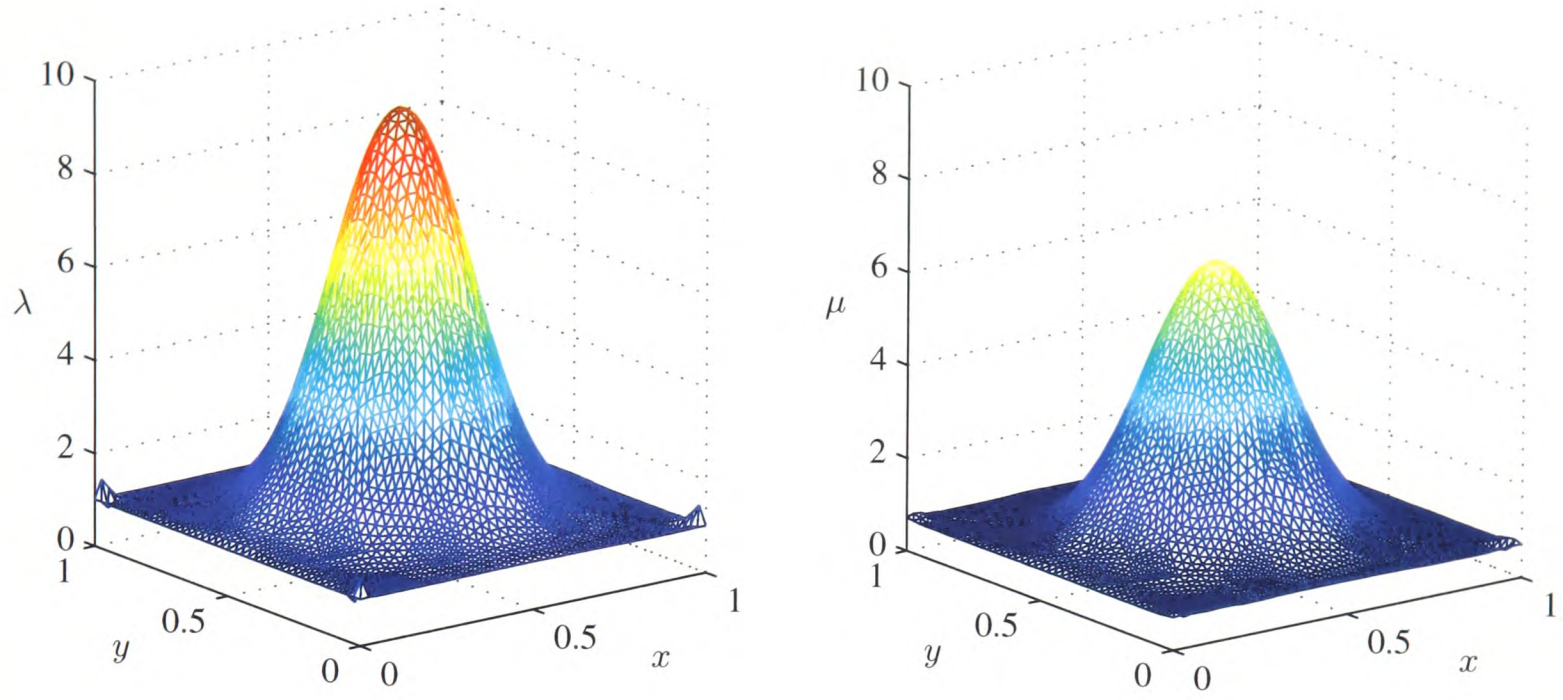


Figure 5.32: Numerical solution of the inverse problem using a displacement field generated from a single finite element solution of the forward problem with Lamé constants as shown in Figure 5.5. The Lamé constants are prescribed to be $\lambda_0 = 1$, $\mu_0 = \frac{2}{3}$ on the boundary.

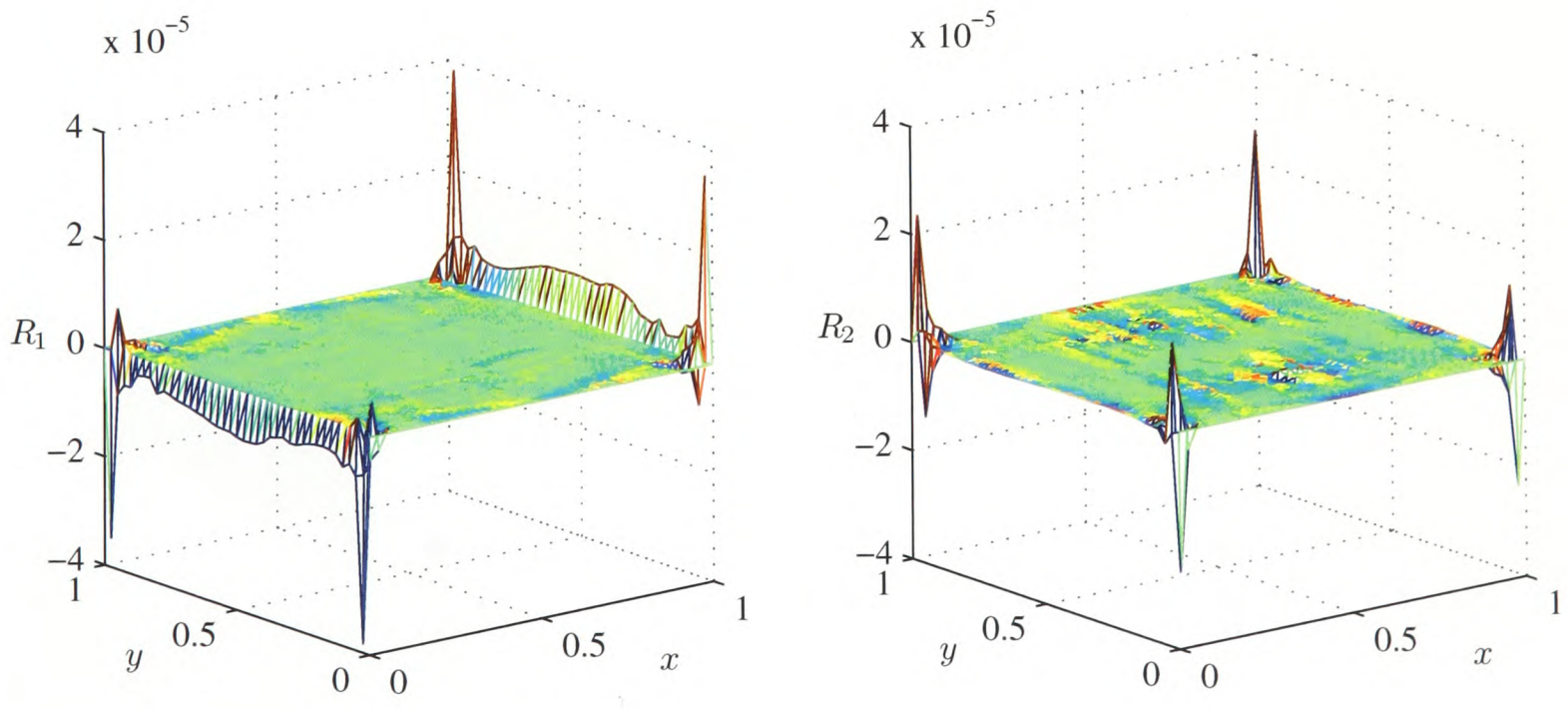


Figure 5.33: Residuals resulting from the numerically determined Lamé constants shown in Figure 5.32. The colour scale runs from -5×10^{-7} to 5×10^{-7}, so for any region which is not coloured dark red or dark blue $|R_i| < 5 \times 10^{-7}$.

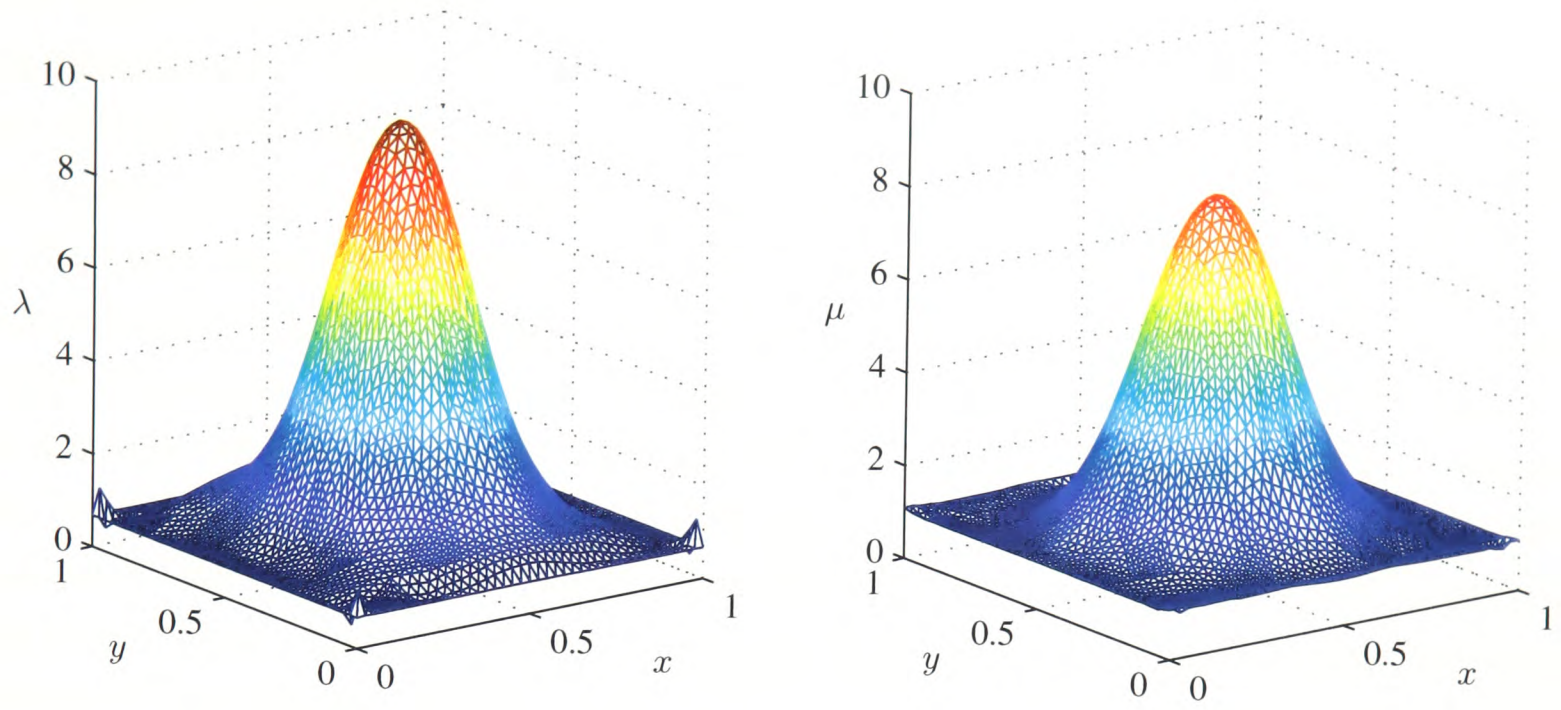


Figure 5.34: Numerical solution of the inverse problem using a displacement field generated from a single finite element solution of the forward problem with Lamé constants as shown in Figure 5.5. The Lamé constants are prescribed to be $\lambda_0 = \frac{2}{3}$, $\mu_0 = 1$ on the boundary.

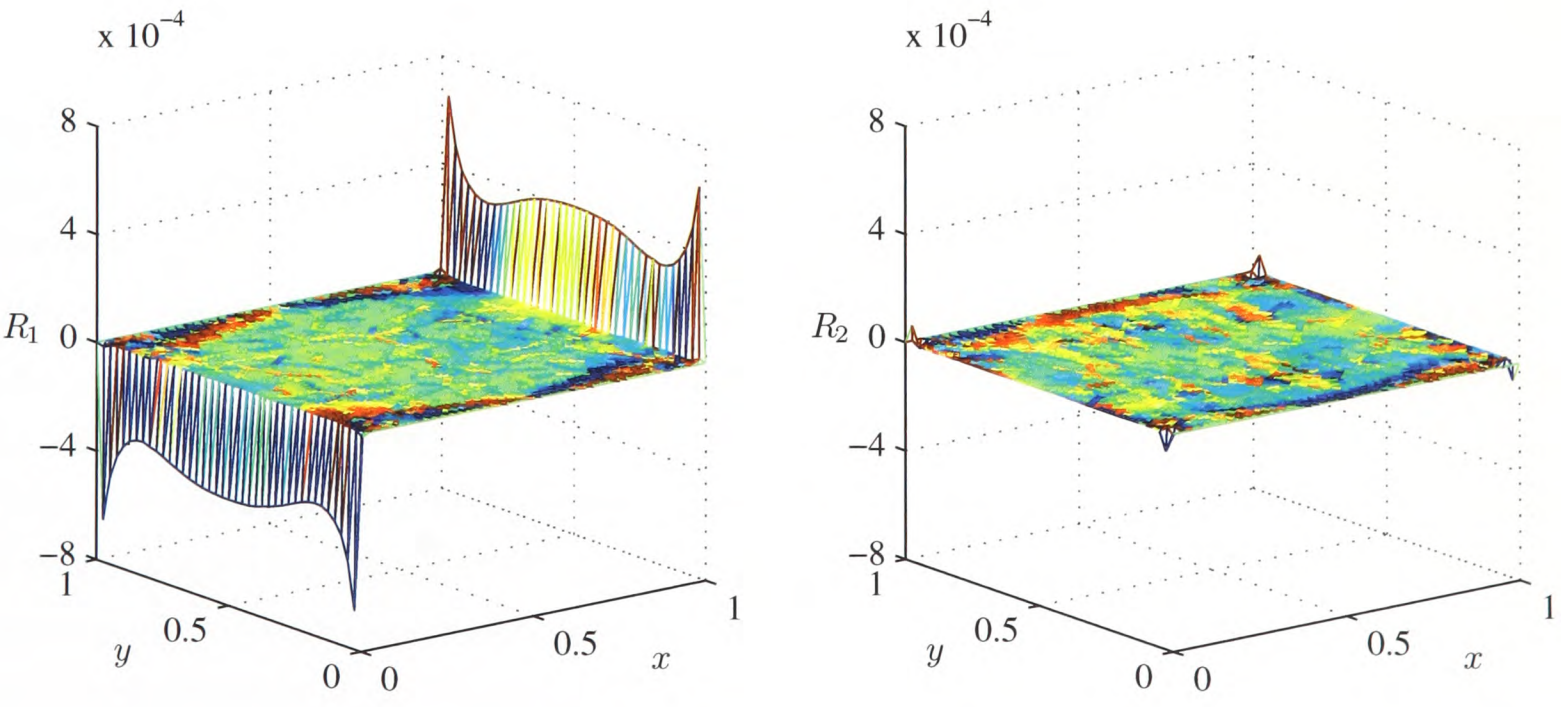


Figure 5.35: Residuals resulting from the numerically determined Lamé constants shown in Figure 5.34. The colour scale runs from -5×10^{-7} to 5×10^{-7}, so for any region which is not coloured dark red or dark blue $|R_i| < 5 \times 10^{-7}$.

are generated by setting $\lambda_0 = \frac{2}{3}$, $\mu_0 = 1$ everywhere on the boundary of the domain. This is an error of approximately 33% in the values of λ and 50% in that of μ. The shape of the solutions for both λ and μ are qualitatively very similar to the input shown in Figure 5.5 in the interior of the domain, although the ratio of λ to μ is incorrect: the values of μ should be smaller. Close to the boundary a thin boundary layer can be seen in which the values of the Lamé constants change very rapidly (λ increases rapidly since its prescribed value is too small, μ decreases).[6] Figure 5.35 shows the residuals corresponding to this solution of the inverse problem. It can be seen that in the centre of the domain the residuals are of the same size as those seen when the boundary values are given correctly. Close to the boundary, however, the residuals are much larger. This confirms that for specified boundary conditions that are significantly different from the actual data used to generate the displacement field the recovery method is still reliable for λ and μ *separately* in the interior of the domain. However, information about the relative sizes of λ and μ is lost.

Figure 5.36 shows the solutions to the linear elasticity inverse problem when the boundary conditions are

$$\lambda_0 = 1 + \tfrac{3}{4}\sin(10\pi y), \quad \text{on } x = 0, 1, \tag{5.37}$$

$$\lambda_0 = 1 + \tfrac{3}{4}\sin(10\pi x), \quad \text{on } y = 0, 1, \tag{5.38}$$

$$\mu_0 = \tfrac{2}{3}, \quad \text{everywhere on the boundary.} \tag{5.39}$$

So, the errors in λ_0 are large (up to 75% in places) and oscillate rapidly along the boundary. The errors in μ_0, on the other hand, are small (a maximum of 6%). Despite the large, oscillatory errors in the boundary data the solutions appear similar to the input data shown in Figure 5.5, although both λ and, to a lesser extent, μ oscillate throughout the domain. The amplitude of these oscillations does not, however, increase in proportion to the value of the Lamé constants (i.e. they do not have a larger amplitude near to the centre of the domain where $\lambda = O(10)$ than near to the boundary where $\lambda = O(1)$). The residuals shown in Figure 5.37 demonstrate that, as for the other examples in this section, the largest errors in the solution occur only near to the boundary — although in this case the residuals near to the centre of the domain are also larger than those seen in Figure 5.12 when the correct boundary conditions are given, roughly by a factor of 10. Nevertheless, the feature

[6]The boundary layers are difficult to see in Figure 5.34 since the sizes of the changes in the values of λ and μ in the boundary layer are much smaller in absolute terms than the difference between the boundary values and the value in the centre of the domain. The easiest place to see the boundary layer is close to the line $y = 0$ in the plot of λ.

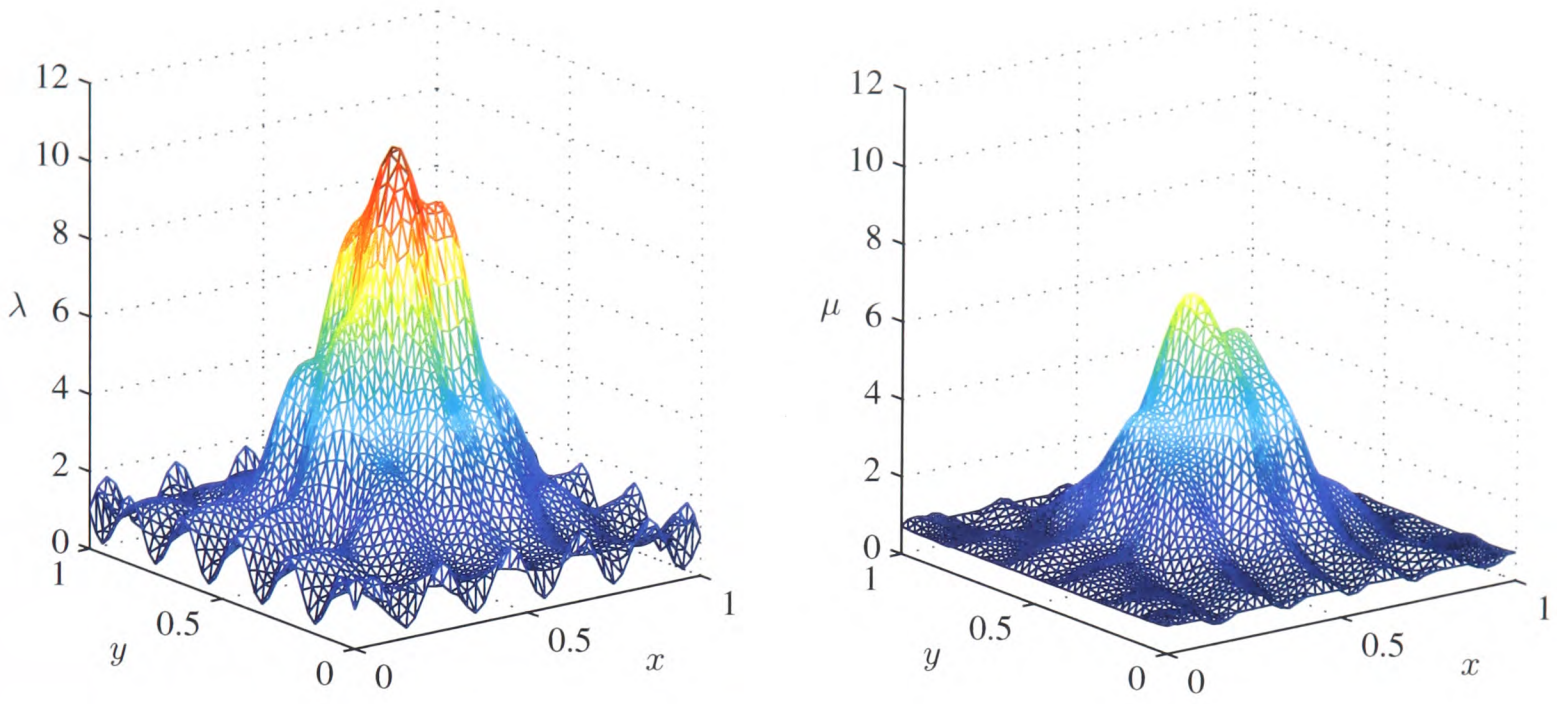


Figure 5.36: Numerical solution of the inverse problem using a displacement field generated from a single finite element solution of the forward problem with Lamé constants as shown in Figure 5.5. The Lamé constants are prescribed to be $\lambda_0 = 1 + \frac{3}{4}\sin(10\pi y)$ on $x = 0, 1$, $\lambda_0 = 1 + \frac{3}{4}\sin(10\pi y)$ on $y = 0, 1$ and $\mu_0 = \frac{2}{3}$ all around the boundary.

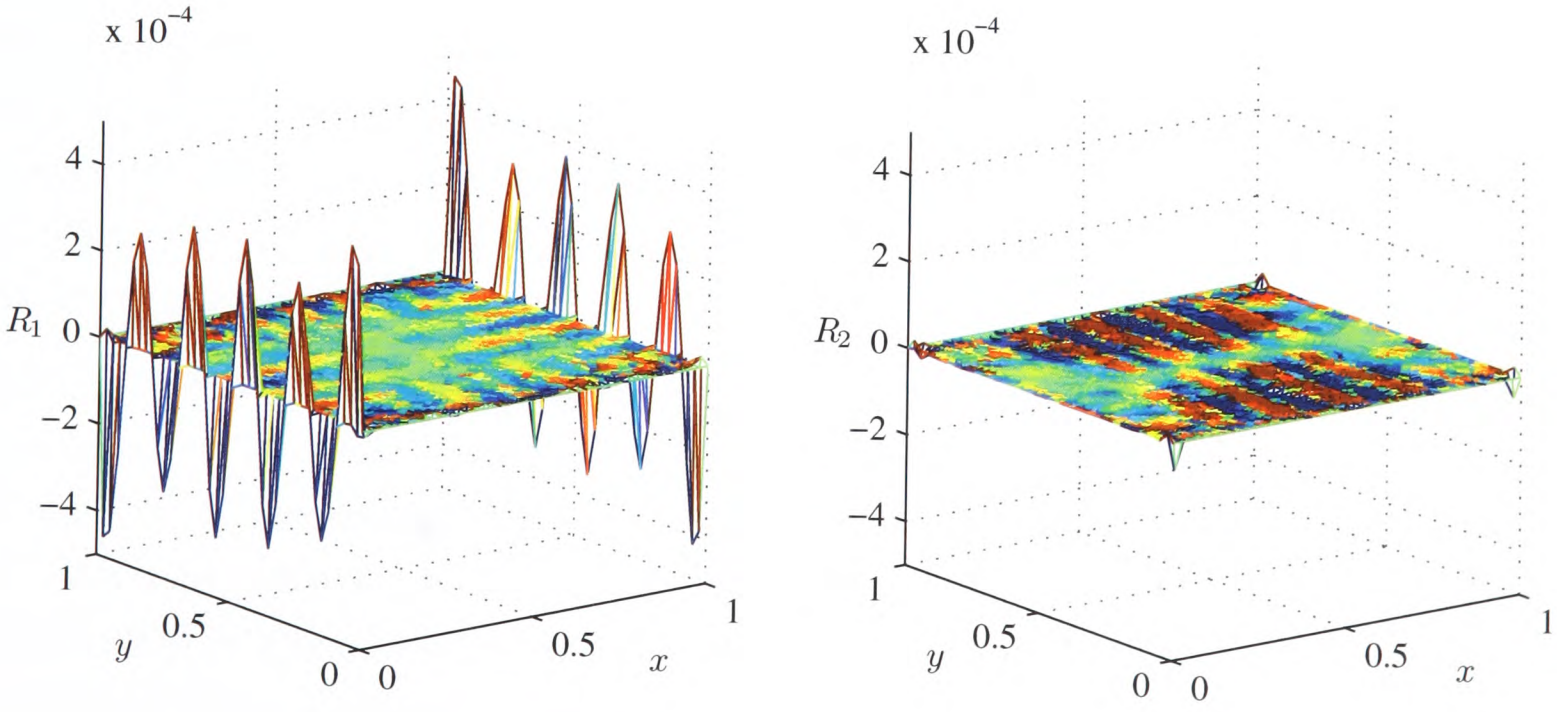


Figure 5.37: Residuals resulting from the numerically determined Lamé constants shown in Figure 5.36. The colour scale runs from 10^{-6} to 10^{-6}, so for any region which is not coloured dark red or dark blue $|R_i| < 10^{-6}$.

in the centre of the domain can be clearly identified in Figure 5.36 and the ratio of λ to μ is preserved away from the boundaries. So, the parameter recovery method is robust even with respect to large, rapidly varying errors in the boundary data. This is encouraging as it indicates that the method should be able to deal with real tissue which, although likely to have an approximately constant stiffness, contains slight inhomogeneities.

5.5 Conclusions

In this chapter the results of testing the parameter recovery method developed in Chapter 4 are described. The method has been shown, in Sections 5.1 and 5.2, to be effective in recovering parameters that vary continuously across the domain, in the cases of both exact and numerical solutions of the forward problem for linear elasticity. It does not work well, however, when the Lamé constants are discontinuous. This is likely to be due to the fact that the finite element scheme uses continuous basis functions, so it cannot deal with sudden changes in the values of λ and μ.

Although the method works well both for a single displacement field with boundary conditions applied and for four independent displacement fields with no boundary conditions, the results are generally slightly more accurate for the single field. In Section 5.3 this was shown to be due to the presence of a stress singularity in the solution of the forward problem, forcing the values of the Lamé constants in the corners of the domain to vanish. For a series of real ultrasound images this problem will not arise as the ultrasound probe can be more correctly modelled as a smooth indenter with a free boundary. In this situation the stress singularity is not seen.

Finally, in Section 5.4, the parameter recovery method was shown to be robust when given input data with errors. If small errors (less than 1% of a typical displacement) were added to the true displacement field then the method was still able to generate a qualitatively accurate representation of the spatial variation of λ and μ. For larger errors, first the relationship between λ and μ was lost (although the presence of features was still apparent in both of them separately) and then the method broke down completely. If incorrect boundary values of the Lamé constants were given then the method was able to deal with very large (in relation to the actual values) errors in λ_0 and μ_0.

Chapter 6

The 3-Dimensional Inverse Problem

In this chapter the validity of the plane strain assumption used in Chapter 4 is considered. In Section 6.1 a typical solution of the 2-d forward problem is compared with the cross-section of the solution to an appropriate 3-d forward problem and the results are shown to be significantly different. In addition, the parameter recovery method of Chapter 4 is used to attempt to recover the input parameters used in a finite element simulation of the 3-d forward problem.

In Section 6.2 the 3-d linear elasticity inverse problem is formulated and shown to be hyperbolic. The resulting system consists of three equations and so is over-specified for the two unknowns, λ and μ. However, for a real displacement field they should be consistent with one another and so the solution of any two of them should also satisfy the third equation. However, the third equation contains extra information that can be used to reduce the need to apply boundary conditions. Further additional information can also be obtained by using multiple displacement fields. These two situations are considered in Section 6.3.

In Section 6.4 the inverse problem is reformulated as a minimization problem. This is analogous to the reformulation of the 2-d inverse problem in Section 4.4.1. The minimization approach is preferred over direct solution of the inverse problem since it allows the information contained in all three equations for λ and μ to be used simultaneously. It also eliminates the need to solve a hyperbolic system numerically. The difficulty of doing this has been discussed previously. The minimization problem can be solved using finite element methods in exactly the same way as described for the 2-d case in Section 4.4.4.

The solution is not implemented in this work, however, as series of 3-d ultrasound images for use in elasticity imaging are not currently available.

6.1 2-D Plane Strain vs 3-D Elasticity

In this section the results of the 2-d plane strain and 3-d models of the deformation of a region of soft tissue by an ultrasound probe are compared. In Section 6.1.1 the results of the two forward problems in the plane of the ultrasound probe ((y, z)-plane) are compared. The 2-d and 3-d models are shown to give very different results: the deformation in the y-direction is much smaller in the 3-d case than the 2-d case and, most significantly, the 2-d deformation is symmetric in both y- and z-directions whilst the deformed cross-section of the 3-d case is not symmetric in the z-direction. This shows that the plane strain assumption is *not valid* for the ultrasound elasticity problem — a full 3-d model is needed.

Nevertheless, existing elasticity imaging techniques do give useful information about tissue properties. So, it may be possible that solving the 2-d inverse problem on a cross-section of the solution of the 3-d forward problem will give a qualitative indication of which regions are stiffer than others — even if this method cannot quantify the difference in stiffness. This possibility is investigated in Section 6.1.2. Finally, a brief discussion of the computational expense of moving from 2-d to 3-d is given in Section 6.1.3.

6.1.1 Comparing the Forward Problems

As a test of the the validity of the plane strain assumption, in this section a 2-d plane strain solution of the elasticity forward problem is compared to an equivalent 3-d simulation. In both cases the deformations are intended to represent the deformation of a region of soft tissue by an ultrasound probe (long, thin indenter). Both simulations use the same parameter values so, if the 2-d plane strain inverse problem is to be valid (i.e. if it can be used to infer the parameter values of the equivalent 3-d problem), then they should give rise to similar displacement fields in the plane of the ultrasound image.

The 3-d simulation was performed in Matlab using a finite element mesh with 8880 tetrahedral elements and 1762 nodes. The undeformed mesh occupied the region $(x_1, x_2, x_3) = (x, y, z) = [-1, 1] \times [0, 1] \times [0, 1]$, which was deformed by displacing the part of the plane

$z = 1$ where $|x| \leqslant 0.05$ so that $z = 0.8$ and there is no displacement in the x- and y-directions. The plane $z = 0$ was held fixed and the remainder of the boundary was assumed to be stress free. The elements are of variable size, with those close to the plane $x = 0$ being much smaller than those close to the faces $x = \pm 1$. There are two reasons for this non-uniform element size:

- The plane $x = 0$ is to be compared to a 2-d deformation, so the displacements here need to be most accurate. By using larger elements further away the computational time is reduced.
- The largest displacements are expected close to the line $(x, z) = (0, 1)$ that is indented. So, using larger elements away from this line should not significantly change the calculated solution.

The Lamé constants were given by

$$\lambda_{\text{in}}(x, y, z) = 1 + 9\exp\big[-20\big(x^2 + (y - 0.5)^2 + (z - 0.5)^2\big)\big], \tag{6.1}$$

$$\mu_{\text{in}}(x, y, z) = \tfrac{1}{50}\lambda_{\text{in}}(x, y, z). \tag{6.2}$$

Note that the shear modulus (μ_{in}) is much smaller in comparison to λ_{in} than the values typically used in Chapter 5. This is because in 3-d the material that extends in the x-direction provides additional resistance to deformation in the y-direction, so a smaller shear modulus is required in order to see any significant lateral displacements. This observation should already begin to indicate that the plane strain assumption is highly unlikely to be valid. The resulting deformed cuboid is shown in Figure 6.1. A discussion of the time required to perform this simulation is postponed until Section 6.1.3.

There are two possible choices of the 2-d Lamé constants that could be considered to be equivalent to the 3-d parameters: either they could be chosen to take the same values as the 3-d Lamé constants or they could be chosen so that the Young's modulus and Poisson's ratio are the same in both 2-d and 3-d.[1] If the Lamé constants are chosen to be the same

[1]Recall that in 3-d equations (2.54) and (2.56) define the Young's modulus and Poisson's ratio in terms of the 3-d Lamé constants. Similar equations exist to relate Young's modulus and Poisson's ratio to the 2-d Lamé constants. The 2-d Lamé constants are chosen so that the 2-d equations define the same values of the Young's modulus and Poisson's ratio as equations (2.54) and (2.56) do for the 3-d Lamé constants.

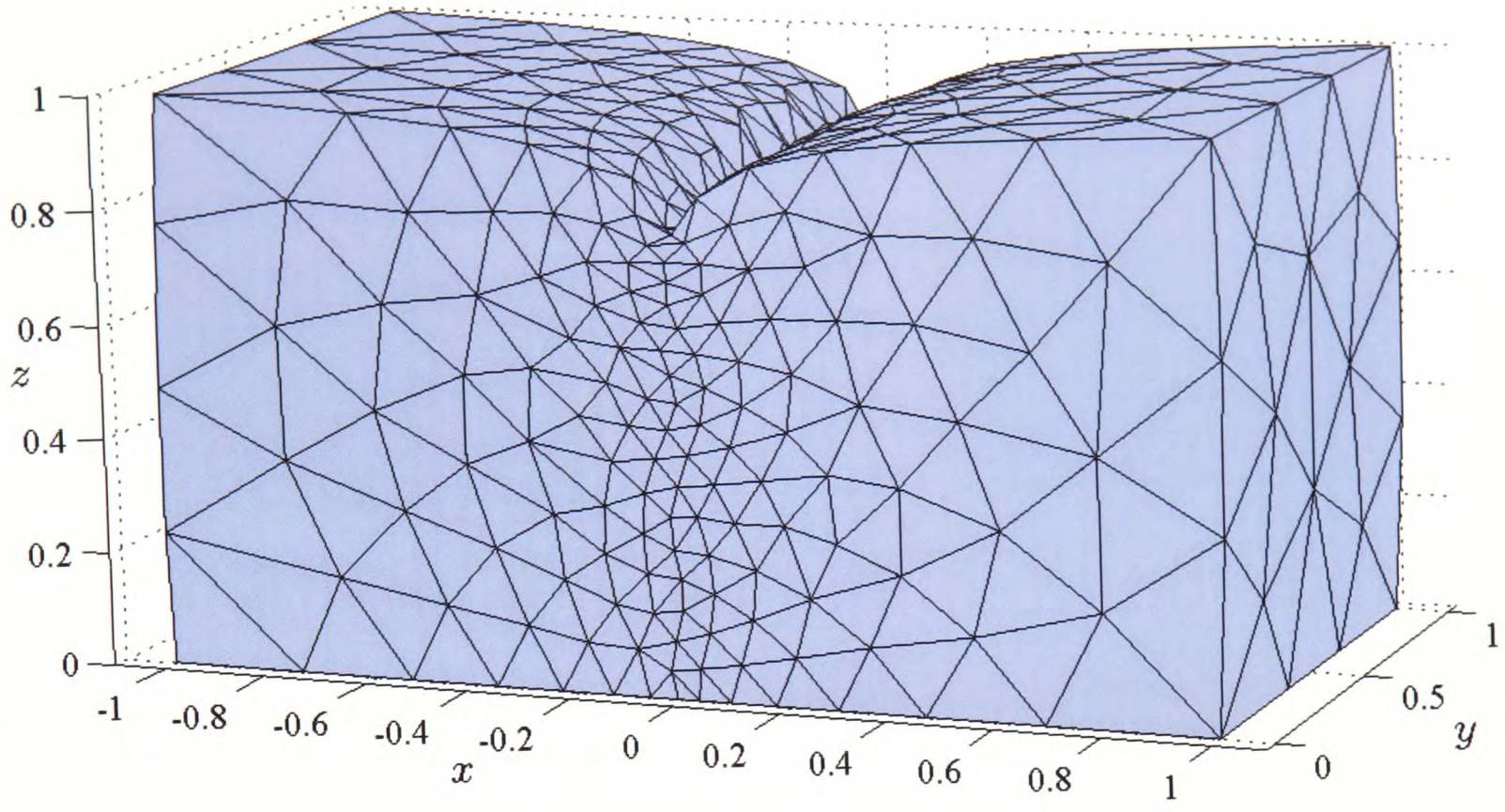


Figure 6.1: Simulated deformation of 'soft tissue' by an 'ultrasound probe' that displaces the part of the plane $z = 1$ where $|x| \leqslant 0.05$ to $z = 0.8$. There is no displacement in the x- and y-directions.

in both 2-d and 3-d then in the plane $x = 0$ they take the values

$$\lambda_{\text{in}}(y, z) = 1 + 9\exp\left[-20\big((y - 0.5)^2 + (z - 0.5)^2\big)\right], \tag{6.3}$$

$$\mu_{\text{in}}(y, z) = \tfrac{1}{50}\lambda_{\text{in}}(y, z). \tag{6.4}$$

When the 2-d Lamé constants are chosen so that the Young's modulus and Poisson's ratio are the same as in 3-d it turns out that the same μ_{in} can be used and in this case $\lambda_{\text{in}} = \frac{25}{13}\mu_{\text{in}}$. The mesh used in Chapter 5, consisting of 5248 elements and 2705 nodes, was again used. The undeformed mesh occupies the region $(y, z) = [0, 1] \times [0, 1]$ and the line $z = 0$ was held fixed and $z = 1$ moved to $z = 0.8$. This is the 2-d equivalent of the motion of the plane $x = 0$ in the 3-d case. The results of solving the 2-d forward problems with these choices for the Lamé constants are shown in Figure 6.2, with the cross-section of the 3-d solution at $x = 0$ overlaid and shaded blue.

Inspection of Figure 6.2 immediately leads to the conclusion that neither of the two 2-d displacement fields is equivalent to the 3-d one. In both cases more lateral displacement is seen in the 2-d fields than in the 3-d one. The difference in lateral displacements is

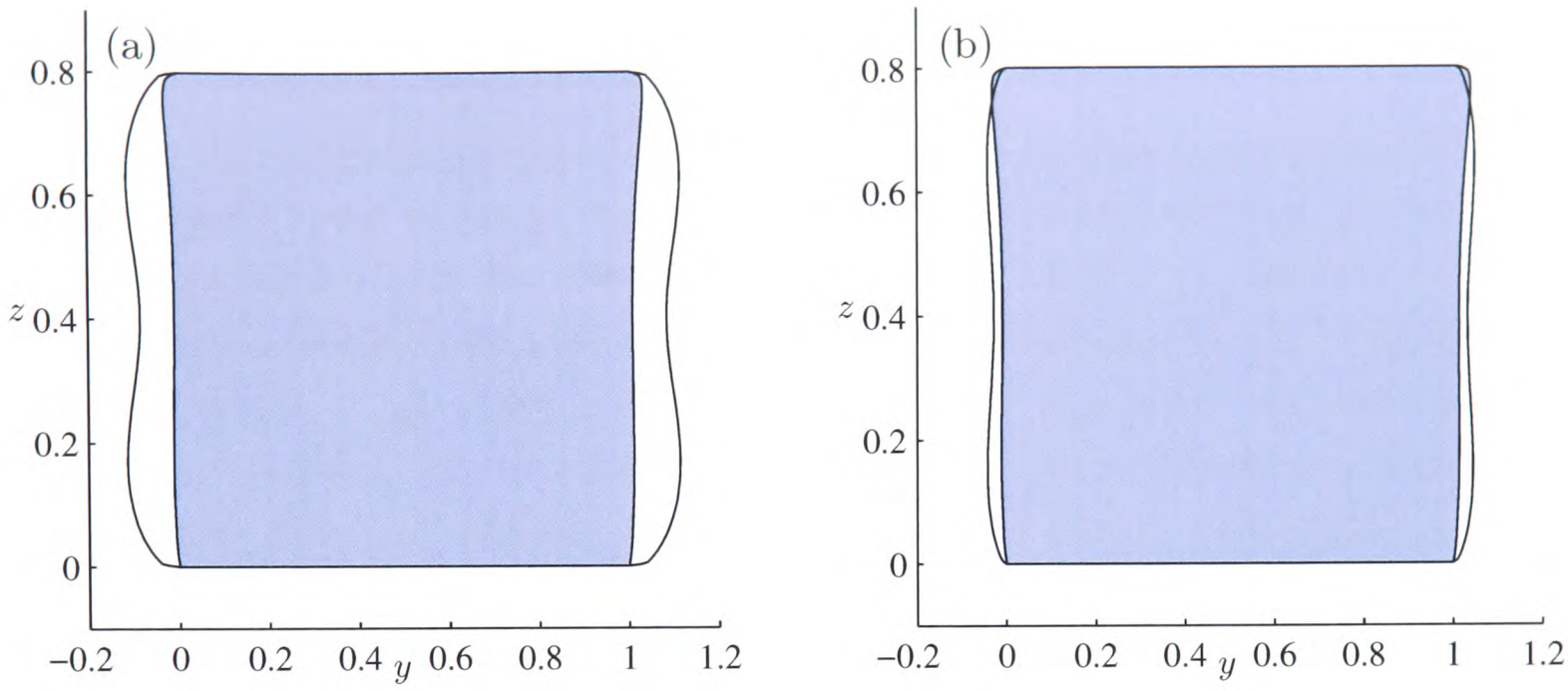


Figure 6.2: Cross-section of the solution of a 3-d linear elasticity forward problem at $x = 0$ (shaded) compared to the solutions of two plane strain problems in 2-d (unshaded). In both 2-d cases the unit square is deformed by moving the edge $z = 1$ to $z = 0.8$. In the 3-d case the part of the plane $z = 1$ where $|x| \leqslant 0.05$ is moved so that $z = 0.8$ and there is no displacement in the x- and y-directions. In the plot labelled (a) the Lamé constants are given by equations (6.3) and (6.4). In the plot labelled (b) they are calculated so that the Young's modulus and Poisson's ratio are the same in 2-d and 3-d.

smaller in the case where the Lamé constants are chosen so that the Young's modulus and Poisson's ratio are the same in both 2-d and 3-d. This suggests that preserving the Young's modulus and Poisson's ratio is the more appropriate way of switching between 2-d and 3-d.

However, discussion about the choice of appropriate 2-d parameters turns out to be irrelevant due to the observation that the 2-d displacement field is symmetric about the line $z = 0.5$ (in the undeformed coordinate system, $z = 0.4$ in the deformed body), whereas the 3-d field is not. In fact, *all* 2-d deformations of bodies with symmetric distributions of λ and μ will be symmetric since there is no difference between the lines $z = 0$ and $z = 1$. This is *not the case* in 3-d — the planes $z = 0$ and $z = 1$ have *different boundary conditions* applied on them away from $x = 0$ (zero displacement on $z = 0$, zero stress on $z = 1$). Hence, the deformation field cannot be expected to be symmetric. Since the 3-d deformation is clearly an accurate representation of the deformation of soft tissue by a linear ultrasound probe and the 2-d deformation is the *only* plane strain equivalent to it the only conclusion that can be drawn from these results is that plane strain is *not appropriate* for modelling tissue deformation in ultrasound elasticity.

6.1.2 2-D Parameter Recovery on a Cross-Section

Recall from the review in Chapter 3 that existing elasticity imaging techniques do "work", i.e. they identify areas of tissue that are stiffer than their surroundings. In this section this fact is reconciled with the observation of Section 6.1.1 that the solutions of the 2-d and 3-d forward problems with the same parameters are not similar, i.e. the plane strain assumption is not valid. This is done by applying the parameter recovery method of Chapter 4 to cross-sections through the 3-d displacement fields generated in this chapter. Since the (2-d) parameter recovery method has no way of dealing with displacements out of the plane of the image all of the cross-sections considered are through the plane $x = 0$, where there is no displacement in the x-direction (since the displacement field is symmetric about this plane).

Figure 6.3 shows the values of the Lamé constants defined by equations (6.3) and (6.4) (and used to generate the deformation shown in Figure 6.1) on the plane $x = 0$. The results of applying the parameter recovery method of Chapter 4 to the displacement field resulting from this (3-d) distribution of Lamé constants (with differing applied Dirichlet boundary conditions) are shown in Figures 6.4 and 6.6. In order to generate Figure 6.4 the values of λ and μ on the boundary were specified to be those given by equations (6.3) and (6.4) respectively. To generate Figure 6.6 the boundary conditions were chosen to correspond to the values of the Lamé constants that ensure that the Young's modulus and Poisson's ratio are the same in both 2-d and 3-d. This distribution is shown in Figure 6.5.

In both cases it can be seen that the parameter recovery method is able to detect that the tissue is much stiffer in the interior of the domain than close to the edges. This is consistent with the observation that imaging techniques using the plane strain assumption have been shown to detect regions of stiff tissue. However, it is clear that the parameter recovery method does not work as well for the cross-section as it did for the plane strain deformations considered in Chapter 5. Most obviously, the shapes of the detected solution surfaces shown in Figure 6.4 are much less similar to those of the true parameter values than was seen for the plane strain solutions. In particular, in Figure 6.5 the recovered values of λ and μ become negative on either side of the central peak.

Further, quantitative information about the values of the Lamé constants or the ratio of λ to μ is lost. In Figure 6.4 the maximum value of λ is more than twenty times the boundary value, compared to a ratio of only ten for the true values illustrated in Figure 6.3. The

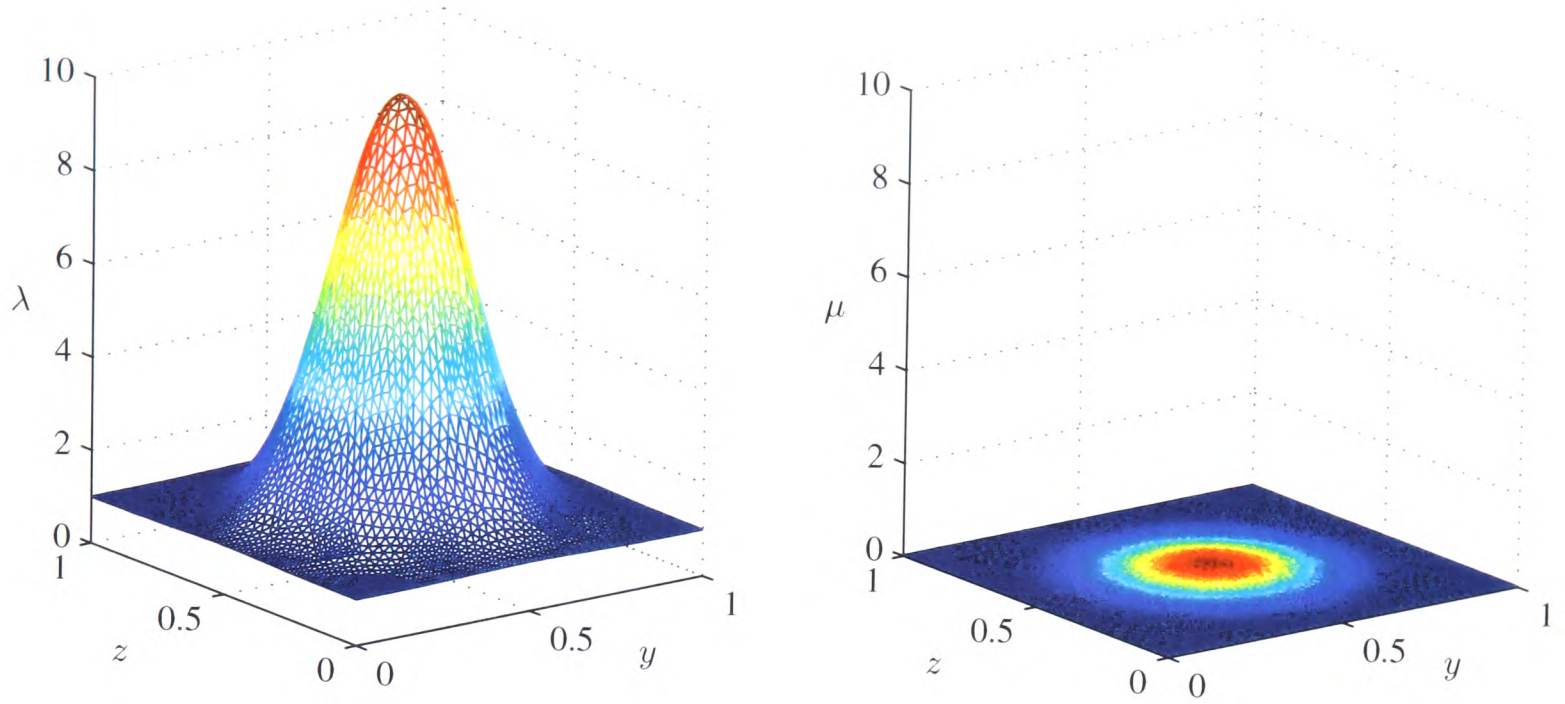


Figure 6.3: Spatial variation across the plane $x = 0$ of the Lamé constants used to generate the deformation shown in Figure 6.1. Here $\lambda_{\text{in}}(x, y, z) = 1 + 9\exp[-20(x^2 + (y - 0.5)^2 + (y - 0.5)^2)]$, $\mu_{\text{in}}(x, y, z) = \lambda_{\text{in}}(x, y, z)/50$.

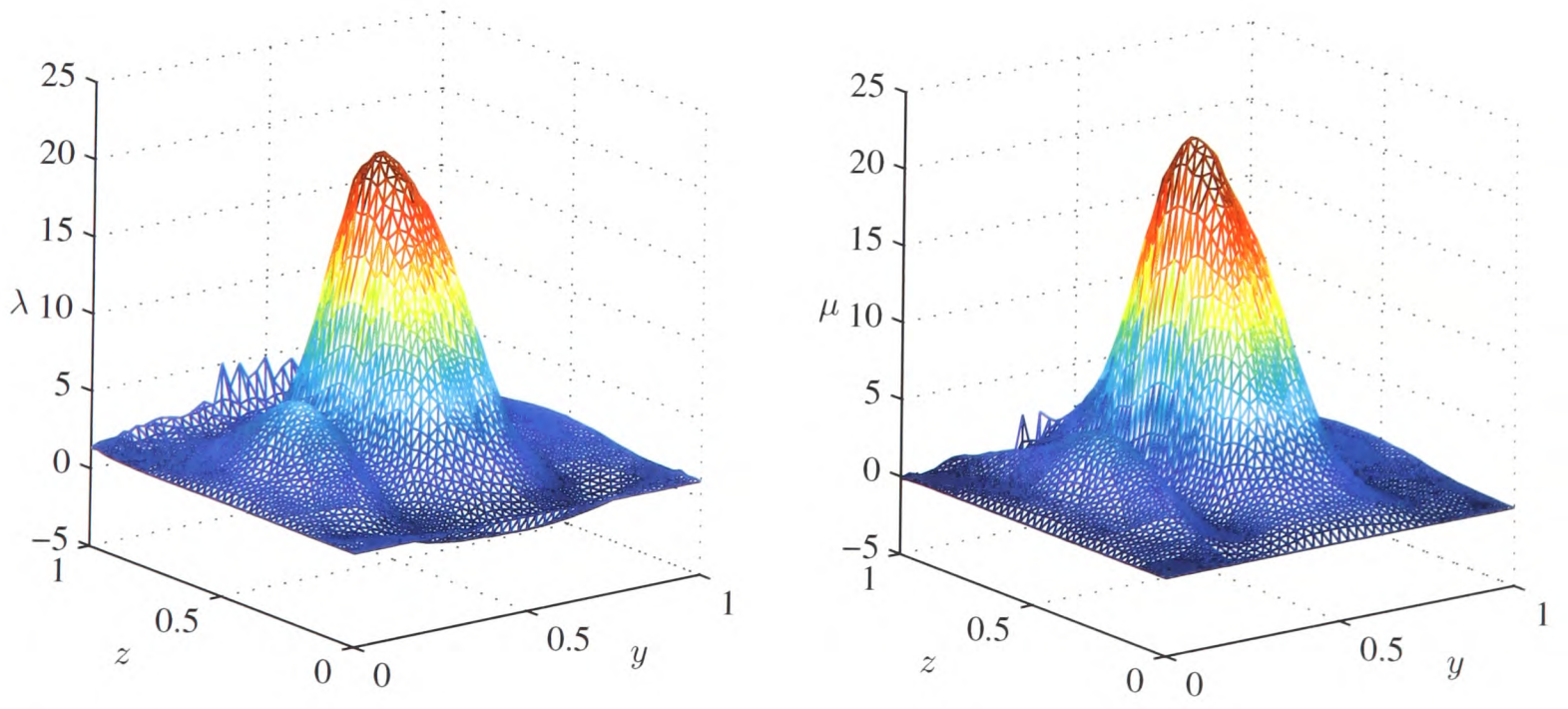


Figure 6.4: Result of applying the 2-d parameter recovery method to the cross-section through the plane $x = 0$ of the 3-d forward problem illustrated in Figure 6.1. Dirichlet boundary conditions $\lambda(y, z) = \lambda_{\text{in}}(0, y, z)$, $\mu(y, z) = \mu_{\text{in}}(0, y, z)$ are applied on the boundary, where λ_{in} and μ_{in} are defined by equations (6.3) and (6.4) respectively.

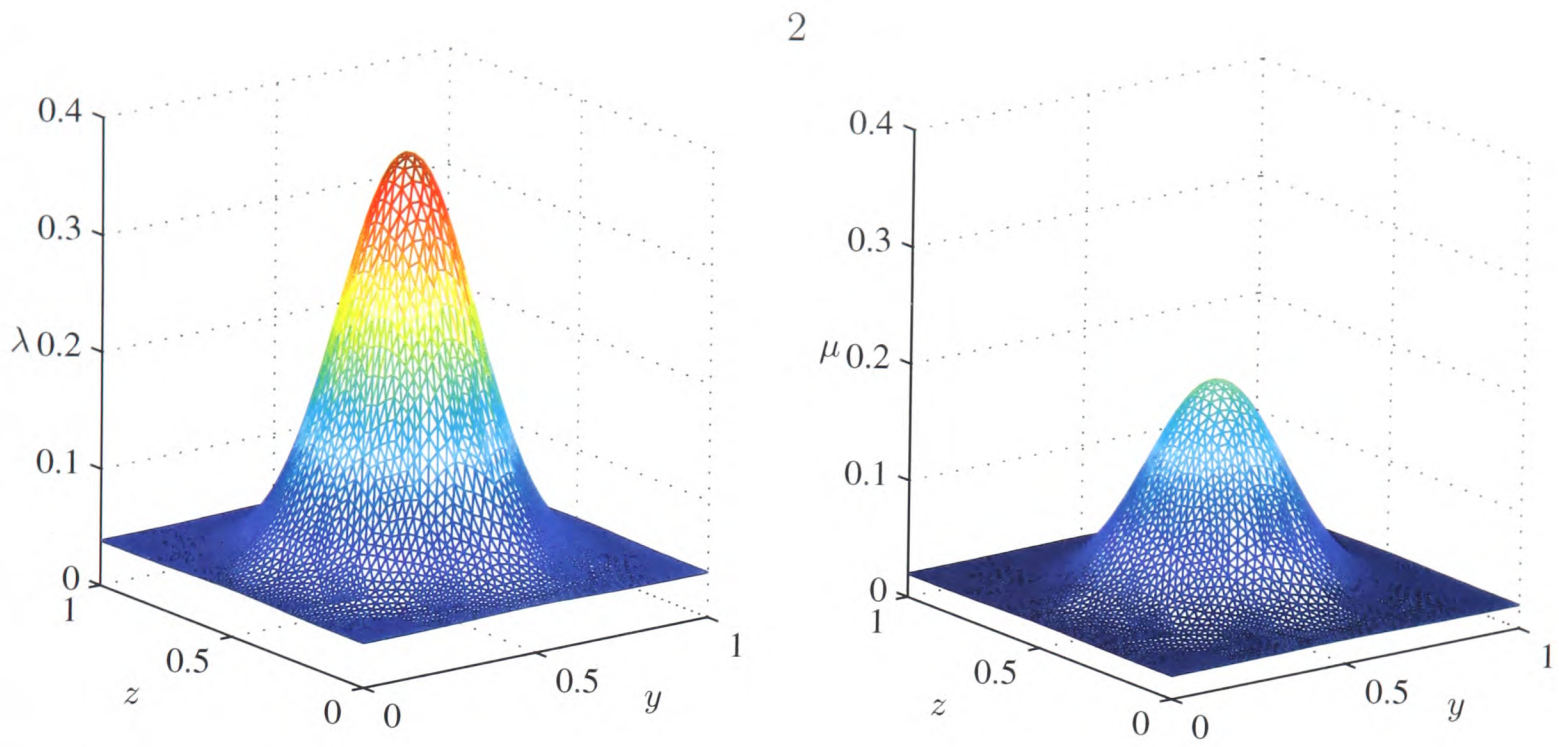


Figure 6.5: Spatial variation across the plane $x = 0$ of the Lamé constants that give the same values of the Young's modulus and Poisson's ratio (in 2-d) as those used to generate the deformation shown in Figure 6.1 (in 3-d).

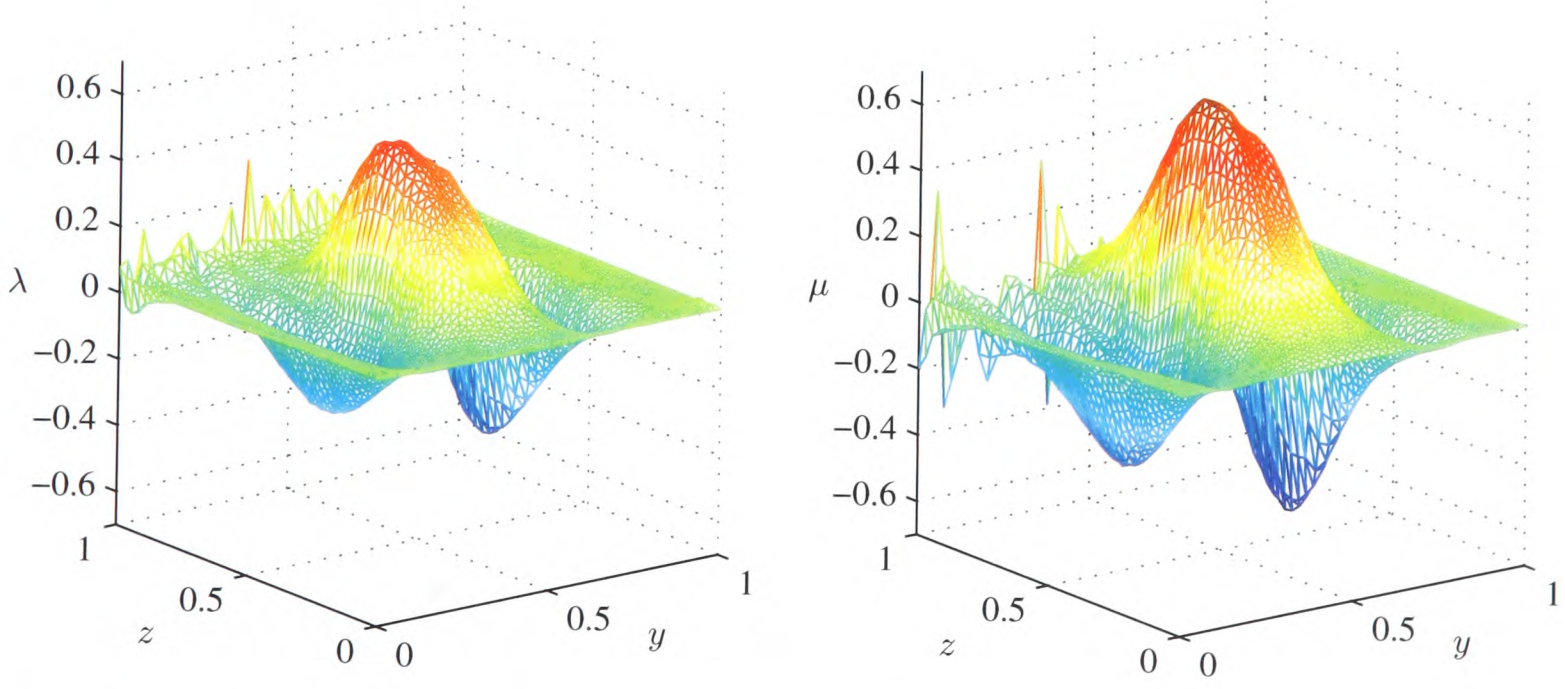


Figure 6.6: Result of applying the 2-d parameter recovery method to the cross-section through the plane $x = 0$ of the 3-d forward problem illustrated in Figure 6.1. Dirichlet boundary conditions are applied with the values of λ and μ on the boundary chosen so that the Young's modulus and Poisson's ratio are the same as those used to generate the displacement field in 3-d.

problem is even more noticeable for μ, where the maximum value is more than a thousand times the boundary values (again the ratio should be ten). Similarly, the value of λ used in generating the displacement field was fifty times that of μ everywhere in the domain. For the recovered parameters, however, this is not the case — indeed, away from the boundary μ is generally larger than λ. For Figures 6.5 and 6.6 these problems are not as apparent. Although μ is again much to large compared to λ the ratio of the values of both in the interior of the domain to those on the boundary is conserved (except for in the regions where λ and μ are predicted to be negative).

6.1.3 Computational Issues

The 3-d simulation took approximately 10 hours to run on a standard desktop computer. Observe that this is significantly larger than the 2-d simulations in Chapter 5 (which took about half an hour on the same computer) even though the 2-d mesh contained more nodes (so the system of equations initially appears likely to be smaller). There are several reasons for this:

- There are three unknowns at each node in 3-d (the displacements in each of the three dimensions) compared to two in 2-d.
- Although there are more nodes in the 2-d mesh there are more elements in the 3-d mesh (and each element contains four nodes in 3-d to three in 2-d). This means that when additional nodes are added to allow the use of quadratic basis functions more must be added in 2-d. Hence, there are actually more nodes in the 3-d mesh used in the solution loop than in the equivalent 2-d mesh.
- In assembling the stiffness matrix the *elements* are looped over rather than the nodes. So, the assembly time depends on the number of elements rather than the number of nodes. Since in practice assembling the matrix takes substantially longer than inverting it this is a major problem in the 3-d case.
- More Gauss points are used per element in the 3-d case (27 vs 8). This means that longer is taken on each element of the 3-d mesh than on those of the 2-d mesh, exacerbating the problem of having more elements.
- There are more nodes per element in the 3-d mesh (10 vs 6). One entry in the stiffness matrix has to be updated per node, so this also required more time per element in

3-d than 2-d.

- The Jacobi matrix is 2-by-2 matrix in 2-d, but 3-by-3 in 3-d. This matrix needs to be inverted on each element when assembling the stiffness matrix — 3-by-3 matrix inversion takes longer than 2-by-2 inversion.
- There are three governing equations in 3-d, compared to two in 2-d. This means that more numerical integrations need to be performed in assembling the 3-d stiffness matrix.

The large amount of computational time required to simulate the 2-d forward problem is one major reason why 3-d elasticity imaging has not yet been developed (the other important one is the lack of 3-d images). This is an especially serious problem for the iterative schemes currently used - there is no way that the thousands of forward problems required could be solved in a reasonable time period (even using a more efficient programming language than Matlab).

Using a direct solution method, however, solving the 3-d inverse problem would be viable. 10 hours is less time than is currently required to produce a 2-d elasticity image iteratively and, in fact, the inverse problem requires less time to solve than the forward problem since linear basis functions can be used for the Lamé constants whereas quadratic ones are required for the displacements (since there is one fewer derivative on the Lamé constants). This means fewer nodes (both in total and per element) and fewer Gauss points are required to solve the inverse problem than the forward problem. There are also only two unknowns per node in the 2-d inverse problem, compared to three in the forward problem. So, both assembly and inversion of the global stiffness matrix can be achieved more quickly. In the 2-d case, the inverse problem takes two to three minutes to solve and the forward problem roughly half an hour (both on the mesh with 5248 elements and 2705 nodes). If the same speed-up is seen in 3-d then the inverse problem could be expected to take of the order of an hour to solve on a mesh containing of the order of 8880 elements and 1762 nodes.

6.1.4 Conclusions

In this section the validity of assuming that the tissue in the plane of an ultrasound image is undergoing plane strain has been investigated. In Section 6.1.1 the results of solving 2-d (plane strain) and 3-d linear elasticity forward problems with corresponding parameter

values were compared. It was seen that the solutions were not at all similar — in particular the 2-d deformation was symmetric whereas the 3-d one was not. Hence, the plane strain assumption is not valid for elasticity imaging. So, the deformation of soft tissue during the imaging process must be modelled in 3-d.

However, current methods that use the plane strain assumption *do* give sensible results. In Section 6.1.2 this was shown to be consistent with the invalidity of the plane strain assumption. The results of applying the 2-d parameter recovery method developed in Chapter 4 were shown to produce images from which the location of the stiffer regions of tissue can be identified (approximately). However, the results have much less quantitative value than the plane strain equivalents discussed in Chapter 5. In particular, it is impossible to use the results of the parameter recovery method on the cross-section to compare the stiffnesses of different regions of tissue (this was possible in the plane strain case). This means that the method would have little use in classifying regions of tissue as either cancerous or benign.

Finally, in Section 6.1.3 the computational requirements of the various simulations performed in this work (both forward and inverse problems) were discussed. Although solving the 3-d inverse problem numerically could be expected to take roughly an hour to solve on a mesh containing fewer than 9000 elements and 2000 nodes it was observed that this is still much less time than the iterative solution methods require to produce an image even when using the plane strain assumption. So, given 3-d ultrasound images, a 3-d parameter recovery method based on the theory of Chapter 4 would be viable. The development of such a method is discussed in the following section.

6.2 The 3-Dimensional Inverse Problem

In this section the theory underlying the parameter recovery method is extended to three dimensions. Much of the theory is analogous to that for the 2-d case presented in Chapter 4. In Section 6.2.1 the inverse problem is formulated in terms of a strain field resulting from the deformation of a material. This gives three equations for the two unknown Lamé constants, so the entire system is over-specified and cannot be classified. However, any two of the equations can be considered and in Section 6.2.2 it is shown that each of these pairs are hyperbolic.

6.2.1 Formulation of the Inverse Problem

Recall from Section 2.1.3.1 that in 3-d the governing equations of an isotropic linearly elasticity solid in equilibrium and subject to no external body forces are

$$0 = \frac{\partial}{\partial x_i}\left(\lambda \frac{\partial u_k}{\partial x_k}\right) + \frac{\partial}{\partial x_j}\left(\mu \frac{\partial u_i}{\partial x_j}\right) + \frac{\partial}{\partial x_j}\left(\mu \frac{\partial u_j}{\partial x_i}\right), \quad i = 1, 2, 3, \tag{6.5}$$

So, the inverse problem can be written in matrix form as

$$A \begin{pmatrix} \lambda_x \\ \mu_x \end{pmatrix} + B \begin{pmatrix} \lambda_y \\ \mu_y \end{pmatrix} + C \begin{pmatrix} \lambda_z \\ \mu_z \end{pmatrix} = \mathbf{d}, \tag{6.6}$$

where the subscripts denote partial differentiation and

$$A = \begin{pmatrix} \epsilon_{11} + \epsilon_{22} + \epsilon_{33} & 2\epsilon_{11} \\ 0 & 2\epsilon_{12} \\ 0 & 2\epsilon_{13} \end{pmatrix}, \tag{6.7}$$

$$B = \begin{pmatrix} 0 & 2\epsilon_{12} \\ \epsilon_{11} + \epsilon_{22} + \epsilon_{33} & 2\epsilon_{22} \\ 0 & 2\epsilon_{23} \end{pmatrix}, \tag{6.8}$$

$$C = \begin{pmatrix} 0 & 2\epsilon_{13} \\ 0 & 2\epsilon_{23} \\ \epsilon_{11} + \epsilon_{22} + \epsilon_{33} & 2\epsilon_{33} \end{pmatrix}, \tag{6.9}$$

$$\mathbf{d} = -\begin{pmatrix} \lambda(\epsilon_{11} + \epsilon_{22} + \epsilon_{33})_x + 2\mu((\epsilon_{11})_x + (\epsilon_{12})_y + (\epsilon_{13})_z) \\ \lambda(\epsilon_{11} + \epsilon_{22} + \epsilon_{33})_y + 2\mu((\epsilon_{12})_x + (\epsilon_{22})_y + (\epsilon_{23})_z) \\ \lambda(\epsilon_{11} + \epsilon_{22} + \epsilon_{33})_z + 2\mu((\epsilon_{13})_x + (\epsilon_{23})_y + (\epsilon_{33})_z) \end{pmatrix}. \tag{6.10}$$

Observe that this is a system of *three* scalar PDEs for the *two* unknowns λ and μ. Hence, it appears that the system is over-specified and so the problem is ill-posed for general ϵ. However, for a *real* strain field ϵ the three equations will be compatible (since the strain field is a solution of the forward problem) and so a solution will exist.

6.2.2 Classification of the Inverse Problem

Since the 3-d inverse problem contains three equations for the two unknown Lamé constants it cannot be classified as hyperbolic, parabolic or elliptic. However, any pair of the three equations can be classified. In this section it is shown that each pair of equations forms a hyperbolic system. Recall that in Section 2.2.1 definitions of hyperbolic, parabolic and elliptic systems for systems of equations in 2-d were given. However, in 3-d the situation is more complicated since the characteristic equations define surfaces rather than curves. So, before attempting to classify the 3-d elasticity inverse problem it is necessary to define exactly how this should be done. The classification of general PDEs in more than two dimensions is described in Section 6.2.2.1 before, in Section 6.2.2.2 any pair of equations from the 3-d inverse problem is shown to be hyperbolic.

6.2.2.1 Classification of Second-Order PDEs in More Than Two Dimensions

Suppose that a general second-order PDE in n-d of the form

$$\tfrac{1}{2}a_{ij}\frac{\partial^2 \mu}{\partial x_i \partial x_j} + \text{lower-order terms} = 0, \quad i,j = 1,\ldots,n. \tag{6.11}$$

is to be classified. The type of the equation depends on the signs of the eigenvalues of the symmetric matrix $A = (a_{ij})$. It is said to be:

(i) *Elliptic* if the eigenvalues of A are all non-zero and have the same sign.

(ii) *Parabolic* if exactly one of the eigenvalues of A is zero and all of the others have the same sign.

(iii) *Hyperbolic* if all of the eigenvalues are non-zero and there is either exactly one positive eigenvalue or exactly one negative eigenvalue.

(iv) *Ultra-hyperbolic* if all of the eigenvalues are non-zero, there is more than one positive eigenvalue and more than one negative eigenvalue.[2]

Recall that the eigenvalues of a real symmetric matrix are always real, so these definitions do make sense.

[2]Observe that a PDE defined in 3-d (leading to a 3-by-3 matrix) cannot be ultra-hyperbolic as there can be at most three eigenvalues — so there cannot be two positive *and* two negative eigenvalues.

6.2.2.2 Hyperbolicity of the 3-Dimensional Inverse Problem

In this section, the definitions given in the previous section are applied to the 3-d linear elasticity inverse problem. In order to do this, first consider the first two rows of equation (6.6), i.e.

$$\frac{\partial}{\partial x}(\lambda(\epsilon_{11}+\epsilon_{22}+\epsilon_{33})+2\mu\epsilon_{11})+2\frac{\partial}{\partial y}(\mu\epsilon_{12})+2\frac{\partial}{\partial z}(\mu\epsilon_{13})=0, \tag{6.12}$$

$$2\frac{\partial}{\partial x}(\mu\epsilon_{12})+\frac{\partial}{\partial y}(\lambda(\epsilon_{11}+\epsilon_{22}+\epsilon_{33})+2\mu\epsilon_{22})+2\frac{\partial}{\partial z}(\mu\epsilon_{23})=0. \tag{6.13}$$

Since the classification theory has been described for second-order PDEs rather than first-order systems it is necessary to eliminate one of the variables. Here, λ is eliminated by differentiating equation (6.12) with respect to y, differentiating equation (6.13) with respect to x and subtracting to give

$$\mathcal{L}_1(\mathbf{u}) \equiv \left(\frac{\partial^2}{\partial y^2}-\frac{\partial^2}{\partial x^2}\right)(\mu\epsilon_{12})+\frac{\partial^2}{\partial x\partial y}(\mu(\epsilon_{11}-\epsilon_{22}))+\frac{\partial^2}{\partial y\partial z}(\mu\epsilon_{13})-\frac{\partial^2}{\partial x\partial z}(\mu\epsilon_{23})=0. \tag{6.14}$$

Then, the matrix $A=(a_{ij})$ defined in equation (6.11) and used to classify the PDE is given by

$$A=\begin{pmatrix} -2\epsilon_{12} & \epsilon_{11}-\epsilon_{22} & -\epsilon_{23} \\ \epsilon_{11}-\epsilon_{22} & 2\epsilon_{12} & \epsilon_{13} \\ -\epsilon_{23} & \epsilon_{13} & 0 \end{pmatrix}. \tag{6.15}$$

Calculating the eigenvalues of this matrix leads to the following lemma showing the hyperbolicity of equation (6.14) and, hence, the system of first-order equations it is equivalent to:

Lemma 6.1. *Given* $\mathbf{u}(\mathbf{x})$ *and* $\epsilon(\mathbf{x})$ *computed from equation (6.5), if* ϵ *is non-zero and regular everywhere inside* Ω *then equation (6.14) (and, hence the system of equations (6.12) and (6.13)) is hyperbolic or it has either three zero eigenvalues or one positive, one negative and one zero eigenvalue (and so does not fit into any of the categories listed above).*

Proof. Here, the eigenvalues of A are calculated as follows:

$$0 = \det(A - kI), \tag{6.16}$$

$$= \det \begin{pmatrix} -2\epsilon_{12} - k & \epsilon_{11} - \epsilon_{22} & -\epsilon_{23} \\ \epsilon_{11} - \epsilon_{22} & 2\epsilon_{12} - k & \epsilon_{13} \\ -\epsilon_{23} & \epsilon_{13} & -k \end{pmatrix}, \tag{6.17}$$

$$\begin{aligned} = -(2\epsilon_{12} + k) \det \begin{pmatrix} 2\epsilon_{12} - k & \epsilon_{13} \\ \epsilon_{13} & -k \end{pmatrix} - (\epsilon_{11} - \epsilon_{22}) \det \begin{pmatrix} \epsilon_{11} - \epsilon_{22} & \epsilon_{13} \\ -\epsilon_{23} & -k \end{pmatrix} \\ - \epsilon_{23} \det \begin{pmatrix} \epsilon_{11} - \epsilon_{22} & 2\epsilon_{12} - k \\ -\epsilon_{23} & \epsilon_{13} \end{pmatrix}, \end{aligned} \tag{6.18}$$

$$\begin{aligned} = (2\epsilon_{12} + k)[-k(2\epsilon_{12} - k) - \epsilon_{13}^2] + (\epsilon_{11} - \epsilon_{22})[-k(\epsilon_{11} - \epsilon_{22}) + \epsilon_{13}\epsilon_{23}] \\ + \epsilon_{23}[\epsilon_{13}(\epsilon_{11} - \epsilon_{22}) + \epsilon_{23}(2\epsilon_{12} - k)], \end{aligned} \tag{6.19}$$

$$= k^3 - \big((\epsilon_{11} - \epsilon_{22})^2 + 4\epsilon_{12}^2 + \epsilon_{13}^2 + \epsilon_{23}^2\big)k + 2\epsilon_{12}\big(\epsilon_{23}^2 - \epsilon_{13}^2\big) + \epsilon_{13}\epsilon_{23}(\epsilon_{11} - \epsilon_{22}). \tag{6.20}$$

This equation is a *depressed cubic* for k (i.e. the coefficient of k^2 is zero). Suppose that it is rewritten as

$$k^3 - pk + q = 0, \tag{6.21}$$

for appropriately chosen $p \geqslant 0$, q. Observe first that $p > 0$ unless $\epsilon_{11} = \epsilon_{22}$, $\epsilon_{12} = \epsilon_{13} = \epsilon_{23} = 0$ (since the strain components must be real-valued), in which case $p = 0$. If these conditions are met then it can be seen that $q = 0$ also. Hence, if $p = 0$, the eigenvalue equation reduces to $k^3 = 0$ and all three eigenvalues are zero.

Suppose now that $p > 0$ and recall that A has three real eigenvalues (since it is real and symmetric), so equation (6.21) must have two turning points. These turning points satisfy the quadratic equation

$$3k^2 - p = 0, \tag{6.22}$$

obtained by differentiating equation (6.21) with respect to k. So, they are located at

$$k = \pm\sqrt{\frac{p}{3}}. \tag{6.23}$$

Observe that, since p is non-zero these k-coordinates are distinct. Further, one is positive and one is negative. For a cubic equation with three real roots, $r_1 \leqslant r_2 \leqslant r_3$, and two turning points, located at $k = k_1$ and $k = k_2$ where $k_1 < k_2$, one root must be located

between the turning points, one to the left of the smallest turning point and one to the right of the largest turning point, i.e. $r_1 \leqslant k_1 \leqslant r_2 \leqslant k_2 \leqslant r_3$. So, in this case, where $k_1 < 0$ and $k_2 > 0$ it can be seen that $r_1 < 0$ and $r_3 > 0$, i.e. there is at least one negative eigenvalue and at least one positive eigenvalue. Further, $r_2 = 0$ if and only if $q = 0$.

If $r_2 < 0$ then there is exactly one positive eigenvalue and hence the system defined by equations (6.12) and (6.13) is hyperbolic. If $r_2 > 0$ then there is exactly one negative eigenvalue and, again, the system is hyperbolic. Finally, if $r_2 = 0$ then there is one positive, one negative and one zero eigenvalue. □

Suppose instead that the classification of the system consisting of the first and third rows of equation (6.6) is to be considered. Then the equivalent of equation (6.14) is

$$\mathcal{L}_2(\mathbf{u}) \equiv \left(\frac{\partial^2}{\partial z^2} - \frac{\partial^2}{\partial x^2}\right)(\mu\epsilon_{13}) + \frac{\partial^2}{\partial x \partial z}(\mu(\epsilon_{11} - \epsilon_{33})) + \frac{\partial^2}{\partial y \partial z}(\mu\epsilon_{23}) - \frac{\partial^2}{\partial x \partial y}(\mu\epsilon_{12}) = 0. \quad (6.24)$$

Observe that this equation is identical to equation (6.14), except that the roles of y and z have been interchanged (note that this means that some of the subscripts in the strain tensor must also be changed, but the exact components of the strain tensor in the coefficients do not affect the classification). Hence, this system must also be hyperbolic (or have positive, negative and zero eigenvalues). Similarly, if the second and third rows of equation (6.6) are used the resulting second-order equation for μ is

$$\mathcal{L}_3(\mathbf{u}) \equiv \left(\frac{\partial^2}{\partial z^2} - \frac{\partial^2}{\partial y^2}\right)(\mu\epsilon_{23}) + \frac{\partial^2}{\partial y \partial z}(\mu(\epsilon_{22} - \epsilon_{33})) + \frac{\partial^2}{\partial x \partial z}(\mu\epsilon_{12}) - \frac{\partial^2}{\partial x \partial y}(\mu\epsilon_{13}) = 0. \quad (6.25)$$

Again, this equation must be hyperbolic (or have positive, negative and zero eigenvalues) as it is equivalent to equation (6.14), but with x, y and z permuted.

6.3 Multiple Displacement Fields

Recall from Section 4.3 that in the 2-d plane strain case information from multiple displacement fields was used in order to reduce the amount of boundary data that was required in order to solve the inverse problem. Recall further that the major problem with specifying boundary data was that it is impossible to have *a priori* knowledge of the tissue parameter values on the boundary. This problem is still present in 3-d. So, in this section

the effect of using multiple displacement fields in the 3-d inverse problem is considered. First, in Section 6.3.1 the case of two displacement fields is considered. In 2-d using two displacement fields restricted the dimension of the solution space to be at most four. In 3-d it is shown that using two displacement fields restricts the dimension of the solution space to be at most three. Then, in Section 6.3.2 the case of three displacement fields is considered. This is shown to be equivalent to having four fields in 2-d, i.e. it gives enough information to determine the solution up to a multiplicative constant.

6.3.1 Two Displacement Fields

In this section the solution space of the inverse problem for two displacement fields, defined by the equations

$$\mathcal{L}_j(\mathbf{u}^{(i)})\mu = 0, \quad i = 1, 2, \quad j = 1, 2, 3, \tag{6.26}$$

where $\mathcal{L}_j$, $j = 1, 2, 3$ are defined by equations (6.14) – (6.25), is shown to be at most 3-dimensional. Analogously to the analysis of Section 4.3.1 for the plane strain case, this is done by showing that the Cauchy data must specify a third-order ODE. This is shown to be the case in the following lemma:

Lemma 6.2. *For a solution to equation (6.26) to exist, the Cauchy data specified must satisfy a given third order ODE.*

Proof. Without loss of generality, suppose that the coordinate system is chosen so that the x-axis is aligned with the C, the curve on which the Cauchy data is prescribed. Then the Cauchy data can be written as

$$\mu(x, 0, 0) = \bar{\mu}(x), \tag{6.27}$$

$$\frac{\partial \mu}{\partial y}(x, 0, 0) = \bar{\mu}_y(x), \tag{6.28}$$

$$\frac{\partial \mu}{\partial z}(x, 0, 0) = \bar{\mu}_z(x), \tag{6.29}$$

To prove the lemma the strategy is to show that $\bar{\mu}$ satisfies a given third order ODE and then that $\bar{\mu}_y$ and $\bar{\mu}_z$ are known functions of $\bar{\mu}$ and its x-derivatives.

The proof follows that of Lemma 4.3 by eliminating the y- and z-derivatives between the six PDEs to obtain an ODE in x. First, consider the terms $(\mu_{xx} - \mu_{yy})$, $(\mu_{xx} - \mu_{zz})$, $(\mu_{yy} - \mu_{zz})$, μ_{xy}, μ_{xz}, μ_{yz} to be linearly independent variables in the system defined by

equation (6.26). Then, any five of these can be eliminated from the system to leave a single equation defining them in terms of μ and its first derivatives. Doing this for each in turn gives the six PDEs:

$$a^{(1)}(\mu_{xx}-\mu_{yy})+b^{(1)}\mu_x+c^{(1)}\mu_y+d^{(1)}\mu_z+e^{(1)}\mu=0, \tag{6.30}$$

$$a^{(2)}(\mu_{xx}-\mu_{zz})+b^{(2)}\mu_x+c^{(2)}\mu_y+d^{(2)}\mu_z+e^{(2)}\mu=0, \tag{6.31}$$

$$a^{(3)}(\mu_{yy}-\mu_{zz})+b^{(3)}\mu_x+c^{(3)}\mu_y+d^{(3)}\mu_z+e^{(3)}\mu=0, \tag{6.32}$$

$$a^{(4)}\mu_{xy}+b^{(4)}\mu_x+c^{(4)}\mu_y+d^{(4)}\mu_z+e^{(4)}\mu=0, \tag{6.33}$$

$$a^{(5)}\mu_{xz}+b^{(5)}\mu_x+c^{(5)}\mu_y+d^{(5)}\mu_z+e^{(5)}\mu=0, \tag{6.34}$$

$$a^{(6)}\mu_{yz}+b^{(6)}\mu_x+c^{(6)}\mu_y+d^{(6)}\mu_z+e^{(6)}\mu=0. \tag{6.35}$$

Next, the second derivatives of μ except for μ_{xx} are eliminated. This is done by cross-differentiating pairs of the equations (6.30) – (6.35) and adding. For example,

$$a^{(4)}\frac{\partial(6.30)}{\partial x}+a^{(1)}\frac{\partial(6.33)}{\partial y} \quad\Rightarrow\quad a^{(7)}\mu_{xxx}+b^{(7)}\mu_{xx}+c^{(7)}\mu_x+d^{(7)}\mu_y+e^{(7)}\mu_z+f^{(7)}\mu=0, \tag{6.36}$$

where this equation has been simplified by using equations (6.33) – (6.35) to rewrite μ_{xy}, μ_{xz} and μ_{yz} in terms of μ and its first derivatives and equation (6.30) used to rewrite μ_{yy} in terms of μ, its first derivatives and μ_{xx}. Similarly,

$$a^{(5)}\frac{\partial(6.31)}{\partial x}+a^{(2)}\frac{\partial(6.34)}{\partial z} \quad\Rightarrow\quad a^{(8)}\mu_{xxx}+b^{(8)}\mu_{xx}+c^{(8)}\mu_x+d^{(8)}\mu_y+e^{(8)}\mu_z+f^{(8)}\mu=0, \tag{6.37}$$

$$a^{(6)}\frac{\partial(6.32)}{\partial y}+a^{(3)}\frac{\partial(6.35)}{\partial z} \quad\Rightarrow\quad a^{(9)}\mu_{yyy}+b^{(9)}\mu_{xx}+c^{(9)}\mu_x+d^{(9)}\mu_y+e^{(9)}\mu_z+f^{(9)}\mu=0, \tag{6.38}$$

where again second derivatives have been rewritten in terms of μ, its first derivatives and μ_{xx} using equations (6.30) – (6.35).

In order to eliminate the μ_{yyy} term from equation (6.38) consider the following:

$$a^{(1)}\frac{\partial(6.33)}{\partial x}-a^{(4)}\frac{\partial(6.33)}{\partial y} \quad\Rightarrow\quad a^{(10)}\mu_{yyy}+b^{(10)}\mu_{xx}+c^{(10)}\mu_x+d^{(10)}\mu_y+e^{(10)}\mu_z+f^{(10)}\mu=0, \tag{6.39}$$

where equations (6.30) – (6.35) have again been used to eliminate the second derivatives

of μ except for μ_{xx}. So, equation (6.38) can be rewritten as

$$b^{(11)}\mu_{xx} + c^{(11)}\mu_x + d^{(11)}\mu_y + e^{(11)}\mu_z + f^{(11)}\mu = 0, \tag{6.40}$$

Then, equations (6.37) and (6.40) can be used to eliminate μ_y and μ_z from equation (6.36), giving

$$a^{(12)}\mu_{xxx} + b^{(12)}\mu_{xx} + c^{(12)}\mu_x + d^{(12)}\mu = 0, \tag{6.41}$$

which is a third-order ODE for μ that holds throughout the domain Ω for any μ satisfying the system defined by equation (6.26). In particular, it must be satisfied by the Cauchy data on $x = 0$. Hence, $\bar{\mu}(x)$ is a solution of equation (6.41). Further, equations (6.37) and (6.40) can be used to determine $\bar{\mu}_y(x)$ and $\bar{\mu}_z(x)$ in terms of $\bar{\mu}(x)$. This completes the proof. □

An immediate corollary of the above lemma is that the Cauchy data can contain at most three arbitrary constants (since it is the solution of a third-order ODE). Since Cauchy data determines the solution of a hyperbolic system of PDEs uniquely, the solution of equation (6.26) also contains at most three arbitrary constants. So, if two linearly independent displacement fields are given, the 3-d inverse problem has a solution space of dimension at most three.

6.3.2 Three Displacement Fields

If three displacement fields are given then an exact solution for μ can be written down in terms of one arbitrary constant and integrals of known functions. This is analogous to the analysis leading to Theorem 4.2 for the 2-d plane strain case.

Theorem 6.1. *Given three linearly independent, mutually compatible equilibrium displacement fields* $\mathbf{u}^{(i)}$, $i = 1, 2, 3$, *then the solution* $\mu(x, y, z)$ *of the set of equations*

$$\mathcal{L}_j(\mathbf{u}^{(i)})\mu = 0, \quad i, j = 1, 2, 3, \tag{6.42}$$

is unique up to a multiplicative constant. Further, the solution is given by

$$\begin{aligned}\mu(x,y,z) = \mu(x_0,y_0,z_0)\exp\Bigg[&\frac{4}{3}\int_{z_0}^{z}\int_{y_0}^{y}\int_{x_0}^{x}\left(\frac{\partial^2 P}{\partial y'\partial z'}+\frac{\partial^2 Q}{\partial x'\partial z'}+\frac{\partial^2 R}{\partial x'\partial y'}\right)dx'dy'dz' \\ &-\frac{4}{3}\int_{y_0}^{y}\int_{x_0}^{x}\left(\frac{\partial Q}{\partial x'}+\frac{\partial P}{\partial y'}\right)dx'dy' \\ &-\frac{4}{3}\int_{z_0}^{z}\int_{x_0}^{x}\left(\frac{\partial R}{\partial x'}+\frac{\partial P}{\partial z'}\right)dx'dz' \\ &-\frac{4}{3}\int_{z_0}^{z}\int_{y_0}^{y}\left(\frac{\partial R}{\partial y'}+\frac{\partial Q}{\partial z'}\right)dy'dz' \\ &+\int_{z_0}^{z}R\,dz'+\int_{y_0}^{y}Q\,dy'+\int_{x_0}^{x}P\,dx'\Bigg],\end{aligned} \tag{6.43}$$

for some given functions $P(x,y,z)$, $Q(x,y,z)$ and $R(x,y,z)$.

Proof. Observe that the system of equations (6.42) is homogeneous system of nine linear equations for ten independent variables (μ, its three first derivatives and six second derivatives). The coefficients of these variables depend on position (i.e. on x, y and z). So, any eight of the equations can be used to eliminate all of the second derivatives and two of the first derivatives from the ninth equation. In this way the system can be reduced to the following three first-order partial differential equations for μ:

$$\frac{\partial\mu}{\partial x}+P(x,y,z)\mu=0, \tag{6.44}$$

$$\frac{\partial\mu}{\partial y}+Q(x,y,z)\mu=0, \tag{6.45}$$

$$\frac{\partial\mu}{\partial z}+R(x,y,z)\mu=0. \tag{6.46}$$

In order for these three PDEs to have a solution it is necessary for the following solvability conditions to be satisfied:

$$\frac{\partial P}{\partial y}=\frac{\partial Q}{\partial x}, \tag{6.47}$$

$$\frac{\partial P}{\partial z}=\frac{\partial R}{\partial x}, \tag{6.48}$$

$$\frac{\partial Q}{\partial z}=\frac{\partial R}{\partial y}. \tag{6.49}$$

Since the three displacement fields are compatible there must be a μ that satisfies equations (6.44) – (6.46). Hence, the solvability conditions must hold. In this case the joint solution of equations (6.44) – (6.46) is given by equation (6.43). As in the proof of Theorem 4.2 this can be verified by differentiating equation (6.43) using the Leibniz integral rule. For example, differentiating the triple integral with respect to x gives

$$\begin{aligned}\frac{\partial}{\partial x}\left[\int_{z_0}^{z}\int_{y_0}^{y}\int_{x_0}^{x}\left(\frac{\partial^2 P}{\partial y'\partial z'}+\frac{\partial^2 Q}{\partial x'\partial z'}+\frac{\partial^2 R}{\partial x'\partial y'}\right)dx'dy'dz'\right.& \\ =\int_{z_0}^{z}\int_{y_0}^{y}\left(\frac{\partial^2 P}{\partial y'\partial z'}+\frac{\partial^2 Q}{\partial x\partial z'}+\frac{\partial^2 R}{\partial x\partial y'}\right)dy'dz',& \qquad (6.50)\\ =P+\int_{z_0}^{z}\int_{y_0}^{y}\left(\frac{\partial^2 Q}{\partial x\partial z'}+\frac{\partial^2 R}{\partial x\partial y'}\right)dy'dz'.& \qquad (6.51)\end{aligned}$$

The x-derivatives of the double integrals are (on applying the solvability conditions)

$$-\frac{\partial}{\partial x}\int_{y_0}^{y}\int_{x_0}^{x}\left(\frac{\partial Q}{\partial x'}+\frac{\partial P}{\partial y'}\right)dx'dy'=-\int_{y_0}^{y}\left(\frac{\partial Q}{\partial x}+\frac{\partial P}{\partial y'}\right)dy'=-2\int_{y_0}^{y}\frac{\partial P}{\partial y'}dy'=-2P, \qquad (6.52)$$

$$-\frac{\partial}{\partial x}\int_{z_0}^{z}\int_{x_0}^{x}\left(\frac{\partial R}{\partial x'}+\frac{\partial P}{\partial z'}\right)dx'dz'=-\int_{z_0}^{z}\left(\frac{\partial R}{\partial x}+\frac{\partial P}{\partial z'}\right)dz'=-2\int_{z_0}^{z}\frac{\partial P}{\partial z'}dz'=-2P, \qquad (6.53)$$

$$-\frac{\partial}{\partial x}\int_{z_0}^{z}\int_{y_0}^{y}\left(\frac{\partial R}{\partial y'}+\frac{\partial Q}{\partial z'}\right)dy'dz'=-\int_{z_0}^{z}\int_{y_0}^{y}\left(\frac{\partial^2 R}{\partial x\partial y'}+\frac{\partial^2 Q}{\partial x\partial z'}\right)dy'dz'. \qquad (6.54)$$

Observe that the right-hand-side of equation (6.54) cancels with the integral term on the right-hand-side of equation (6.51) when the derivative of equation (6.43) is calculated. Finally, the derivatives of the single integrals in equation (6.43) with respect to x are (again, applying the solvability conditions)

$$\frac{\partial}{\partial x}\int_{z_0}^{z}Rdy'=\int_{z_0}^{z}Rdz'=\int_{z_0}^{z}\frac{\partial R}{\partial x}dz'=\int_{z_0}^{z}\frac{\partial P}{\partial z'}dz'=P, \qquad (6.55)$$

$$\frac{\partial}{\partial x}\int_{z_0}^{z}Rdy'=\int_{y_0}^{y}Qdy'=\int_{y_0}^{y}\frac{\partial Q}{\partial x}dy'=\int_{y_0}^{y}\frac{\partial P}{\partial y'}dy'=P, \qquad (6.56)$$

$$\frac{\partial}{\partial x}\int_{x_0}^{x}Pdx'=P. \qquad (6.57)$$

Putting all of these results together to differentiate equation (6.43) with respect to x leads

to

$$\frac{\partial \mu}{\partial x} = \left(\tfrac{4}{3}P - \tfrac{8}{3}P - \tfrac{8}{3}P + P + P + P\right)\mu, \tag{6.58}$$

$$= -P\mu, \tag{6.59}$$

which is equivalent to equation (6.44). Equivalent results can be derived for the y- and z-derivatives to show that equations (6.45) and (6.46) hold. □

Further, it can be shown (again analogously to the 2-d problem) that no more than three independent displacement fields can be given in 3-d. So, we can do no better than determining the solution in terms of a single multiplicative constant. This is demonstrated in the following lemma:

Lemma 6.3. *The maximum number of linearly independent, mutually compatible equilibrium displacement fields that can exist in an isotropic, linearly elastic 3-d material is three.*

Proof. The proof is similar to that of Lemma 4.4 in 2-d. Suppose that in addition to the set of equations (6.42) the three equations

$$\mathcal{L}_j(\mathbf{u}^{(4)})\mu = 0, \quad j = 1, 2, 3, \tag{6.60}$$

derived from a fourth displacement field $\mathbf{u}^{(4)}$ also hold. Then these form a set of twelve linear equations for the ten unknowns (recall that these are μ, its three first derivatives and six second derivatives). Since the displacement fields are assumed to be mutually compatible a nontrivial solution exists. Hence the coefficient matrix formed by selecting any ten of the equations must be singular, contradicting the assumed linear independence of the displacement fields. So, no four mutually compatible displacement fields can be linearly independent in 3-d. □

6.4 The Minimization Problem

Since the system defined by equation (6.6) consists of three equations for the two unknown Lamé constants it cannot be directly solved using numerical methods (since the resulting matrices will not be square and so cannot be inverted) even in the case of a real strain field

for which the equations are compatible. Further, it has been shown that a system consisting of any two of the components of equation (6.6) will be hyperbolic and, so, difficult to solve numerically. Together these two observations suggest that (as in the 2-d case) the best way to proceed is to reformulate the system as a minimization problem analogously to the 2-d case described in Section 4.4.1.

With this in mind, define the residuals

$$R_1 = \frac{\partial}{\partial x}(\lambda(\epsilon_{11} + \epsilon_{22} + \epsilon_{33}) + 2\mu\epsilon_{11}) + 2\frac{\partial}{\partial y}(\mu\epsilon_{12}) + 2\frac{\partial}{\partial z}(\mu\epsilon_{13}), \tag{6.61}$$

$$R_2 = 2\frac{\partial}{\partial x}(\mu\epsilon_{12}) + \frac{\partial}{\partial y}(\lambda(\epsilon_{11} + \epsilon_{22} + \epsilon_{33}) + 2\mu\epsilon_{22}) + 2\frac{\partial}{\partial z}(\mu\epsilon_{23}), \tag{6.62}$$

$$R_3 = 2\frac{\partial}{\partial x}(\mu\epsilon_{13}) + 2\frac{\partial}{\partial y}(\mu\epsilon_{23}) + \frac{\partial}{\partial z}(\lambda(\epsilon_{11} + \epsilon_{22} + \epsilon_{33}) + 2\mu\epsilon_{33}). \tag{6.63}$$

Then, if λ and μ satisfy the system of equations (6.6) each of these residuals will vanish. Hence, the solution of the system can be reformulated as the following minimization problem:

$$\text{Find } \lambda,\ \mu \text{ to minimize } I = \int_\Omega \left(R_1^2 + R_2^2 + R_3^2\right) dxdydz. \tag{6.64}$$

Recall from Section 4.4.2 that in the 2-d case it was necessary to include a Lagrange multiplier term to ensure that the trivial case $\lambda \equiv \mu \equiv 0$ is not a minimizer. As in the 2-d case this is achieved by imposing the constraint

$$J(\lambda) = \int_\Omega \lambda dxdydz = 1, \tag{6.65}$$

and adding a Lagrange multiplier to the integral in equation (6.64) so that

$$I = \int_\Omega \left(R_1^2 + R_2^2 + R_3^2 - k\lambda\right) dxdydz. \tag{6.66}$$

This minimization problem can be solved directly using a finite element method analogous to that described for the 2-d plane strain case in Section 4.4.4.

Recall that in the 2-d case it was necessary either to specify the values of the Lamé constants on the boundary of Ω or to use four displacement fields in order to ensure that the minimization problem had a unique solution. Similarly, in the 3-d case extra information is required in order to guarantee the uniqueness of the solution. This can take one of the following forms:

(i) The imposition of boundary conditions on the Lamé constants. As in the 2-d case, appropriate boundary data is to specify λ and μ everywhere on $\partial\Omega$.

(ii) The use of multiple displacement fields. In 2-d four fields were required. However, in 3-d Theorem 6.1 shows that *three* displacement fields are enough.

The finite element solution method is not implemented in this work. This is because the method has already been shown to work in principle in the 2-d case (see the results of Chapter 5), which requires far fewer computing resources, and there are no 3-d ultrasound images on which to validate it.

Chapter 7

Validation: Real Ultrasound Images

In this chapter the parameter recovery method developed in Chapter 4 is tested in preliminary experiments on real ultrasound images. The images have been acquired by Mike Kadour and other members of the Wolfson Medical Vision Laboratory, Engineering Department, University of Oxford (WMVL) as part of a clinical trial of a new device for producing elasticity images at the Breast Care Unit of the Churchill Hospital in Oxford. This new device consists of an ultrasound probe mounted on a moving platform. The platform can be moved up and down with a high degree of precision to generate a series of images of the same region of tissue under varying amounts of compression.

The ultrasound probe and its associated software stores a large quantity of data relating to each series of images. This data is described in Section 7.1. In Section 7.2 the results of using the parameter recovery method on a variety of different series of images are described. The images are of a gelatin phantom with a stiff cylindrical inclusion, a scar and an invasive tumour. It was expected that the inclusion in the phantom would be the easiest feature to detect, followed by the scar and then the tumour. This is due to a combination of the sharpness of the interfaces (interfaces between tumour and healthy tissue are not usually well-defined) and the regularity of the shape of the feature (tumours are usually very irregularly shaped).

7.1 Format of the Input Data

In this section the data stored by the imaging device is described. The data is made up of three parts: a series of two or more RF-mode ultrasound images, one or more displacement estimates and strain fields corresponding to the displacement estimates. For this work the displacement fields are the most important part of the data as they are used as the input for the parameter recovery method. Here, the RF-mode images and displacement estimates are described. The strain data is not used in this work. A software package (written by members of the WMVL) is also provided with the data in order to facilitate its analysis. Where relevant some of the functionality of this package is also discussed.

7.1.1 RF-Mode Images

In this section the format of the RF-mode ultrasound data is described, together with the software provided to convert it to a B-mode image for visualization. Usually a series of RF-mode images are provided, each stored as a matrix. Each column of the matrix corresponds to a single element of the (linear) ultrasound array. The elements are uniformly spaced along the array with a separation of approximately 0.15 mm between neighbouring elements. Within a column, the entries correspond to samples of the received wave. The exact frequency of the samples varies from image to image, but in a typical image, the wave is sampled so that successive rows of the matrix correspond to a difference in depth of 0.02 mm within the tissue.

Within the associated software package a Matlab function `rf2b` converts the RF-mode data to B-mode. This essentially involves envelope detection of the RF-mode waves. Similarly to RF-mode, the B-mode data is stored as a matrix with columns representing the elements of the ultrasound array. The B-mode image can then be displayed using Matlab's built-in `imagesc` function and selecting a grey-scale colourmap. During this image processing step some of the spatial resolution in the z-direction (distance into the tissue) is lost. Again, the exact resolution varies from image to image, but in the image considered in the previous paragraph, successive rows of the B-mode matrix represent a difference of approximately 0.08 mm in depth within the tissue (compared to 0.02 in the RF-mode matrix). The resolution in the x-direction is unchanged.

7.1.2 Displacement Data

The second component of the experimental data is the estimated displacement field. This is the part of the data that will be used in Section 7.2 as input data for the parameter recovery method of Chapter 4. The displacement data consists primarily of two structures defining the position of the points at which the displacement was estimated (in (x, z)-coordinates) and the components of displacement in each of these directions at each point. The data also contains a measure of the confidence in the displacement estimates at each point. Currently this is not used in the parameter recovery method, but it would be a logical future extension to include it.

The displacement estimates are given on a uniform grid that needs to be converted to a finite element mesh in order to use the parameter recovery method. This is done by constructing a regular mesh around the points at which the displacement estimates are given. 3-by-3 blocks of these points are converted into two triangular elements: the first element consists of the six nodes closest to the bottom right of the block and the second the six nodes closest to the top left (note that the three nodes on the bottom left to top right diagonal are in both elements). The mesh is formed as a 2-d array of these blocks covering the whole of the region of interest (see Figure 7.1).

7.2 Results

In this section the results of applying the parameter recovery method described in Chapter 4 to data of the type discussed in Section 7.1 are described. Three series of images are used, representing a gelatin phantom in Section 7.2.2 and two different types of feature within real tissue in Sections 7.2.3 and 7.2.4. Before describing the results of applying the parameter recovery method to these series of ultrasound images a discussion of the application of either boundary data or multiple displacement fields is given in Section 7.2.1.

7.2.1 Boundary Data and Multiple Displacement Fields

Recall from Chapter 4 that in order for a unique solution to the (plane strain) linear elasticity inverse problem for one displacement field to exist boundary conditions must be

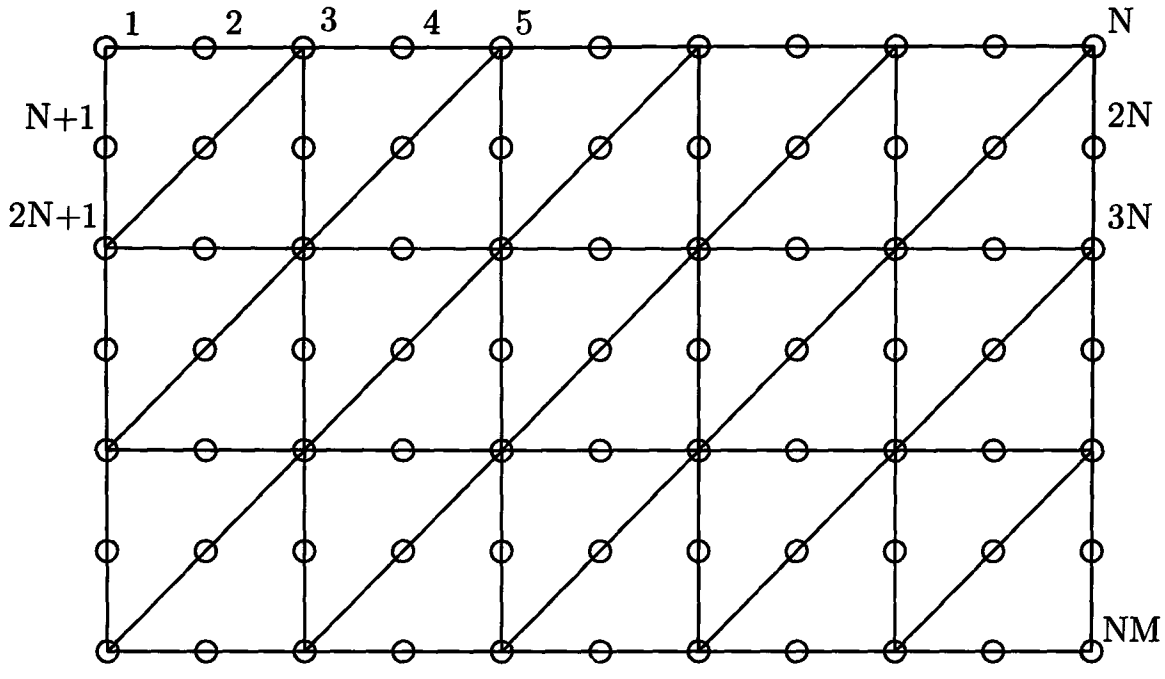


Figure 7.1: An example of the type of regular mesh used to recover the stiffness parameters for displacement fields generated from real ultrasound images. The nodes form a rectangular array with N nodes in the x-direction and M in the y direction. The origin is at the top left corner and nodes are numbered within rows (from left to right) starting with the top row. Each element contains six nodes since quadratic basis functions are used to interpolate the strain on the element. However, only the three corner nodes are used to calculate the parameter values as the finite element method uses linear basis functions.

specified. Alternatively, using four linearly independent, compatible displacement fields gives a solution that is unique up to a multiplicative constant without the need to specify boundary conditions. Unfortunately, for the data available here neither of these methods can be used satisfactorily. Clearly boundary data is generally unavailable for real tissue so multiple displacement fields must be used. However, although there are a series of displacement estimates given they are not linearly independent. This is because in linear elasticity if the boundary displacements are multiplied by a factor k then so are all of the displacements throughout the region of interest, i.e. all displacement fields due to a compression by an ultrasound probe are proportional to one another and, hence, linearly dependent.

In this work the parameter recovery method is tested using both boundary conditions and an integral constraint. In the case of the integral constraint only one displacement field is used, so the calculated solution is not necessarily unique. Although this is not an ideal situation it reflects the type of data that is likely to be given in a real clinical situation so it is worth evaluating how the method performs with only one linearly independent

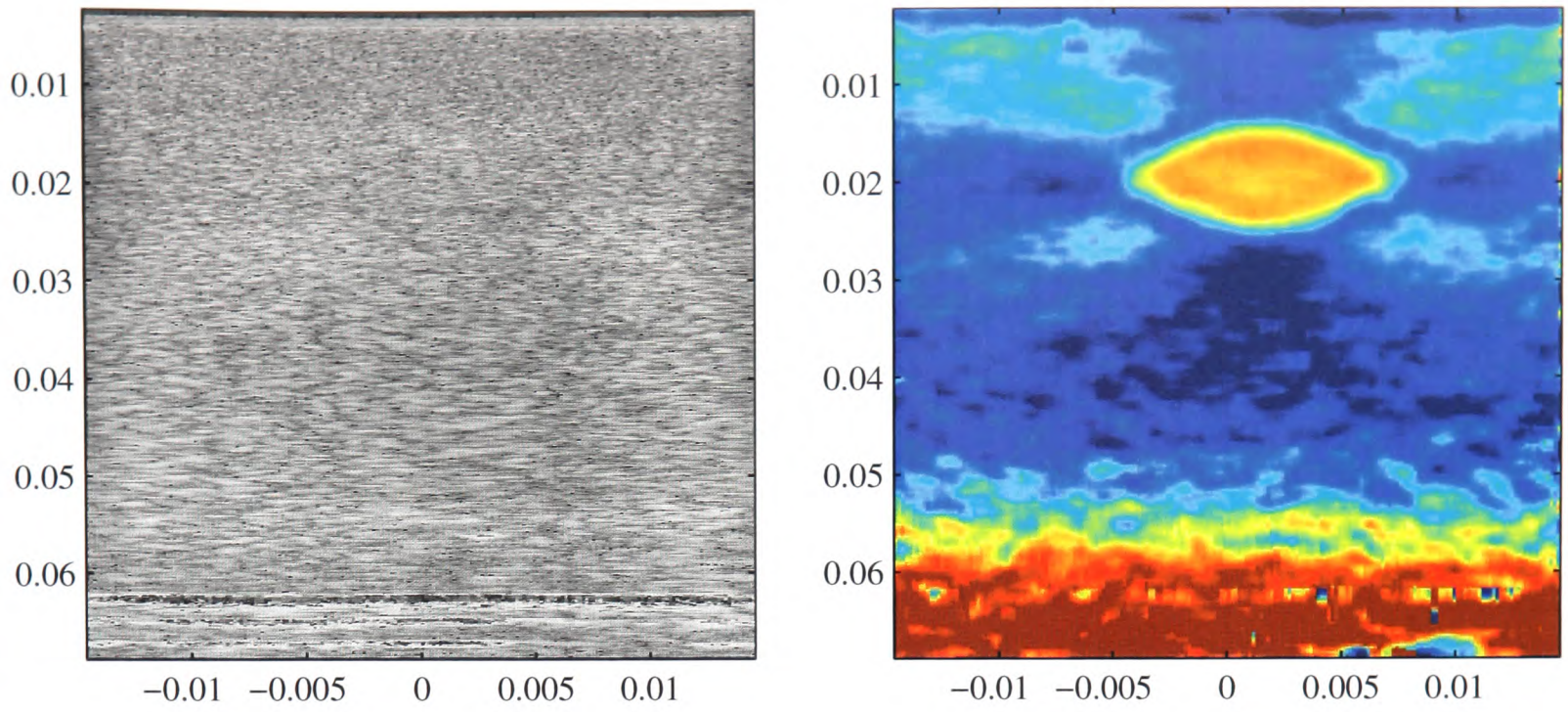


Figure 7.2: Ultrasound (left) and strain (right) images of a gelatin phantom with a cylindrical inclusion. The ultrasound probe is located at the top of the image and the numbers along the axes represent lengths in metres. In the strain image, regions where the (axial) strain is lower (i.e. stiffer regions) are shown in red, regions of higher strain in blue.

displacement field available. It turns out that the method works reasonably well in this situation. However, it does not work well at all if boundary conditions are imposed: the recovered solution rapidly approaches zero away from the boundary. A discussion of why this may be the case is given in Section 7.2.2.2.

7.2.2 Gelatin Phantom

The first series of images that the method is tested on were taken of a gelatin phantom. The phantom is made up of a cuboid of gelatin containing a stiff cylindrical inclusion running through its centre perpendicular to the plane of the ultrasound image. An image typical of the series is shown in Figure 7.2. The ultrasound probe is located at the top of the image and the inclusion has diameter 10 mm and its centre is located a depth of 20 mm below the centre of the ultrasound probe (see Figure 7.3 for a schematic representation of the phantom). The inclusion cannot be seen in the raw ultrasound image. A strain image generated from the series of ultrasound images is also shown in Figure 7.2. The inclusion can be seen as the region of low strain (red) in the expected location.

Gelatin has the advantage of being homogeneous (as opposed to real tissue which is very

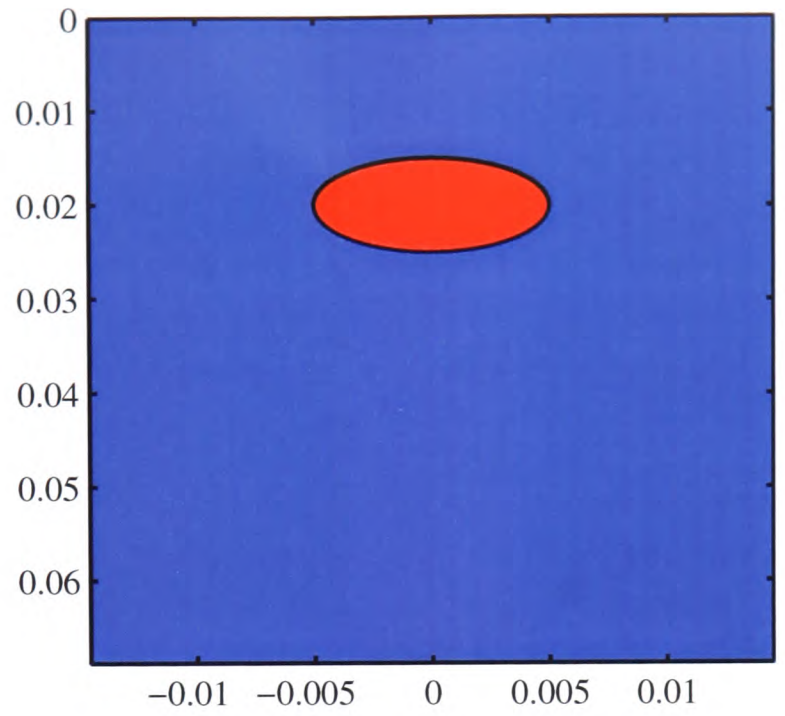


Figure 7.3: The stiffness distribution within the gelatin phantom. Values along the axes represent lengths in metres. The region coloured red is an inclusion that is three times stiffer than the rest of the phantom (coloured blue). The inclusion is circular, with diameter 0.01 m and is located 0.02 m below the top surface of the gelatin.

inhomogeneous) so the images, and, hence, displacement estimates, are expected to be much cleaner. The phantom is intended to be an intermediate step between the idealized displacement fields of Chapter 5 and the displacement fields estimated from images of real tissue in Sections 7.2.3 and 7.2.4. The parameter recovery method is tested using an integral constraint in Section 7.2.2.1 and with applied boundary conditions in Section 7.2.2.2. In both cases only one displacement field is used since no more are available. In both sections the displacement field used results from an estimate obtained by considering two images, one taken with an applied displacement of 20 mm and the other with an applied displacement of 25 mm.

7.2.2.1 Integral Constraint

The results of applying the parameter recovery method using a single displacement field and applying an integral constraint to guarantee that the solution is non-zero are shown in Figure 7.4, with the recovered λ shown on the left and μ on the right. The chosen constraint is

$$\int_{\Omega} \lambda(x,z)dxdz = \text{Area of } \Omega, \tag{7.1}$$

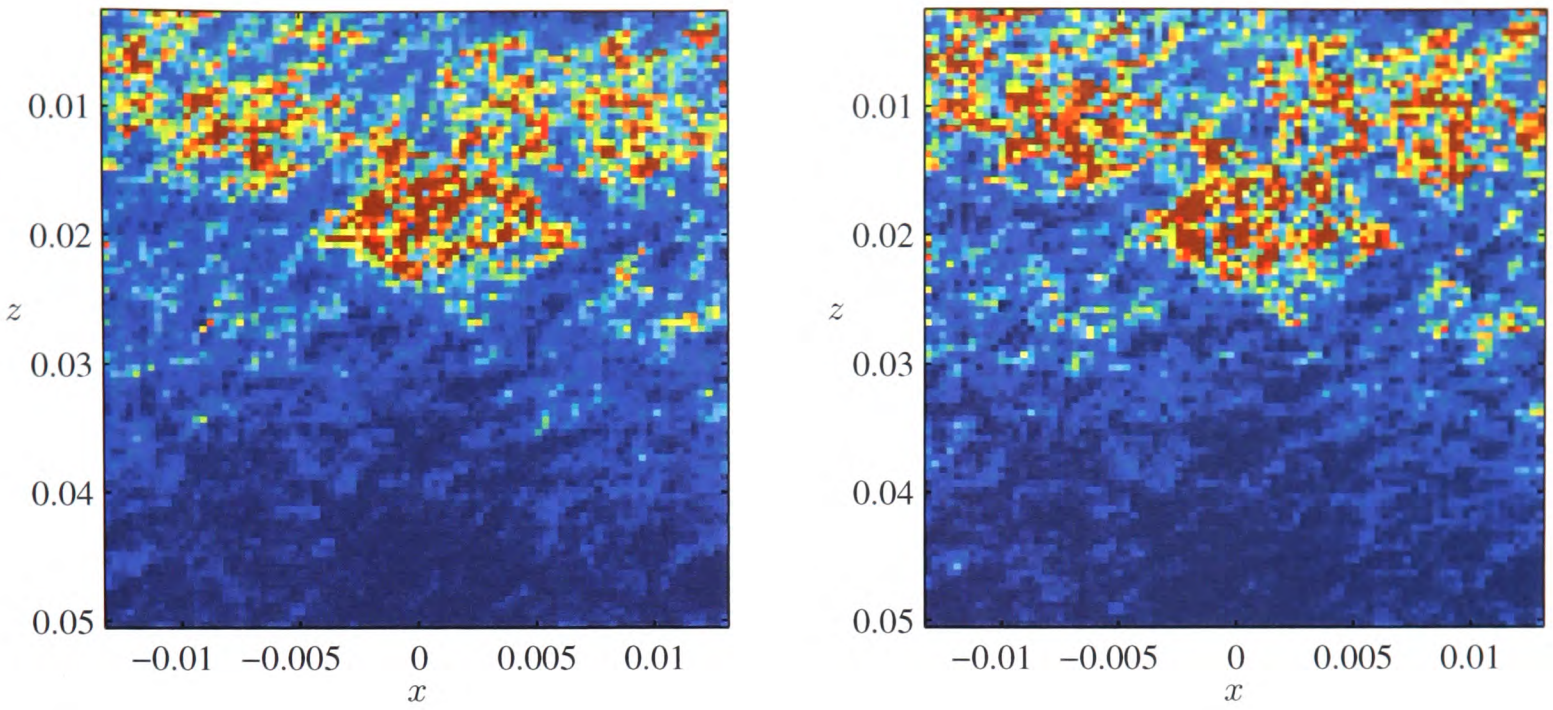


Figure 7.4: Results of applying the parameter recovery method to a displacement field estimated from successive ultrasound images of the gelatin phantom with cylindrical inclusion. No boundary conditions were applied and only one displacement field used, with an integral constraint imposed in order to ensure that the solution is non-zero. The image on the right shows the estimated values of λ and that on the left those of μ (both in arbitrary units). Red represents areas where the Lamé constants take high values and blue areas where they are close to zero. The images were taken with displacements of 20 mm and 25 mm. The ultrasound probe is located at the top of the image ($z = 0$) and the labels on the axes represent lengths in metres.

where Ω is the region on which the recovery method is applied. This constraint is chosen to ensure that the average value of λ over the domain is 1 (i.e. so that the solution $\lambda \equiv 0$ does not satisfy it). In practice this constraint is equivalent to specifying the net force applied to the sample (note that this could be measured by fitting a force probe to the ultrasound probe). The results are represented as 2-d images where variation in colour represents the variation of the values of the Lamé constants rather than as a 3-d surface as in the previous chapter. Relatively stiff regions (high λ or μ) appear red, while less stiff regions are shown in blue. This reflects the type of image that is likely to be familiar to a clinician. It also reduces the visibility of noise in the recovered Lamé constants (the solution surfaces are not at all smooth). Images of this type are used throughout the rest of this chapter.

Note that the scale is not given in the figure (since it is completely determined by the choice of the scalar constant on the right-hand-side of equation (7.1)), but the values of λ are typically of the order of 50 times those of μ (this ratio is not dependent on the integral constraint). Observe that this agrees with the ratio of Lamé constants that was used in

Section 6.1.1 to simulate realistic deformations of a 3-d region of tissue.

Recall that the centre of the inclusion is located approximately at $(x, z) = (0, 0.02)$ and it has radius 0.01 (where all of these lengths are given in metres). The inclusion can be seen clearly in both images as a red circle in the expected location and with the correct dimensions. However, the locality of the inclusion is not the only region which appears stiffer than the surrounding tissue — most noticeably the tissue close to the ultrasound probe appears to be much stiffer than that further away from the probe. The cause of these regions that appear stiffer than they actually are is not clear. One possibility is that the artifact is caused by errors in the displacement estimation method. However, this is unlikely — the displacement estimation method should be at its most accurate close to the ultrasound probe (since the true displacement on the probe is known exactly). It is more likely that the artifact is due to the 2-d approximation used in the parameter recovery method. Recall that when the parameter recovery method was tested on a cross-section of a simulated 3-d deformation in Section 6.1.2 a false peak in the solution was observed close to the boundary on which the displacement was applied. This is a similar type of artifact to that seen here and provides further support for the argument that a 3-d parameter recovery method should be used.

7.2.2.2 Boundary Conditions

An alternative to applying the integral constraint is to specify boundary conditions around the edge of the region of interest Ω. Theoretically this is a more valid method as it gives rise to a unique solution to the elasticity inverse problem. However, for real tissue it is not clear what boundary values should be used for the Lamé constants. Figure 7.5 shows the results of applying the parameter recovery method using boundary values $\lambda = 50$, $\mu = 1$ everywhere on $\partial\Omega$. These boundary values were used since they reflect the typical ratio of λ to μ seen in the results of the previous section when the integral constraint method was used. It can be seen that the method is not successful in locating the inclusion. Away from the boundary of the domain the solutions for both λ and μ tend to zero. Recall that $\lambda(x, z) = \mu(x, z) = 0$ for all (x, z) *is* a solution of the inverse problem. However, it does not satisfy the boundary conditions (and is not physically realistic).

It appears that this behaviour is due to errors in the displacement estimation technique. Recall that the parameter recovery method works by minimizing the integral of a residual

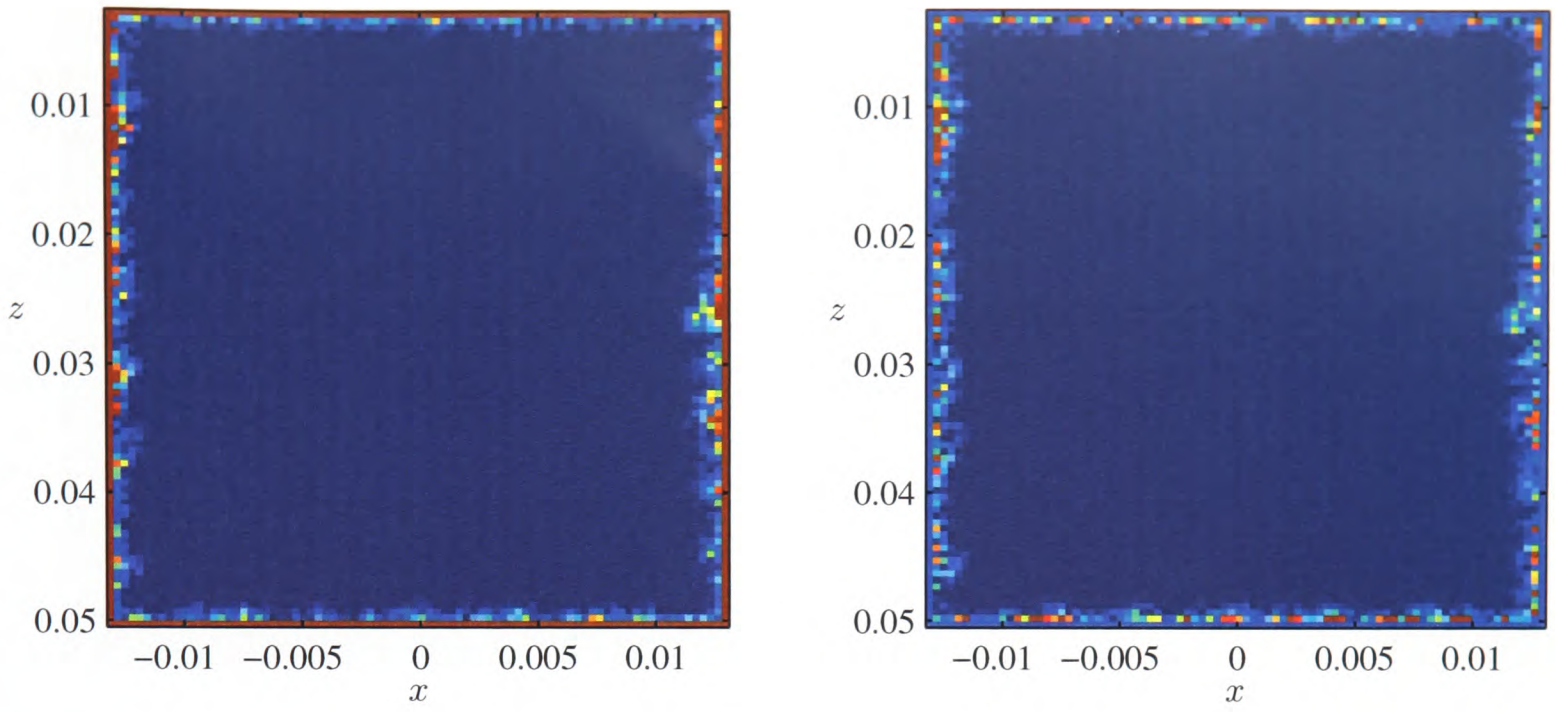


Figure 7.5: Results of applying the parameter recovery method to a displacement field estimated from successive ultrasound images of the gelatin phantom with cylindrical inclusion. Boundary conditions $\lambda = 50$, $\mu = 1$ were applied. The image on the right shows the estimated values of λ and that on the left those of μ (both in arbitrary units). Red represents areas where the Lamé constants take high values and blue areas where they are close to zero. The images were taken with displacements of 20 mm and 25 mm. The ultrasound probe is located at the top of the image ($z = 0$) and the labels on the axes represent lengths in metres.

over the region of interest. In regions where there are large variations in strain over small length scales it is difficult to approximate the Lamé constants with continuous finite element basis functions. So, in these regions the residuals are likely to be comparatively large for non-zero λ and μ. On the other hand, since $\lambda(x, z) = \mu(x, z) = 0$ for all $(x, z) \in \Omega$ is a solution of the inverse problem the residuals vanish in regions where the Lamé constants are zero. Hence, the integral of the residuals over the region is minimized by setting the Lamé constants to be zero everywhere except for a boundary layer where the residuals are comparatively large.

Similar behaviour is seen when the parameter recovery method is used with applied boundary conditions on the displacement estimates derived from series of ultrasound images of features in real soft tissue considered in the following sections. Since the results are of no use in detecting features they are not shown (or discussed) in these sections. All of the elasticity images shown in the rest of this chapter are produced using the parameter recovery method with an integral constraint.

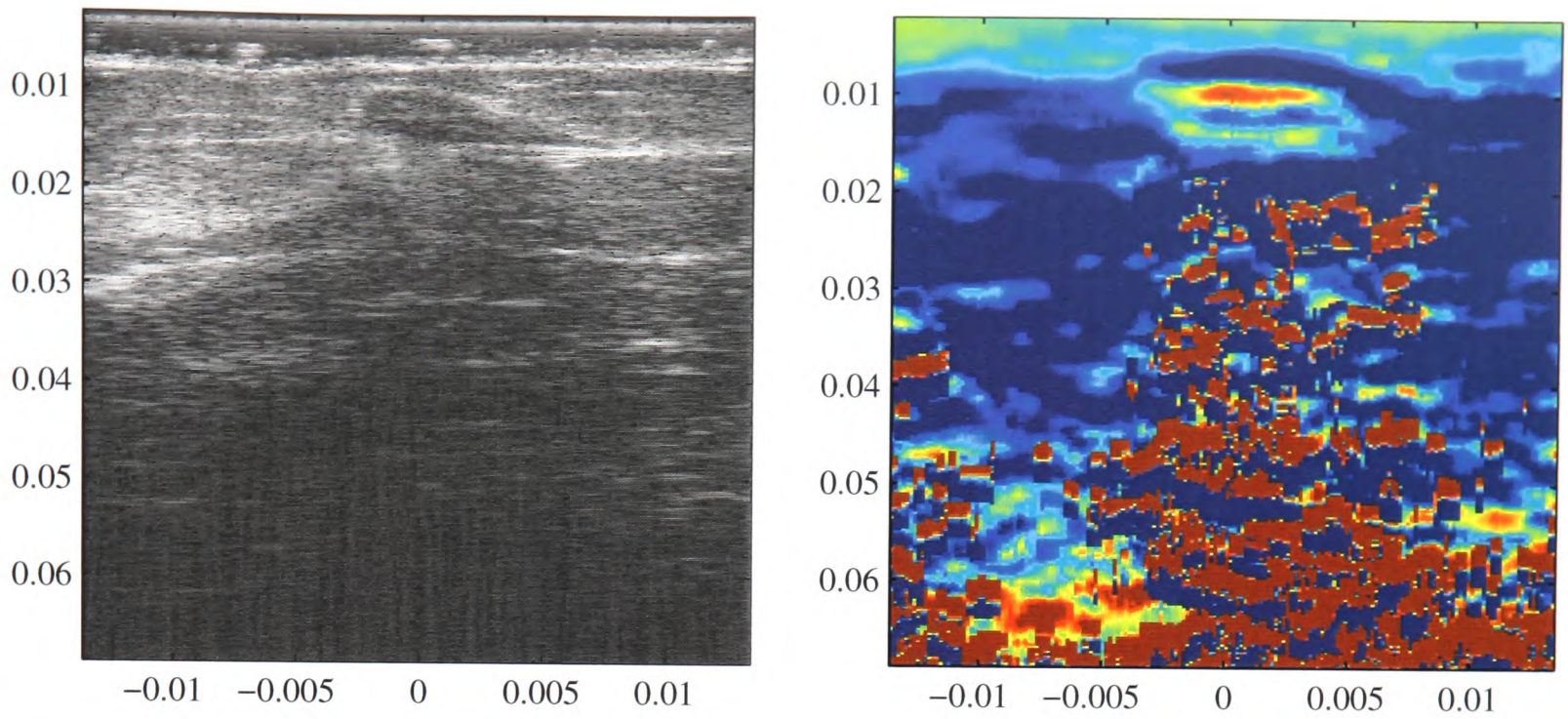


Figure 7.6: Ultrasound (left) and strain (right) images of a scar. The ultrasound probe is located at the top of the image and the numbers along the axes represent lengths in metres. In the strain image, regions where the (axial) strain is lower (i.e. stiffer regions) are shown in red, regions of higher strain in blue.

7.2.3 Scar Tissue

Figure 7.6 shows an ultrasound image of a scar surrounded by otherwise healthy tissue. Scar tissue is generally 'harder' than healthy tissue so the scar is expected to show up as a region with comparatively large Lamé constants. The ultrasound probe is located along the top edge of the image and the scar can be seen identified as the faint dark oval close to the top edge in the centre of the image. This image is taken from a series acquired by the WMVL. In this section the parameter recovery method is analyzed on this series of images. A strain image generated from the series of ultrasound images is also shown in Figure 7.6. Observe that the scar can be seen more clearly in the strain image. However, many other areas also appear stiffer than their surroundings in the strain image, especially in the 'shadow' of the scar (i.e. the region directly beneath it).

Scars are generally more regularly shaped than tumours and the interfaces between them and healthy tissue are more pronounced than those between healthy tissue and tumours. Hence, the parameter recovery method should be expected to detect the scar if it is to have any hope of detecting a tumour. The displacement field used in this section is estimated using two images, one of the undeformed tissue and one with an applied displacement of 2.5 mm.

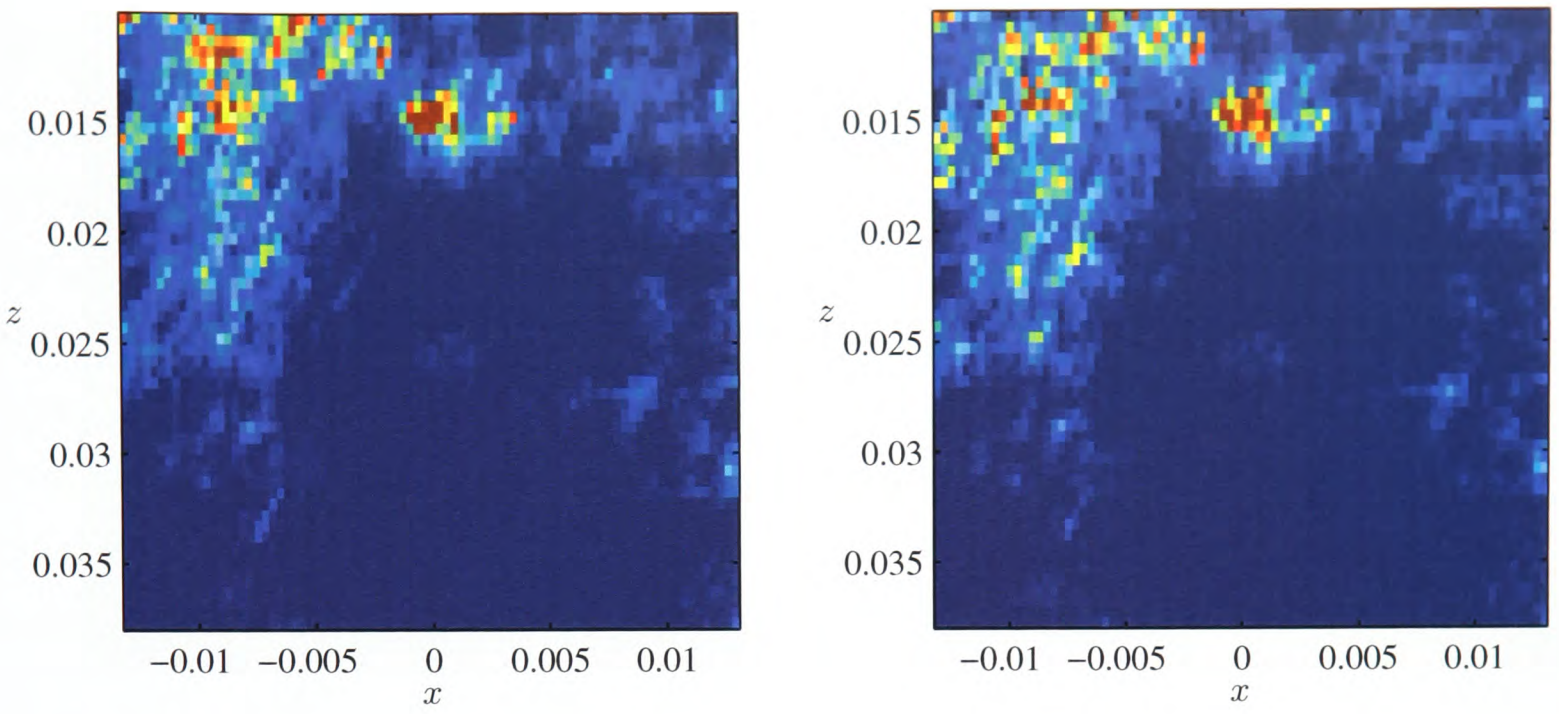


Figure 7.7: Results of applying the parameter recovery method to a displacement field estimated from successive ultrasound images of a scar. No boundary conditions were applied and only one displacement field used, with an integral constraint imposed in order to ensure that the solution is non-zero. The image on the right shows the estimated values of λ and that on the left those of μ (both in arbitrary units). Red represents areas where the Lamé constants take high values and blue areas where they are close to zero. One image was of the undeformed tissue, the other taken with a displacement of 2.5 mm. The ultrasound probe is located at the top of the image ($z = 0$) and the labels on the axes represent lengths in metres.

The results of applying the parameter recovery method to the displacement field estimated for the region containing the scar are shown in Figure 7.7. The results are similar to those for the gelatin phantom discussed in Section 7.2.2.1. The scar tissue can be clearly seen as the dark red (and green) region approximately 15 mm below the ultrasound probe and just to the right of the centre of the image. This feature is located exactly where the scar was expected to be seen. Observe also that the boundary of the scar appears much sharper in the elasticity images than in the strain image shown in Figure 7.6. However, it is not the only region where the recovered stiffness is significantly higher than that of the surrounding tissue. There is another, much larger, feature in the top left of the image. Again, the reason why this artifact is seen is not clear. However, as in the case of the gelatin phantom, its location close to the ultrasound probe suggests that it may be due to the approximation of the full 3-d inverse problem by its plane strain equivalent. Note, however, that in this case it is possible (although perhaps unlikely) that this region really is stiffer than the surrounding tissue since real tissue generally exhibits a degree of inhomogeneity.

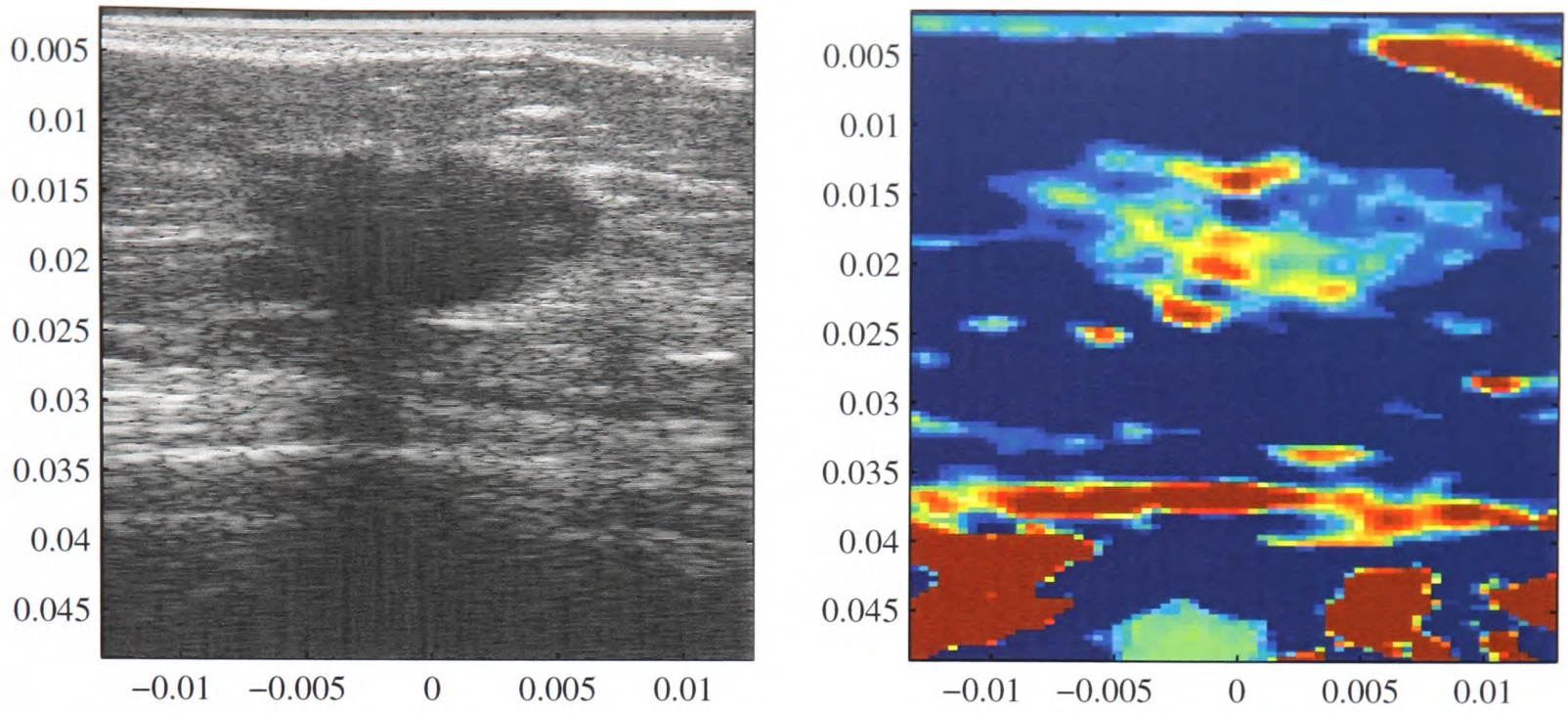


Figure 7.8: Ultrasound (left) and strain (right) images of an invasive tumour. The ultrasound probe is located at the top of the image and the numbers along the axes represent lengths in metres. In the strain image, regions where the (axial) strain is lower (i.e. stiffer regions) are shown in red, regions of higher strain in blue.

7.2.4 Invasive Tumours

Finally, the parameter recovery method is tested on two series of images of invasive tumours. Again these images have been acquired by the WMVL. One of the images from the first series of images and a corresponding strain image are shown in Figure 7.8. The tumour is very large and can be identified as the dark shadow occupying most of the width of the ultrasound image at a depth of approximately 20 mm. A stiff region can also be seen in the strain image at this location. The size of this tumour and its clarity in the raw ultrasound image suggest that it should be relatively easy to detect (although observe that its boundary is not sharply defined in the strain image). The displacement field used for the elasticity imaging is determined using an image of the undeformed tissue and a second taken when the tissue is subject to an applied displacement of 2.5 mm. Since tumours are generally made up of stiffer tissue than their surroundings it is expected that in the elasticity image the tumour will show up as a red region.

Figure 7.9 shows the results of applying the parameter recovery method to the first series of images of an invasive tumour. The tumour can be clearly seen as a dark red region in the images for both λ and μ, located at a depth of approximately 20 mm and occupying most of the width of the image exactly as expected. Its boundary is much sharper than was

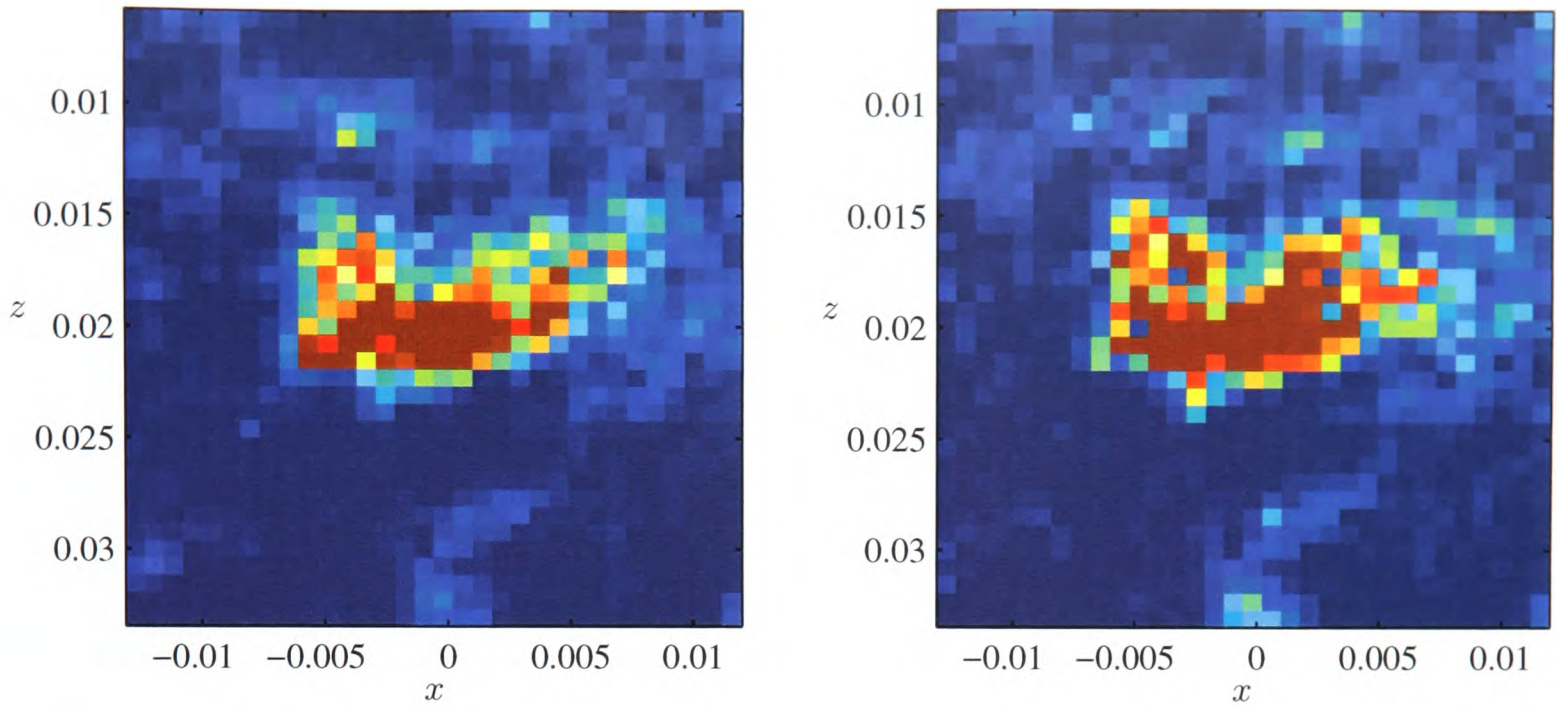


Figure 7.9: Results of applying the parameter recovery method to a displacement field estimated from successive ultrasound images of a scar. No boundary conditions were applied and only one displacement field used, with an integral constraint imposed in order to ensure that the solution is non-zero. The image on the right shows the estimated values of λ and that on the left those of μ (both in arbitrary units). Red represents areas where the Lamé constants take high values and blue areas where they are close to zero. One image was of the undeformed tissue, the other taken with a displacement of 2.5 mm. The ultrasound probe is located at the top of the image ($z = 0$) and the labels on the axes represent lengths in metres.

the case in the strain image shown in Figure 7.8. Note that the shape appears much more irregular in the elasticity images than it did in the ultrasound image. This is in agreement with the true nature of tumours: they are almost always irregularly shaped (especially when they are large) as new cells tend to grow in the direction of blood vessels in order to gain access to nutrients.

An ultrasound image from the second series and a corresponding strain image are shown in Figure 7.10. The tumour can be seen as a faint shadow on the right-hand-side of the ultrasound image, about half way up. However, it is not visible in the strain image. As in the previous images, the ultrasound probe is located along the top edge of the image. The displacement field used in this section is again estimated using two images, one of the undeformed tissue and one with an applied displacement of 2.5 mm. This image is much more typical of those seen in cancer screening than the one shown in Figure 7.8. It is anticipated that the smaller size and less well-defined (in the image) nature of the tumour will make it harder to detect using elasticity imaging.

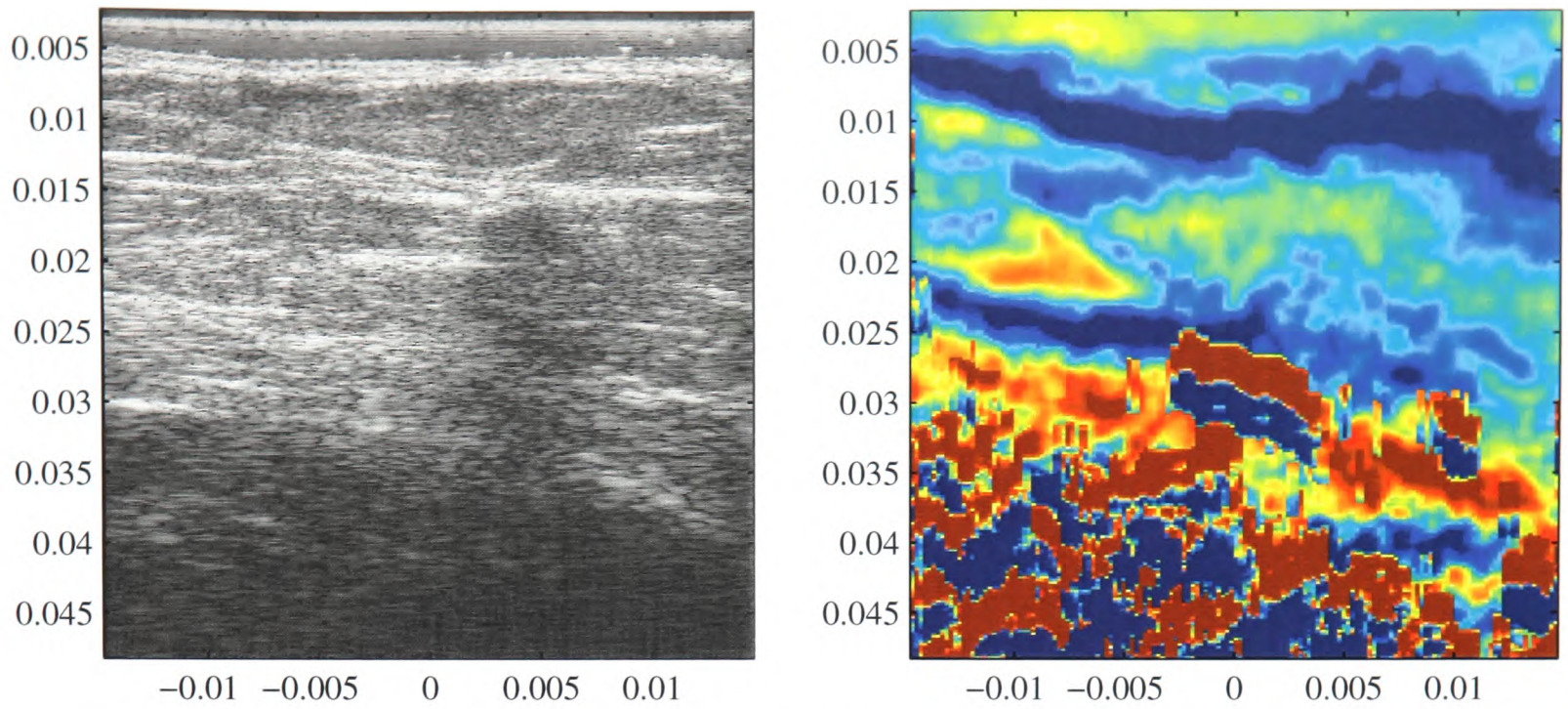


Figure 7.10: Ultrasound (left) and strain (right) images of an invasive tumour. The ultrasound probe is located at the top of the image and the numbers along the axes represent lengths in metres In the strain image, regions where the (axial) strain is lower (i.e. stiffer regions) are shown in red, regions of higher strain in blue.

Figure 7.11 shows the results of applying the parameter recovery method to the displacement estimate corresponding to the second series of images of an invasive tumour. A feature can clearly be seen centred at $(x, z) \approx (-0.005, 0.013)$ in the images for both λ and μ. Unfortunately this does not correspond to the location of the tumour, which can be identified as the bright area centred on $(x, z) \approx (0.004, 0.015)$, immediately above the large dark shadow in the right of the image shown in Figure 7.10. This area does appear to be slightly stiffer than the surrounding tissue in the elasticity images, but not to a sufficient degree to be reasonably identified as a feature of interest — especially as another region that is apparently much stiffer is visible in the image. The results illustrated in Figure 7.11 indicate that the parameter recovery method still requires significant development before it could be utilized in a clinical environment.

7.3 Conclusions

In this chapter the parameter recovery method developed in Chapter 4 has been tested on displacement fields estimated from series of real ultrasound images. In Section 7.2.2 the parameter recovery was tested on a series of images of a gelatin phantom containing

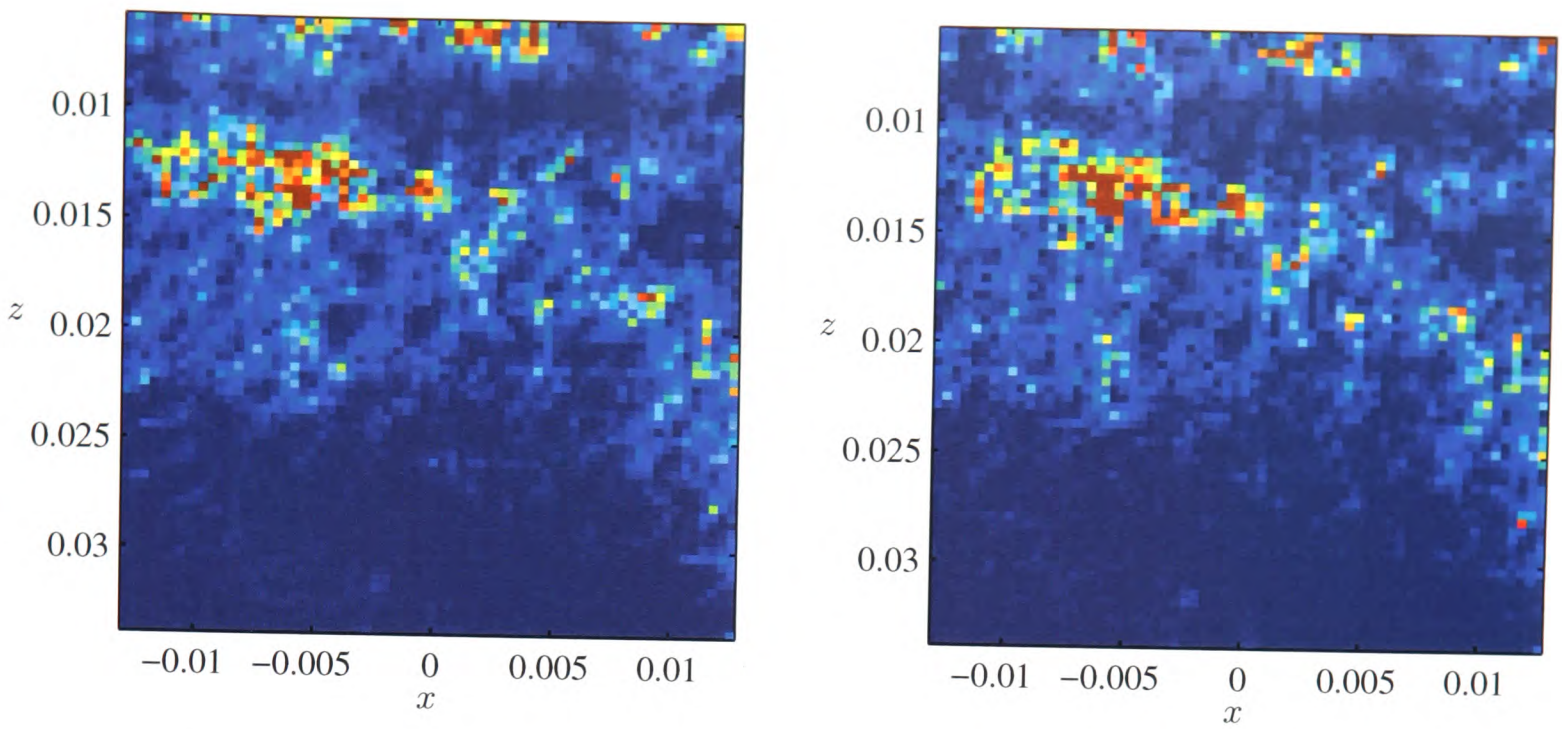


Figure 7.11: Results of applying the parameter recovery method to a displacement field estimated from successive ultrasound images of an invasive tumour. No boundary conditions were applied and only one displacement field used, with an integral constraint imposed in order to ensure that the solution is non-zero. The image on the right shows the estimated values of λ and that on the left those of μ (both in arbitrary units). Red represents areas where the Lamé constants take high values and blue areas where they are close to zero. One image was of the undeformed tissue, the other taken with a displacement of 2.5 mm. The ultrasound probe is located at the top of the image ($z = 0$) and the labels on the axes represent lengths in metres.

a stiff cylindrical inclusion. The method was used both with an integral constraint and with applied boundary conditions. Both of these methods were intended to ensure that the solution $\lambda \equiv \mu \equiv 0$ was not a valid solution of the inverse problem. The theory of elliptic eigenvalue problems guarantees a unique solution when boundary conditions are applied, but not when the integral constraint is used with only one displacement field.

In Section 7.2.2.1 it was shown that the elasticity images resulting from using the integral constraint (even with only one displacement field) can be used to locate the inclusion — although other regions close to the ultrasound probe appeared to be stiffer than was expected. Section 7.2.2.2, on the other hand, shows that even though a unique solution is guaranteed when boundary conditions are applied this solution turns out to be zero throughout the domain, except for a thin boundary layer close to its edge. This, together with the difficulty in determining appropriate boundary values, suggests that the method is best used with no boundary data at all and (if possible) multiple displacement fields in order to ensure the solution is unique.

Next, the method was applied to ultrasound images of real tissue. First, in Section 7.2.3 a displacement field was estimated from a series of images of a region of tissue containing scar tissue. As in the case of the gelatin phantom the expected feature could be identified as a stiffer region in the expected location. However, once again other areas of the region of interest — away from the scar tissue — appeared to be unexpectedly stiff in comparison to their surroundings. Then, in Section 7.2.4, series of ultrasound images of two regions of tissue containing invasive tumours were used to estimate displacement fields. For one of these series of images the method was successful in locating the tumour. For the other series, however, it did not clearly identify the tumour — instead appearing to identify a different region as being much stiffer than the rest of the image (recall, however, that the tumour was not visible in the strain image either). This suggests that further work is needed to improve the method. Finally, note that in the cases where the method did detect the feature present in the ultrasound image the boundary appeared much sharper in the elasticity images than in the strain images.

Overall, the results were encouraging, and suggest that the method, although still in need of significant improvement, has the potential to be applied successfully in a clinical environment. It was generally successful in identifying the expected features within the tissue, although other regions also appeared stiffer than would otherwise be expected. It was suggested that these artifacts are due to modelling the 3-d inverse problem in 2-d. The development of a 3-d parameter recovery method to be used in conjunction with 3-d images would represent a significant improvement on the current method.

Chapter 8

Summary and Future Work

In this thesis a parameter recovery method based on a direct solution of the linear elasticity inverse problem has been developed and validated with the aim of assisting in the analysis of ultrasound images of regions of soft tissue thought to contain tumours. This work is summarized in Section 8.1. Then, in the remainder of this chapter, some suggestions for the future direction of this work are described. This chapter is intended to provide a framework for the future development of the method towards a clinical application.

There are still some important problems with the method that need to be addressed before it could begin to be used in a clinical setting. In particular it does not deal well with sharp changes in the parameter values, it uses an invalid 2-d approximation, and it is very sensitive to errors in the displacement estimate. Possible solutions to these problems are discussed in Section 8.2. Once these problems have been solved there may be other improvements to the method that are desirable. These may be to increase its accuracy and reliability or to make the results easier to understand or analyze. Some examples of extensions of this type are proposed in Section 8.3.

8.1 Summary of Key Contributions

The major original contribution of this thesis is to develop a new method for the direct numerical solution of the plane strain linear elasticity inverse problem, primarily for use in ultrasound elasticity imaging. The solution method works by reformulating the inverse problem (which is hyperbolic) as a minimization problem (which can be regularized so that

it has elliptic Euler-Lagrange equations) which can be solved using finite element methods. A strategy for overcoming the lack of boundary data for the Lamé constants for a typical series of ultrasound images by using multiple displacement fields has also been developed. The theory behind these methods was explained in Chapter 4 and they were validated on displacement fields calculated by solving the plane strain linear elasticity forward problem in Chapter 5.

A further original contribution is the demonstration (in Chapter 6) that plane strain is not a valid assumption for the ultrasound imaging process. This was demonstrated both by comparison of the solution of the appropriate 2-d and 3-d forward problems and by attempting (unsuccessfully) to reconstruct the elastic properties of a 2-d cross-section from a numerical solution of the 3-d linear elasticity forward problem. In response to this the inverse problem solution method was extended to 3-d by the formulation of an appropriate minimization problem.

Finally, in Chapter 7 the (plane strain) solution method was used to generate elasticity images corresponding to displacement fields calculated from real ultrasound images. Although 3-d ultrasound images (and the 3-d inverse problem solution method) would have been required in order to quantify the differences in tissue stiffness at different locations the method was still able to detect stiffer regions of tissue qualitatively in some images (although it failed to do so other images).

8.2 Addressing Problems With the Current Method

Although the results of applying the parameter recovery method to both idealized displacement fields in Chapter 5 and those arising from real ultrasound images in Chapter 7 were encouraging they also highlight a number of problems with the method as it is currently formulated. In this section possible solutions to these problems are considered. There were three major problems with the method and each of these are addressed in turn.

8.2.1 Dealing with Discontinuous Lamé Constants

In Sections 5.1.2 and 5.2.3 it was shown that the parameter recovery method was unable to detect discontinuities in the values of the Lamé constants. Since this type of variation in

λ and μ is likely to be seen at interfaces between different tissue types it is important that the method should be able to cope with it. In Section 8.2.1.1 the use of discontinuous finite element basis functions is proposed as a way of dealing with this problem. An alternative proposal is considered in Section 8.2.1.2: smoothing the displacement fields before applying the parameter recovery method.

8.2.1.1 Discontinuous Finite Element Methods

Recall from Section 2.4 that conventional finite element methods assume that each node that is on the boundary between two elements is a member of both elements and that the value of the numerical solution at the node is the same in both elements. This means that the entire numerical solution is *continuous*[1] (since the conventional basis functions are continuous within each element and the interpolation method along the intersection of two elements is the same in each element). Hence, the method cannot deal with variation in the values of the Lamé constants on a length scale smaller than the element size.

Suppose instead that the value of the numerical solution at each node can take different values in each element that the node is a member of, perhaps with some penalty function applied to ensure that continuous solutions are preferred to discontinuous ones wherever possible. Then, where appropriate, the numerical solution can exhibit sharp changes in value across element boundaries. This allows the method to deal much more efficiently with regions where the solution does have discontinuities. Techniques of this type have been developed over recent years (see, e.g., Eriksson *et al.* [24] for an introduction to disconstinuous finite element methods and Ten Eyck and Lew [82] for an example of their application in elasticity theory) and it is suggested here that adapting the parameter recovery method to use them would significantly increase its ability to deal with the sharp changes in the values of λ and μ typically seen at tissue boundaries. However, no preliminary work has yet been carried out to investigate the feasibility of such an approach as it would involve a complete rewrite of the code base.

[1]Observe that an exception to this is the piecewise constant basis functions. For these basis functions the nodes are located at the centre of the elements rather than on their boundaries, so the solution can be discontinuous even though the values at the nodes are single-valued. However, piecewise constant basis functions are not of a high enough order to be used in the parameter recovery method.

8.2.1.2 Smoothing the Displacement Fields

Although the parameter recovery method fails when the Lamé constants it is attempting to reconstruct are discontinuous (in the sense that the change in values is over a length scale smaller than the element size), recall from Section 5.2 that it works well when the true parameter distributions are continuous but contain rapid variations in λ and μ (but over a length scale that is larger than the element size — see Figures 5.11 and 5.13). Discontinuities in the Lamé constants lead to discontinuities in the first derivatives of the strain components, whereas continuous distributions of λ and μ give rise to strains with continuous first derivatives. This suggests that an alternative way of dealing with discontinuous Lamé constants could be to pre-smooth the displacement fields before applying the parameter recovery method. It is hoped that this will give rise to a continuous approximation to the true, discontinuous distribution of the stiffness parameters. This idea has one key drawback, however: it may remove some of the more subtle information contained in the displacement fields — in particular, the size of the jump in the values of the strain derivatives provide information about the size of the jump in the Lamé constants. For this reason the use of discontinuous finite element basis functions appears more likely to result in a successful parameter recovery method.

8.2.2 3-Dimensional Implementation

Recall that the parameter recovery method developed in Chapter 4 assumed that the material being imaged was undergoing plane strain (this assumption was discussed in Section 4.1.2). It was later demonstrated, in Section 6.1.1, that the assumption is not valid for the types of deformation usually applied to a region of tissue during ultrasound imaging — the cross-sections resulting from a 3-d deformation are not at all similar to the deformations seen in a material undergoing plane strain. Further, it was shown in Section 6.1.2 that whilst applying the 2-d parameter method to a cross-section of a 3-d deformation may give results that appear plausible (i.e. a feature in the correct region), all quantitative information about relative stiffness is lost. This makes it impossible to use the method to distinguish between different types of feature (e.g. to decide whether a stiff inclusion is a cyst or a tumour).

The conversion of the 2-d parameter recovery method to 3-d was addressed in some detail in Section 6.2. The 3-d inverse problem was shown to be hyperbolic and reformulated as

a minimization problem. The use of multiple displacement fields to give a unique solution without needing to specify boundary data was investigated and it was shown that three displacement fields are required in order to guarantee that the solution is unique up to a multiplicative constant. In fact, the theory was developed to the same level as that of the 2-d inverse problem and all that remains to be done is to write and test the code. This was not pursued since 3-d ultrasound images are not currently available to validate the method. An estimate of the time that the 3-d solution method would take to produce a stiffness field was also addressed — it was thought to be of the order of an hour, which is much quicker than most of the 2-d iterative methods. So, the key requirement to develop an effective 3-d parameter recovery method is the availability of series of 3-d ultrasound images. These are likely to be available in the very near future.

8.2.3 Sensitivity to Displacement Errors

Recall from Section 5.4.1 that the parameter recovery method did not appear to be very robust with respect to errors in displacement estimation, here modelled primitively by adding normally distributed errors to an ideal displacement field. This is likely to be a serious problem in a clinical setting since the noise inherent in ultrasound images makes it extremely difficult to calculate a perfect displacement field.[2] One way of reducing the errors would be to smooth the displacement field. Recall that this was also suggested in Section 8.2.1.2 as a method of dealing with discontinuities in the Lamé constants. Then it was suggested that the loss of subtle information about the distribution of λ and μ made the method less attractive than the alternative of using discontinuous finite element basis functions. However, if there are sizeable errors in the displacement estimate then this information may already have been lost and so pre-smoothing may help to improve the results of the parameter recovery method.

Alternatively, the value of δ used in equation (4.123) could be chosen to be non-zero. Recall that this introduces terms corresponding to $\delta(\epsilon_{11}+\epsilon_{22})\nabla^2\lambda$ and $\delta(\epsilon_{11}+\epsilon_{22})\nabla^2\mu$ to the Euler-Lagrange equations (4.93) and (4.95) respectively. Hence, as the value of δ

[2]Note, however, that when the method was applied to displacement fields estimated from real ultrasound images (which would certainly have contained errors) in Chapter 7 the results were much more encouraging than those seen in Section 5.4.1. This indicates that the method may be more robust with respect to the type of errors typically found in ultrasound displacement estimation (which may be expected to be, for example, continuous rather than totally random) than to those considered in Section 5.4.1. Nevertheless, the method would still benefit from being as insensitive as possible to errors in displacement estimation.

increases non-smooth solutions for λ and μ are penalized increasingly. For very small δ the solution will not differ significantly from the $\delta = 0$ case since the inverse problem is well-posed and so has continuous dependence on its parameters. As $\delta \rightarrow \infty$, on the other hand, the Euler-Lagrange equations tend towards the diffusion equation. So, λ and μ will tend towards constant functions. It may be possible to determine an intermediate value of δ where the solution is not approximately constant, but the penalty for non-smooth λ and μ somehow 'balances' with the requirement that the original Euler-Lagrange equations and the method gives a smooth approximation of the actual, discontinuous, Lamé constants.

Preliminary results on the solutions illustrated in Chapter 4 show the asymptotic behaviour but it has not proved possible to generate any type of continuous approximation to the true solution. It is not clear, however, whether this is due to a fundamental flaw in the proposal or to the effects of the stress singularities at the corners of the domain diffusing and having a global effect on the solution. Further investigation (on simulated displacement fields *without* the stress singularities) is required in order to properly evaluate the proposal.

8.3 Possible Extensions to the Method

In Section 8.2 some ideas for solving the major problems with the current parameter recovery method have been proposed. It is anticipated that the solution of these problems will result in a method that is able to provide reliable, quantitative tissue stiffness data for a 3-d region of soft tissue much more quickly than the existing iterative methods can. Nevertheless, further improvements may be desirable in order to improve the robustness and accuracy of the method and to help the user to interpret the elasticity image produced. Extensions of this type are discussed in this section. These proposals do not impact directly on the fundamental structure of the parameter recovery method, but instead either make small alterations to the detail of the method, provide improved input data or make the interpretation of the output simpler.

First, in Section 8.3.1 the advantages that could result from an improved displacement estimation technique are considered. In Section 8.3.2 the introduction of constraints on the allowable values of the parameter values is discussed. Section 8.3.3 describes the possibility of using more realistic models of soft tissue mechanics than linear elasticity. Finally, in Section 8.3.4 the incorporation of a confidence measure in the solution at each point in the domain is considered. This would allow regions where the stiffness data was

considered unreliable to be easily identified and allow the user to concentrate on regions were the results are known with a greater degree of certainty.

8.3.1 Improved Displacement Estimation Techniques

The sensitivity of the parameter recovery method to errors in the displacement estimates was discussed in Section 8.2.3, and possible methods for reducing this sensitivity were suggested. However, the need to do this could be reduced if the techniques used for displacement estimation could be made more accurate. The development of displacement estimation techniques has been an integral part of research into elasticity imaging techniques since the inception of the field in the late 1980's. This part of the process of elasticity imaging has so far been ignored in this work — the displacement estimates corresponding to the real ultrasound images in Chapter 7 were automatically generated as part of the imaging process. In this section a brief overview of displacement estimation is given, but no methods for improving the existing methods are proposed.

Displacement estimation techniques used have been separated into three main classes by Barron *et al.* [7]. *Differential* (or gradient-based) methods make the assumption that brightness does not change with time. The displacement estimation is calculated by calculating the derivatives of the intensity and often using a knowledge-based regularization procedure. This class of methods do not produce good results on ultrasound images due to their low signal-to-noise ratio. The second class of methods are based on *frequency*, exploiting either energy or phase data. The theory suggests that phase-based displacement estimation should work best for ultrasound. However, the methods exploit filters that are tuned to a velocity range and it is difficult to design optimal filters. The final class is *block-matching* displacement estimation. Here the displacement is estimated by maximizing a *similarity measure* between the images of interest. These methods are generally computationally expensive and are best suited to tracking the position of one feature (e.g. a lesion) between the images.

Many of the advanced displacement estimation techniques used today are based on the maximum likelihood block-matching algorithm of Singh [78]. These methods combine motion estimates based on image similarity and smoothness of motion in a probabilistic framework in order to determine an optimal estimate with the following (highly desirable) properties:

- Regions of one image are displaced to regions of the second image that 'look similar' to them.
- Displacements do not give rise to 'holes' (pixels that are close together in the undeformed image being displaced to totally different parts of the deformed image) or 'folds' (a pixel that is, say, to the left of another pixel in the undeformed image being moved so that it is to the right of the second pixel in the deformed image) in the tissue.

Most image similarity measures are based on a statistical model of the speckle pattern of ultrasound images — since the same particles scatter the ultrasound wave in both images they should give rise to speckle patterns in each image that are somehow statistically similar. As understanding of the way in which soft tissue scatters ultrasound waves improves these models, and hence the resulting similarity measures and displacement estimates, should become more accurate.

8.3.2 Realistic Parameter Values

One of the major advantages of the iterative elasticity imaging methods discussed in Section 3.1 is that it is easy to ensure that they return values of the Lamé constants that fall within the physically allowable range.[3] It is clear before each simulation of the forward problem whether or not the values of the Lamé constants being used are allowable — and if they are not then they can easily be altered to ensure that they are. For the parameter recovery method developed in this thesis, however, the values of the Lamé constants cannot be checked until the method has finished. Also, if the values returned are not valid then as the method stands there is not a way to deal with this.[4] Although for a perfect displacement field the method will return the correct solution (assuming it has no discontinuities) it is unrealistic to expect that any displacement estimation technique will not contain some

[3]Thermodynamic considerations mean that for a real material the Young's modulus and shear modulus must both be positive (so that strain energy cannot be stored within the material without work being done on it). In terms of the Lamé constants these constraints become $\mu > 0$, $3\lambda + 2\mu > 0$. So, λ can take a negative value (although it must be greater than $-\frac{2}{3}\mu$) but μ cannot.

[4]Observe that examples of the method returning negative values of the Lamé constants can be seen in various places in this thesis. With the ideal displacement fields used in Chapter 5 they appear in the cases where the method fails to detect the correct solutions (due to its discontinuities or the addition of noise to the displacement fields). In Chapter 6 regions in which $\mu < 0$ are present as a result of the 2-d parameter recovery method being used on the cross-section of a 3-d displacement field.

errors, however small. It is, therefore, desirable that the parameter recovery method should be altered to include some way of ensuring that the solution is physically valid.

One way to constrain the solution given by the parameter recovery method to lie in the physically acceptable range would be to add an appropriate penalty function to the function to be minimized, i.e. (in 3-d) replace the minimization problem given in equation (6.64) by the following:

$$\text{Find } \lambda(x,y,z),\ \mu(x,y,z) \text{ to minimize } \int_\Omega \left(R_1^2 + R_2^2 + R_3^2 + P(\lambda,\mu)\right) dxdydz, \qquad (8.1)$$

where R_1, R_2 and R_3 are the residuals associated with each of the governing equations and $P(\lambda,\mu)$ has the following properties:

- $P(\lambda,\mu) \gg 1$ if $\mu < 0$ or $3\lambda + 2\mu < 0$,
- $P(\lambda,\mu) \ll 1$ if $\mu > 0$ and $3\lambda + 2\mu > 0$.

These requirements may be relaxed close to $\mu = 0$ and $3\lambda + 2\mu = 0$ in order to allow a continuous function to be used for P. Note, however, that $P(\lambda,\mu)$ cannot be represented by a quadratic form in λ and μ (and their first derivatives). So, its Gâteaux derivatives will be nonlinear in λ and μ (and their first derivatives). Hence, the system of algebraic equations resulting from the finite element method will also no longer be linear. So, in order to incorporate a constraint on the values of λ and μ a nonlinear solution method would have to be incorporated into the parameter recovery method. This should be fairly straightforward to do (e.g., using a simple Newton's method solver) and assuming a sensible penalty function is chosen should not result in a significant increase in the time required to obtain a solution.

8.3.3 Improved Modelling of Soft Tissue Mechanics

In developing the parameter recovery method in this thesis it has been assumed that soft tissue is an isotropic, homogeneous elastic solid. All of these assumptions are invalid: soft tissue *in vivo* does *not* behave as an isotropic material (see, e.g. Fung [30]) due to the complex structure of the tissue, which is usually fibre-reinforced; it is almost certainly inhomogeneous at least on a small scale (since ultrasound images of it contain scatter); it is fundamentally inelastic (experimental stress-strain curves show hysteresis, creep and stress relaxation — again, see Fung [30]); and it is made up mostly of fluid surrounding a

solid matrix. Some possible candidate models that could be developed in order to deal with these problems are discussed in Appendices A (elastic models) and B (non-elastic models). The technique of formulating the inverse problem as a minimization problem and solving this using finite element methods should, in theory, work for any elastic or mixture model, since for an n-parameter model a system of exactly n PDEs can be generated by taking the Gâteaux derivatives of the minimization problem with respect to each parameter in turn. Implementation may be more difficult for a viscoelastic model since these require knowledge of the history of the material being imaged which may not be available.

Of the types of elastic models available, only *Fung elasticity* (see Appendix A.3.2) has been developed with the deformation of soft tissue *in vivo* in mind. Hence, this appears to be the most promising type of elastic model to develop for future use in elasticity imaging. However, although Fung elasticity deals with the problem of nonlinearities in the deformation of soft tissue it does not address the issues of inhomogeneities, anisotropy and viscoelasticity. Each of these problems could be addressed individually and a complex model of fluid flow around an anisotropic inhomogeneous viscoelastic solid could be constructed, however this is likely to be incredibly complex requiring the determination of many parameter values and a large amount of computing time in order to reach a solution.

Instead, the most appropriate way of describing the dynamics of real soft tissue may be to use a mixture model. In such a model, each element of the finite element mesh would contain a fixed amount of (elastic) solid matrix and fluid would be free to flow between the elements subject to various constraints (e.g. the fluid fraction of each element would need to be related to its volume at any given time). Anisotropy could be included by modelling the mesh as an anisotropic solid (the fluid would be expected to be isotropic as it is mostly water). Inhomogeneities could be dealt with on an element-by-element basis at the beginning of the modelling process, possibly by a homogenization process within the element, as they would necessarily be fixed within the matrix. Initially such a model could be developed as an extension of the KLM model (see Appendix B.2.3.1). However, this was designed with 1-d cartilage in mind and careful thought will need to be given to how to extend it to a 3-d region with necessarily more complex geometry (e.g. the inclusion of different tissue types, fibre-reinforcing, etc.). Whilst the development of such a model can be considered as a possible ultimate aim of a soft tissue modelling project it is very ambitious and is likely to still be years away.

8.3.4 Incorporate Confidence in Stiffness Estimates

As part of the displacement estimation process described in Section 8.3.1 it is often possible to provide a level of confidence in the accuracy of the estimate at each point. Since more accurate displacement estimates give rise to more accurate elasticity images this raises the question of whether it is possible to use the confidence measure for the displacement estimate to construct a similar measure of the accuracy of the stiffness estimates throughout the region of interest. If available this information could be incorporated in the visualization process to make the image easier to understand. For example, the elasticity image could be overlaid on an ultrasound image of the corresponding region and its transparency adjusted depending on the degree of confidence of the solution at each point (regions where there is a high level of confidence in the solution would appear more opaque). The idea would be that any artifacts caused by errors in the displacement estimation technique would be highly transparent and so not visible in the final elasticity image — making the regions that really are stiffer than the surrounding tissue stand out more clearly.

It is not clear whether or not this type of confidence measure could be developed for the elasticity image. What is certain is that the errors in displacement estimation are not proportional to the resulting errors in the solution of the inverse problem. This can be illustrated by considering the results of solving the inverse problem using periodic boundary conditions shown in Figure 5.34. In this figure it can be seen that the errors in the stiffness imposed on the boundary give rise to errors throughout the region of interest (the solutions show oscillations everywhere). So, suppose that the displacement field is known exactly everywhere except for one small region in which there are small errors in the displacement estimate. In this region the reconstructed tissue stiffness will be close to the true distribution, but not exactly correct. These errors in the stiffness estimate will then propagate through the domain as seen in Figure 5.34. Hence, there are non-zero errors in the reconstructed parameter values away from the small region in which the displacement is not known exactly, i.e. the confidence in the stiffness estimate is not proportional to that in the displacement. Nevertheless, there is an extensive theory of error estimation for finite element methods. It is possible that this could be used to generate a confidence measure for the parameter recovery method.

Another possible method to generate a measure of confidence in the solution of the inverse problem would be to perturb the initial displacement estimation within the error bounds given by the displacement estimation technique. The inverse problem could then be solved

several times (using different perturbed displacement fields) and the results compared. Regions in which the solutions of the inverse problem for all (or most) of the perturbed displacement fields are similar can be given a high level of confidence while regions where the solutions of the inverse problems differ substantially could be assigned a lower confidence level.

8.4 Concluding Remarks

Ultrasound elasticity imaging provides a framework for quantifying the observation that tumours feel 'harder' than healthy tissue. In this thesis a new method for rapidly generating elasticity images from a series of ultrasound images has been developed. The method has been tested, with some success, on series of ultrasound images of real tumours. However, it is not alway successful in detecting features in the ultrasound images — a significant amount of further work is needed before the method could be reliably used in a clinical setting. It is anticipated that the future development of the method will rely on three separate strands of research: the increased availability of high quality 3-d ultrasound images, improved algorithms for registering the images to one another, and better models of the deformation of soft tissue coupled with methods for reliably solving the resulting inverse problems. Work is progressing rapidly in each of these fields and, hence, methods that improve on the results presented in this work can be expected to be developed in the near future.

Appendix A

Elastic Models of Soft Tissue Mechanics

In this appendix various different models that are applied to soft tissue mechanics are described. The models can be split into three general classes, each of which is discussed separately. The simplest models (that are used almost exclusively in existing elasticity imaging work) are based on linear elasticity theory and are discussed in Section A.1. However, these models may not be appropriate for elasticity imaging of real tissue if the deformations fall outside of the linear range. Nonlinear solids can be modelled by finite elasticity techniques and these are discussed in Section A.2. All of the models of linear and finite elasticity are standard and have been developed to predict the behaviour of metals and rubber. In Section A.3 some newer models of elasticity developed specifically for soft tissue are considered. These models include some of the viscoelastic behaviour of real soft tissue (e.g. hysteresis).

A.1 Linearly Elastic Models

It has been known that soft tissues exhibit nonlinear and inelastic behaviour over finite strains since the middle of the 19th century (see, e.g., Wertheim [87] and Roy [73]). However, much work has been done on modelling soft tissue as a linear elastic material and this continues up to the present day. This appears to be due to the simplicity of the model in comparison with those required for finite elasticity. In this section linear elastic models

for the deformation of soft tissue are described.

A.1.1 Classical Linear Elasticity

The simplest way to model soft tissue is as an isotropic linear elastic solid. This is the approach taken by Schnabel *et al.* [77, 76] and Pathmanathan [60] to model the deformation of breast tissue. In the work of Schnabel *et al.* [77, 76] it is also assumed that the Poisson's ratio of the tissue is given by $\nu = \frac{1}{2} - \epsilon$ where $0 < \epsilon \ll 1$. This has the effect of making the tissue nearly incompressible. Models of this type have the advantage of depending on only two parameters for each tissue type — the Young's modulus E and Poisson's ratio ν (which is known if the tissue is assumed to be nearly incompressible). As has been discussed above, linear elasticity is valid only for sufficiently small deformations. Hence, it is to be expected that models of this type will only be valid when the soft tissue is displaced by a small amount. This is indeed the case — comparison with MRI images of small deformations in Pathmanathan [60] shows good agreement between theory and experiment whereas Whiteley *et al.* [89] show that the deformations seen in clinical deformation of breast tissue cannot be considered to be small enough to justify the use of linear elastic theory.

A.1.2 Azar's Pseudo-Nonlinear Model

Work by Azar *et al.* [3, 4] attempts to combine the simplicity of a linear model with the ability to deal with large deformations. This is done by modelling a large deformation as a series of much smaller displacements with the elastic properties being re-calculated at each step as a function of strain. In Azar *et al.* [3, 4] it is assumed that soft tissue is isotropic, homogeneous and incompressible. The incompressibility condition is enforced (approximately) by setting $\nu = \frac{1}{2} - \epsilon$ as in the models of Schnabel *et al.* [76, 77] in the previous section. On making these assumptions the only remaining parameter of the model is the Young's modulus E_Y. This is, by definition, the gradient of the stress-strain curve and here is assumed to be an exponential function of the strain, i.e.

$$E_Y = \frac{d\sigma}{dE_{\max}} = c_1 e^{c_2 E_{\max}}, \tag{A.1}$$

where σ is the Cauchy stress, $E_{\max}$ the largest eigenvalue of Cauchy's infinitesimal strain tensor and c_1, c_2 are constants. Azar *et al.* [4] calculated the additional strain at each step

using Cauchy's infinitesimal strain tensor and the total strain by adding up the strains generated at each step. Integrating the previous equation and noting that $T = 0$ when $E = 0$ gives the constitutive equation

$$T = \frac{c_1}{c_2}\left(e^{c_2 E_{\max}} - 1\right). \tag{A.2}$$

Integrating again gives the strain energy function

$$W = \frac{c_1}{c_2}\left(\frac{1}{c_2}e^{c_2 E_{\max}} - E_{\max} - \frac{1}{c_2}\right), \tag{A.3}$$

where the constant of integration is chosen so that $W = 0$ when $E = 0$, i.e. when the material is unstrained it has no strain energy.

Simulations of the deformation of soft tissue in the absence of gravity using this model (with extremely coarse meshes) are presented in Azar *et al.* [4] and it is claimed that the model gives accurate results. However, gravity does exist and simulations by Whiteley *et al.* [89] that include it show serious flaws in the model and its underlying assumptions. Firstly, modelling incompressibility by setting $\nu = \frac{1}{2} - \epsilon$ is not ideal — using this method means that the material is actually slightly *compressible* and small changes in the choice of $\epsilon \ll 1$ can have a significant effect on the value of the Lamé constant λ and hence on the size of the displacements seen. Alternative methods for enforcing the incompressibility constraint more rigorously do exist, e.g. one is described in Whiteley [89]. The key observation here, however, is that the simulations differ depending on how incompressibility is modelled.

A second problem with the method is that calculating the strain at successive steps by repeated use of Cauchy's infinitesimal strain tensor will lead to the accumulation of errors in the total strain being calculated. This will affect the ability of the model to predict accurately the effects of large deformations — although not to the same extent as in a classical linear model. Finally, the assumption of isotropy is not valid (even for a genuinely isotropic material) for any of the steps except the first: if the material is already stressed then this will cause it to respond differently to forces applied in various different direction (see Whiteley *et al.* [89]). The consequences of this are seen in comparisons of the results of the pseudo-linear model with a fully nonlinear finite elasticity model (which will be discussed later) by Whiteley *et al.* [89]. The results are seen to be significantly different indicating that the pseudo-linear model is not a good approximation to a fully nonlinear model.

A.2 Finite Elasticity

As discussed in Section A.1 soft tissues are not, in general, linearly elastic under finite strains. In this section models based on the theory of finite elasticity are discussed. Work on modelling soft tissue using finite elasticity has been done since the 1960's. In light of this observation it is difficult to understand the motivation behind the recent work on classical linear elasticity described in Section A.1.1. Much of the work in finite elasticity involves the empirical determination of constitutive equations or strain energy functions for soft tissue.

A.2.1 Isotropic Hyperelasticity

The first nonlinear elastic material to be studied in any detail was rubber. Materials with rubber-like behaviour are referred to as *hyperelastic* (or *Green elastic*). Early textbooks such as Green and Adkins [34] were motivated by the study of hyperelastic materials. Much of this work included an assumption of isotropy based in the observed behaviour of rubber. As a result of this isotropic hyperelasticity is well understood in comparison with more general finite elasticity. This familiarity has, as in the case of linear elasticity, led to the development of isotropic hyperelasticity models for the behaviour of soft tissue even though the tissue itself is *not* a rubber-like material or, in general, even isotropic.

A *hyperelastic* material is defined to be a material whose elastic behaviour is determined by an analytic strain energy function W dependent on position and the right Cauchy-Green deformation tensor, i.e. $W = W(\mathbf{X}, C)$. In particular, it is independent of the stress. Now according to the polar decomposition theorem (Spencer [80]) the deformation gradient tensor can be written as

$$F = RU, \tag{A.4}$$

where R is an orthogonal (rotation) tensor and U is a symmetric and positive definite tensor. We define a *principal value* λ and principal direction $\mathbf{A}$ to be solutions of

$$(U - \lambda I)\mathbf{A} = \mathbf{0}, \tag{A.5}$$

i.e. the line with initial direction $\mathbf{A}$ stretches by a factor λ but does not rotate during the deformation described by U. Spencer [80] shows that since U is symmetric and positive

definite it has 3 positive real principal values $\lambda_1 \geqslant \lambda_2 \geqslant \lambda_3$, known as the *principal stretches* of U, and, if the principal values are distinct, U also has three distinct, orthogonal principal directions $\mathbf{A}_1$, $\mathbf{A}_2$, $\mathbf{A}_3$. These define the *principal axes* of U. If the coordinate system is chosen to coincide with the principal axes then

$$U = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}. \tag{A.6}$$

So, U is simply a series of stretches with magnitudes given by its principal stretches along the corresponding principal axes. Hence, it can be seen that F consists of these three stretches followed by a rotation defined by R.

Notice that since λ_1, λ_2, λ_3 are invariants of F, λ_1^2, λ_2^2, λ_3^2 are invariants of the Cauchy-Green deformation tensors B and C, which define the strain. Then clearly the strain energy function of a hyperelastic material can be written as

$$W(\mathbf{X}) = W(\mathbf{X}, \lambda_1, \lambda_2, \lambda_3), \tag{A.7}$$

since the rotation will not contribute to the strain on the material and U is completely specified by its principal stretches. Physical considerations (zero strain energy and zero stress in the unstrained state) require that:

$$W(\mathbf{X}, 1, 1, 1) = \frac{\partial W}{\partial \lambda_i}(\mathbf{X}, 1, 1, 1) = 0, \quad i = 1, 2, 3. \tag{A.8}$$

Three *strain invariants* can be defined (see Spencer [80]) by:

$$I_1 = \operatorname{tr} C = \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \tag{A.9}$$

$$I_2 = \tfrac{1}{2}\left((\operatorname{tr} C)^2 - \operatorname{tr} C^2\right) = \lambda_1^2\lambda_2^2 + \lambda_2^2\lambda_3^2 + \lambda_3^2\lambda_1^2, \tag{A.10}$$

$$I_3 = \det C = \lambda_1^2\lambda_2^2\lambda_3^2. \tag{A.11}$$

For an isotropic material, W must be symmetric in λ_1, λ_2, λ_3, so, in this case, an alternative formulation for W is

$$W(\mathbf{X}) = W(\mathbf{X}, I_1, I_2, I_3), \tag{A.12}$$

with the physical requirements becoming

$$W(3,3,1) = \frac{\partial W}{\partial \lambda_i}(3,3,1) = 0, \quad i = 1,2,3, \tag{A.13}$$

where the derivatives with respect to λ_i can be calculated using the chain rule.

A.2.2 Modelling the Strain Energy Function

Since W is assumed to be infinitely continuously differentiable it can be written as a Taylor series about the equilibrium point $(I_1, I_2, I_3) = (3,3,1)$ (see Ogden [55]). This leads to the infinite sum:

$$W(I_1, I_2, I_3) = \sum_{p,q,r=0}^{N\to\infty} c_{pqr}(I_1 - 3)^p (I_2 - 3)^q (I_3 - 1)^r. \tag{A.14}$$

Physical constraints require $c_{000} = 0$ (strain energy vanishes in undeformed state) and $c_{100} + 2c_{010} + c_{001} = 0$ (stress free when undeformed). Ogden [55] decomposes the strain energy further to obtain

$$W(I_1, I_2, I_3) = \sum_{p,q=0}^{N\to\infty} c_{pq0}(I_1 - 3)^p (I_2 - 3)^q + \sum_{r=0}^{N\to\infty} c_{00r}(I_3 - 1)^r. \tag{A.15}$$

Here the first term on the right-hand-side represents the *isochoric* strain energy (i.e. work done in deformation that does not alter the volume of the material) and the second term, often written simply as $g(I_3)$, the *volumetric* strain energy (i.e. work done to alter the volume of the material during deformation). For an incompressible material recall that $\det F = \lambda_1\lambda_2\lambda_3 \equiv 1$. Hence $I_3 = 1$ and the volumetric term vanishes (as expected, since the volume of an incompressible material must remain constant). Most models of hyperelastic materials terminate the sum at some finite value of N, giving rise to a strain energy function that is a polynomial in I_1, I_2 and I_3 (and, hence, in λ_1, λ_2 and λ_3).

Various models for the strain energy function of an incompressible, isotropic hyperelastic material have been proposed including:

- The neo-Hookean model [54]:

$$W = \frac{G}{2}(I_1 - 3), \quad I_3 \equiv 1, \tag{A.16}$$

where G is the shear modulus of the material. This model is derived from the Gaussian molecular statistical theory of rubber. However, while it is accurate for small deformations its qualitative and quantitative predictions for large deformations are inaccurate.

- The Mooney-Rivlin model [72]:

$$W = C_1(I_1 - 3) + C_2(I_2 - 3), \quad I_3 \equiv 1, \tag{A.17}$$

where C_1, C_2 are material parameters. This model results from truncating the summation in equation (A.15) by setting $N = 1$ and then choosing $c_{100} = C_1$, $c_{010} = C_2$, $c_{110} = 0$ (recall that the second sum vanishes since we are considering an incompressible material). According to Ogden [55] the Mooney-Rivlin model predicts quantitatively correct behaviour for small deformations of rubber and qualitatively correct behaviour for large deformations.

- The Ogden model [55]:

$$W = \sum_{p=1}^{N\to\infty} \frac{\mu_p}{\alpha_p}\left(\lambda_1^{\alpha_p} + \lambda_2^{\alpha_p} + \lambda_3^{\alpha_p} - 3\right), \quad I_3 \equiv 1, \tag{A.18}$$

where α_p, μ_p are material parameters. The Ogden model is in effect a generalization of the neo-Hookean and Mooney-Rivlin models. Ogden [55] claims that it gives a good quantitative prediction of the deformation of rubber when $N \geqslant 3$.

Many other models for the isochoric strain energy function exist. The behaviour of compressible materials can be predicted by the development of a model for the volumetric strain energy function. Many such models also exist and a fuller discussion of them can be found in Ogden [54]. According to Humphrey [39] the use of the above three models in soft tissue modelling is relatively common. However, all models of this type assume that the hyperelastic material is isotropic. As has been discussed above this is not, in general, the case for soft tissue. In addition, soft tissue does not behave like rubber — it does not deform elastically — so care must be taken when using hyperelastic models.

A.3 Pseudoelasticity and Fung Elasticity

All of the models discussed so far have been developed theoretically. Since they were developed primarily to describe the behaviour of metals (linear elasticity) and rubber (hyperelasticity) any comparison with experimental results during the development of the model would not have included tests on soft tissue. As subsequent work has been done on modelling soft tissue using these methods some comparison between the observed behaviour of these materials and the models has been done. For example, in Section A.1.1 Pathmanathan's [60] validation of the classical linear elastic model was discussed. Lots of other examples of this type of comparison can be found in the literature; however, whilst they are useful for optimizing a particular model for the study of soft tissue they do not consider the possibility that an altogether different model may be more appropriate. In the case of soft tissue modelling this is very likely to be the case when elastic models are being considered since, as has been discussed previously, soft tissue does not behave in an elastic manner. To overcome this difficulty a model must be constructed with reference to the results of experiments on soft tissue. In this section a model of this type is described.

A.3.1 Uniaxial Loading Experiments

The simplest physiological experiment that can be performed on soft tissue is to stretch repeatedly a cylindrical sample in a direction parallel to the axis of the cylinder. This is known as *uniaxial loading.*[1] Experiments of this type show, according to Fung [30], that many soft tissues undergo hysteresis when subject to cyclic uniaxial loading. Soft tissue can also be seen to exhibit creep and relaxation — further characteristics of viscoelasticity. However, Fung [30] observed that in most cases the hysteresis was largely independent of the strain rate. Additionally, Fung [30] showed that, although initially the soft tissue traced out a different hysteresis loop under each cycle of loading and unloading, it often converged to a well-defined and repeatable loop after a number of *preconditioning* cycles. Fung [30] refers to such materials as *pseudoelastic.* A pseudoelastic material has, after preconditioning, predictable stress-strain relationships for loading and unloading. The behaviour of the inelastic soft tissue can then be modelled by treating it as *two separate*

[1]Note that all of these experiments are done *in vitro.* Since the conditions within live tissue cannot be recreated *in vitro* it is not clear how good the results of these experiments are for predicting the behaviour of real tissue *in vivo.*

elastic materials: one for loading and one for unloading.

A.3.2 Fung's 3-d Exponential Strain Energy Function Postulate

In his experiments Fung [30] also quantified the stress-strain relationship of various pseudoelastic soft tissues under loading and unloading. This work showed that a good first approximation for the stress-strain relationship is:

$$\frac{dT_{11}}{d\lambda} = \alpha(T_{11} + \beta), \tag{A.19}$$

where T_{11} is the uniaxial second Piola-Kirchoff stress, λ is the stretch ratio (see Section 2.1.1.2) and α, β are material constants. This relationship is valid, with different values of α and β for both the loading and unloading phases of the cycle. On integrating, it can be shown that the stress-strain relationship is exponential, since

$$T_{11} = Ae^{\alpha\lambda} - \beta, \tag{A.20}$$

where A is a constant of integration and the stretch ratio can be used as a measure of strain. For increased accuracy Fung [30] suggests fitting a series of different exponential curves to various segments of the experimentally observed stress-strain curve. Based on this exponential relationship in 1-d between the second Piola-Kirchoff stress and the stretch ratio, the existence of 3-d pseudoelastic materials and a 3-d pseudostrain energy function can be postulated (see Humphrey [39]):

$$W = c\big(e^{Q} - 1\big), \tag{A.21}$$

where $Q = Q(E)$ is some function of the Lagrange-Green strain tensor and c is a material parameter. This gives an exponential stress-strain relationship for a 3-d pseudoelastic material since, from equation (2.70), it can be seen that

$$T_{MN} = \frac{\partial W}{\partial E_{MN}} = ce^{Q}\frac{\partial Q}{\partial E_{MN}}. \tag{A.22}$$

Materials with constitutive equations of this form are often referred to as *Fung elastic*. To complete the constitutive equation determination of Q is necessary. The quadratic form

$$Q = \tfrac{1}{2} a_{ABCD} E_{AB} E_{CD}, \tag{A.23}$$

where a is a rank 4 tensor of material parameters, is suggested by Fung [29] and Fung *et al.* [31] as an appropriate form for Q for most soft tissues. This is based on fits to experimental stress-strain curves and the ease of estimating parameters. As was the case for the linearly elastic model the number of material parameters in a can usually be shown to be significantly less than the theoretical maximum of 81. In particular, a must be symmetric. In fact, only 9 material parameters are required for an orthotropic material, 5 for a transversely isotropic material and 2 for an isotropic material. Even for a completely anisotropic material it can be shown that only 21 parameters can be freely chosen.[2] These numbers are exactly the same as was the case for a linear elasticity model.[3]

Other forms of Q have also been proposed. Whiteley *et al.* [89] use

$$Q = b(I_1 - 3), \tag{A.24}$$

where b is a constant and I_1 the first strain invariant, defined in terms of the components of strain by

$$I_1 = 3 + 2E_{MM}. \tag{A.25}$$

The strain energy function resulting from this choice of Q was first proposed by Veronda and Westmann [86] and can be seen to be an isotropic hyperelastic model.

---

[2]Note here that a material is defined to be *isotropic* if, at any given point, it has the same mechanical properties in all directions. It is said to be *transversely isotropic* if, for any given point, there exists a vector **a** (known as the *axis of transverse isotropy*) such that the material has the same mechanical properties in all directions lying in the plane normal to **a**. An *orthotropic* material has three mutually perpendicular planes of elastic symmetry at any given point. If the coordinate planes are chosen to coincide with the planes of symmetry then this statement is equivalent to saying that the axial stresses depend only on the axial strains, i.e. $T_{MM} = T_{MM}(E_{11}, E_{22}, E_{33})$, $M = 1, 2, 3$ (no summation), and that T_{MN}, $M \neq N$, depends only on E_{MN}. A material with no elastic symmetries at all is referred to as *anisotropic*.

[3]Recall, however, that separate models are required for the loading and unloading of a Fung elastic model. This effectively doubles the number of material parameters required in comparison with the linear elasticity model.

A.3.3 Polynomial Pseudoelasticity

A second class of pseudoelastic models also exists in which the strain energy function takes the form of a polynomial in the components of strain (see Patel and Vaishnav [59], Vaishnav *et al.* [85], Wesley *et al.* [88]), i.e.

$$W = P, \tag{A.26}$$

where $P = P(E)$ is a polynomial function. Observe that the requirement that the strain energy function vanishes in the undeformed state ($E = 0$) means that the constant term in the polynomial must be zero. A strain energy function of this form gives rise to a polynomial stress-strain relationship since

$$T_{MN} = \frac{\partial W}{\partial E_{MN}} = \frac{\partial P}{\partial E_{MN}}. \tag{A.27}$$

Recall that in the discussion of hyperelasticity in Section A.2.1 it was also assumed that there should be no stress in the unstrained state. This condition can be enforced by the inclusion of an internal pressure term, which is included when the material is incompressible.

Polynomial pseudoelasticity is similar to the approach taken in Section A.2.1 for an isotropic hyperelastic material, except that for a pseudoelastic material *different* polynomials are required for the loading and unloading of the material and here no assumption of symmetry in the strain components is made (i.e. anisotropy is allowed). Since anisotropic hyperelastic models are much less common than isotropic ones (and since rubber is isotropic) the theory has developed much more with soft tissue in mind than was the case with the models of Section A.2.1.

A.3.4 Combining the Two Pseudoelastic Models

Finally, the two classes can be combined to produce a model in which the strain energy function takes the form

$$W = P + c(e^Q - 1), \tag{A.28}$$

where $P = P(E)$ is a polynomial and $Q = Q(E)$ a quadratic form. For example, Zhou and Fung [92] have shown that the strain energy function

$$W = c(e^Q - Q - 1) + q, \tag{A.29}$$

where q is a second quadratic form,

$$q = b_{ABCD}E_{AB}E_{CD}, \tag{A.30}$$

can be used to produce a very accurate prediction of the (anisotropic) behaviour of blood vessels. Notice that since Q is a quadratic form (and, hence, a polynomial) this model is of the form given in equation (A.28). Recall also that, as for all pseudoelastic materials, different models (in this case different quadratic forms) are used to describe the behaviour under loading and unloading. This example emphasizes the need to construct an appropriate strain energy function for each material studied rather than trying to alter the parameter values of a particular model to ensure that it fits every material.

Appendix B

Viscoelastic and Mixture Models of Soft Tissue

A selection of non-elastic models with potential applications for describing the mechanics of soft tissue are described in this appendix. These models can be separated into those that consider soft tissue to be a single-phase viscoelastic material and those that model it as a two-phase solid-fluid mixture. The appendix is split into two sections: first Section B.1 introduces the mathematical theory of viscoelasticity and mixture models is introduced; then, in Section B.2 this theory is applied to develop some possible models appropriate for the simulation of the deformation regions of soft tissue during the ultrasound imaging process.

B.1 Theory

The mathematical theory of elastic solids was described in Section 2.1.3. In this section that theory is extended to deal with materials that, although having the outward appearance of solids (e.g. they maintain their shape rather than flowing like fluids), do not deform in the way that would be expected for a classical solid. For example, they may display hysteresis, creep or stress relaxation — see below for descriptions of these phenomena. There are two ways to deal with this type of behaviour, which are dealt with in turn here. Section B.1.1 describes the theory of *viscoelasticity*, where the way in which a material deforms depends on how it has been deformed in the past. Alternatively, the material could be modelled as

a mixture of solid and fluid, so that it behaves like a solid in some respects and like a fluid in others. The theory that can be used to describe these mixtures is discussed in Section B.1.2.

B.1.1 Viscoelasticity

The descriptions of fluid and solid mechanics given in the Sections 2.1.2 and 2.1.3 respectively are idealizations. Real materials do not behave exactly as suggested by the theories, although in certain conditions their behaviour may be accurately predicted by an appropriate law. In some cases, however, entirely different theories are required. One example of such a theory is *viscoelasticity*. A viscoelastic material will exhibit behaviour that is a combination of that expected for fluids and solids. For example, a solid is expected to maintain a constant deformation under a constant stress. However, a viscoelastic solid may continue to deform slowly over time when under constant stress as shown in Figure B.1. This phenomenon is referred to as *creep*. In a viscoelastic fluid creep is manifested as a gradual increase in the rate of strain that continues some time after a sudden increase in the stress of the system (in a Newtonian fluid the rate of strain would be expected to be proportional to the stress). Similarly, when a constant strain is applied to a solid it is expected that this will induce a constant stress. For a viscoelastic solid this may not be the case — the stress may gradually decrease with time. This is known as *stress relaxation* or often simply *relaxation* and is shown in Figure B.2. Similarly, for a viscoelastic fluid the abrupt cessation of a steady flow (constant rate of strain) leads to stress relaxation rather than an immediate return to a stress free state as would be predicted by the Newtonian model. Finally, if a viscoelastic material is subjected to cyclic loading, the stress-strain relationship can be seen to differ between the processes of loading and unloading. This effect is called *hysteresis*. Creep, relaxation and hysteresis characterize viscoelastic behaviour. Theories of viscoelastic behaviour can be broadly separated into two classes — differential and integral models. Each class is discussed in turn in this Section. A fuller discussion of viscoelasticity can be found in Christensen [19].

B.1.1.1 Differential Models

The goal of a *differential model* of viscoelasticity is to derive differential equations that describe the motion. One way of doing this is by considering mechanical analogues of the

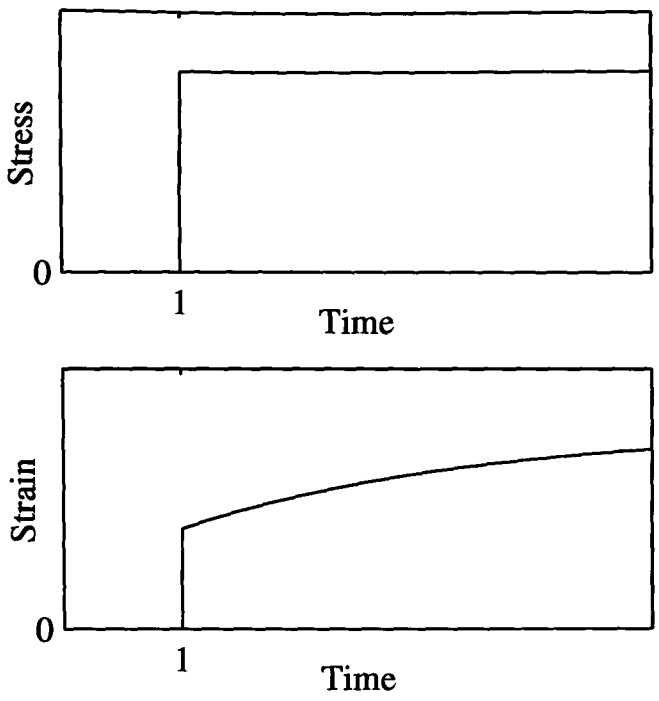


Figure B.1: Creep in a viscoelastic solid: strain increases over time in a material subject to a constant stress.

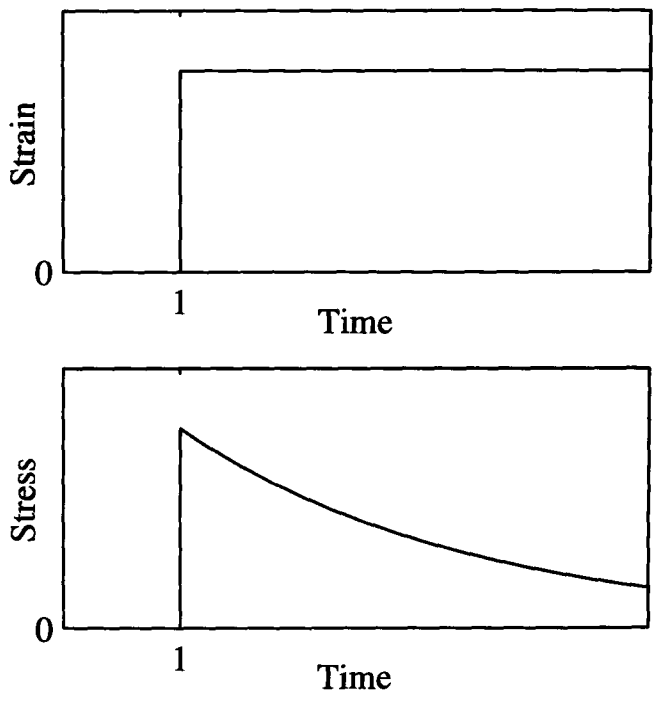


Figure B.2: Stress relaxation in a viscoelastic solid: stress decreases over time in a material where the deformation remains constant.

system under consideration. The behaviour of an elastic solid is analogous to that of a linear spring — deformation is instantaneous and proportional to the applied load, i.e. $F = \mu u$, where F is the force, u the extension of the spring and μ the elastic modulus. For a Newtonian viscous fluid the analogue is with a dashpot — this gives rise to a deflection with velocity in proportion to the load, i.e. $F = \eta \dot{u}$, where $\dot{u}$ is the velocity of deflection (throughout this section dots will be taken to denote differentiation with respect to time) and η is a constant of proportionality. A viscoelastic material can be approximated by connecting a series of springs and dashpots and considering the overall behaviour of the system. Three possible models are shown in Figure B.3.

In the *Maxwell model* [19] a spring and a dashpot are connected in series. This model was introduced to describe a fluid that is to some extent elastic. The same force $F = F(t)$, where t represents time, is transmitted from the spring to the dashpot, producing a displacement $u = F/\mu$ in the spring and a velocity $\dot{u} = F/\eta$ in the dashpot. On differentiating the displacement of the spring it can be seen that the governing equation for the velocity of the Maxwell model is

$$\dot{u} = \frac{\dot{F}}{\mu} + \frac{F}{\eta}. \tag{B.1}$$

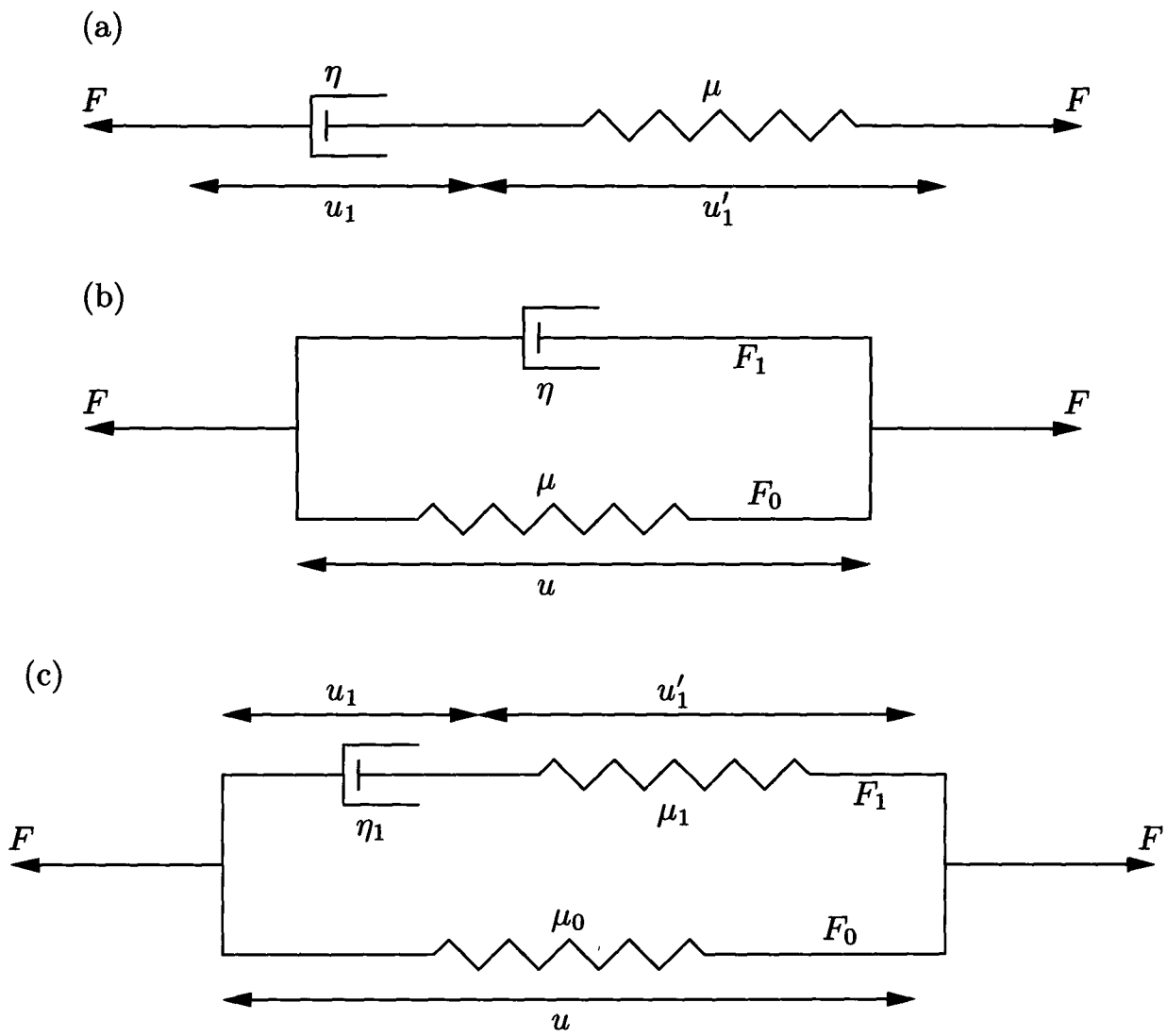


Figure B.3: Three mechanical models of a viscoelastic material consisting of springs and dashpots: (a) the Maxwell model, (b) the Voigt model, and (c) the Kelvin (or standard linear) model.

Further, if the force is instantaneously applied at $t = 0$ then the spring will suddenly deform with extension $F(0)/\mu$ but the dashpot will have no time to deform. Hence, an appropriate initial condition for the Maxwell model is

$$u(0) = \frac{F(0)}{\mu}. \tag{B.2}$$

In the *Voigt model* [19] the spring and dashpot are connected in parallel and thus have the same displacement u. So, they produce forces μu and $\eta \dot{u}$ respectively, giving the governing equation

$$F = \mu u + \eta \dot{u}. \tag{B.3}$$

If the force is suddenly applied an appropriate initial condition for the Voigt model is

$$u(0) = 0, \tag{B.4}$$

since there is no time for the dashpot to move and so the spring cannot deform either (so that their displacements remain equal). Hence, all of the initial force is channelled through the dashpot.

Finally, the *Kelvin* or *standard linear model* consists of a spring (elastic modulus μ_1) and dashpot in series (Maxwell element) connected in parallel with a spring (elastic modulus μ_0). This model was introduced when it was shown that the Maxwell and Voigt models could not account for the rate of dissipation of energy in some materials when subjected to cyclic loading. It is the most general model to include the load, the deflection and their first (or 'linear' derivatives). The displacement of the single spring u can be broken down into the displacements of the dashpot u_1 and the spring u_1' in the Maxwell element. Similarly, the total force F is the sum of the forces in the single spring $F_0(=\mu_0 u)$ and the Maxwell element $F_1(=\eta_1\dot{u}_1 = \mu_1 u_1')$. These observations allow the total force to be calculated as

$$F = \mu_0 u + \mu_1 u_1' = (\mu_0 + \mu_1)u - \mu_1 u_1. \tag{B.5}$$

This expression can be differentiated and, after some algebra, the governing equation for the standard linear model can be written in the form

$$F + \tau_\epsilon \dot{F} = E_R(u + \tau_\sigma \dot{u}), \tag{B.6}$$

where

$$\tau_\epsilon = \frac{\eta_1}{\mu_1}, \quad \tau_\sigma = \frac{\eta_1}{\mu_0}\left(1 + \frac{\mu_0}{\mu_1}\right), \quad E_R = \mu_0. \tag{B.7}$$

For a suddenly applied force at $t = 0$ the appropriate initial condition can be seen to be

$$\tau_\epsilon F(0) = E_R \tau_\sigma u(0). \tag{B.8}$$

The constant τ_ϵ is called the *relaxation time for constant strain*, τ_σ is the *relaxation time for constant stress* and E_R the *relaxed elastic modulus*. The reasons for this will become apparent in the subsequent discussion. More general models of viscoelastic materials can be constructed by adding additional Maxwell or Voigt elements to the Kelvin model.

Solving equations (B.1), (B.3) and (B.7) for $u(t)$ with $F(t) = H(t)$, where $H(t)$ is the *Heaviside function*, defined to be zero when $t < 0$ and one when $t > 0$, gives results known as *creep functions* $c(t)$ which show the time dependence of the elongation seen after the application of a constant force at $t = 0$. For the models discussed above the creep functions

are:

- Maxwell model:

$$c(t) = \left(\frac{1}{\mu} + \frac{t}{\eta}\right) H(t). \tag{B.9}$$

- Voigt model:

$$c(t) = \frac{1}{\mu}\left(1 - e^{-\mu t/\eta}\right) H(t). \tag{B.10}$$

- Standard linear solid:

$$c(t) = \frac{1}{E_R}\left[1 - \left(1 - \frac{\tau_\epsilon}{\tau_\sigma}\right) e^{-t/\tau_\sigma}\right] H(t). \tag{B.11}$$

Hence, a sudden application of force on a Maxwell solid induces an immediate deformation (of the spring) followed by linear 'creep' (of the dashpot). For the Voigt model there is no immediate deflection (since the dashpot cannot move instantaneously) but instead, a gradual increase in displacement (as the spring takes a greater share of the load). The rate of increase in displacement decreases exponentially, with characteristic time scale η/μ. This time scale is referred to as the *relaxation time.* In the standard linear model a combination of these two behaviours is seen — there is an initial deflection which is then followed by an increase in deflection, the rate of which decreases exponentially, with relaxation time τ_σ. As $t \to \infty$ the deflection tends to that which would be expected in the single spring — hence E_R represents the elastic modulus of the system in its fully relaxed state.

Similarly, solving for $F(t)$ when $u(t) = H(t)$ gives the *relaxation functions* $k(t)$ of the models. Physically, these represent the force that must be applied to maintain a unit elongation for $t > 0$. They are given for the models discussed here by:

- Maxwell model:

$$k(t) = \mu e^{-\mu t/\eta} H(t). \tag{B.12}$$

- Voigt model:

$$k(t) = \eta\delta(t) + \mu H(t), \tag{B.13}$$

  where $\delta(t)$ is the delta function, defined to be zero everywhere except at $t = 0$, but with the property $\int_{-\infty}^{\infty} \delta(t)dt = 1$.

- Standard linear solid:

$$k(t) = E_R\left[1 - \left(1 - \frac{\tau_\sigma}{\tau_\epsilon}\right)e^{-t/\tau_\epsilon}\right]H(t). \tag{B.14}$$

So, the sudden deformation of a Maxwell solid gives rise to a sudden increase in force (in the spring), which then decreases exponentially towards zero with relaxation rate η/μ (as the dashpot dissipates the energy). For a Voigt solid the force becomes very large at $t = 0$ and then instantly reduces to a constant value μ. The force in the linear elastic solid increases immediately and then decays away exponentially, with relaxation rate τ_ϵ, towards the force that would be present instantaneously in the single spring if the Maxwell element was absent.

B.1.1.2 Integral Models

The result of an *integral model* for a viscoelastic material is the derivation of an integral equation that describes the behaviour of the material. Here one example of such a method is given (for a fuller description of the method see, e.g., Fung [27]). In 1-d consider a bar subjected to a force $F(t)$ and having extension $u(t)$ which is dependent on the entire history of the loading up until time t. If $F(t)$ is differentiable then in a small time interval $(\tau, \tau + d\tau)$ the increment of loading is $F'(\tau)d\tau$. This increment continues to act on the bar for all $t > \tau$ and contributes an extension $du(t)$ to the elongation at time t. If it is assumed that the size of the extension is proportional to the increment of loading that induced it then it can be seen that

$$du(t) = c(t - \tau)\frac{dF(\tau)}{d\tau}d\tau, \tag{B.15}$$

where $c(t-\tau)$ is a (time dependent) constant of proportionality. Suppose it is assumed that the motion begins at $t = 0$ and that the individual extensions can be integrated over the history of the bar to determine the extension at time t (this is known as the *superposition principle*). Then,

$$u(t) = \int_0^t c(t - \tau)\frac{dF(\tau)}{d\tau}d\tau. \tag{B.16}$$

A similar argument with the roles of F and u interchanged can be used to show that

$$F(t) = \int_0^t k(t - \tau)\frac{du(\tau)}{d\tau}d\tau, \tag{B.17}$$

where $k(t-\tau)$ is a second (time dependent) constant of proportionality. The functions $c(t-\tau)$ and $k(t-\tau)$ are known as the *creep* and *relaxation functions* respectively. This is the most general formulation of a viscoelastic material in which it is assumed that instantaneous force and extension are proportional to one another and is due to Boltzmann [12]. The *principal of fading memory* [20] imposes an additional constraint on the formulation of a constitutive equation for a viscoelastic material. It states that "deformations which occurred in the distant past have a smaller effect on the present forces than have recent deformations". The models of the previous section can be seen to be special cases of this integral formulation for a general viscoelastic material.

B.1.1.3 Linear Viscoelasticity

A theory of *linear viscoelasticity* can be developed by generalizing equations (B.16) and (B.17) into constitutive equations for a 3-d material. For infinitesimal deformations this turns out to be very simple: according to Fung [27] if we replace F and u by the Cauchy stress tensor and Cauchy's infinitesimal strain tensor respectively then replacing c and k with appropriate tensors defining the creep and relaxation of the material turns out to be sufficient to give constitutive equations. So, a material is said to be *linear viscoelastic* if

$$\sigma_{ij}(\mathbf{x},t) = \int_{-\infty}^{t} G_{ijkl}(\mathbf{x},t-\tau)\frac{\partial \epsilon_{kl}}{\partial \tau}(\mathbf{x},t)d\tau, \tag{B.18}$$

or, equivalently,

$$\epsilon_{ij}(\mathbf{x},t) = \int_{-\infty}^{t} J_{ijkl}(\mathbf{x},t-\tau)\frac{\partial \sigma_{kl}}{\partial \tau}(\mathbf{x},t)d\tau, \tag{B.19}$$

where G is the *tensorial relaxation function* and J the *tensorial creep function.* The lower limit of integration is taken as $-\infty$ since *all* previous motion must be taken into account. Care must be taken with finite deformations since the stress and strain rates may then depend on the rotation of the material rather than just its extension.

B.1.2 Mixture Theory

An alternative way of accounting for stress relaxation, creep and hysteresis is to consider the viscoelastic material as a mixture of solid and fluid. The behaviour of the two different constituents of the mixture can be modelled separately and then combined to produce a

model of the composite. The differential theory of viscoelasticity can be considered to be a mixture model since it considers a mixture of springs (solid) and dashpots (fluid). However, most complex viscoelastic models are formulated as integrals so this relationship is not immediately apparent. In this section one of the earliest mixture models (Biot's theory of poroelasticity) is discussed and a general framework for the construction of a general mixture model is described.

B.1.2.1 Biot's Theory of Poroelasticity

The theory of fluid flow through a porous solid medium can be considered to begin with the development of Darcy's law (see Darcy [21], Appendix, Note D) in 1856. In 1-d Darcy's law relates the pressure in the fluid (as a function of distance) $p(x)$ to the fluid flow rate $v(x)$ according to:

$$-\frac{dp}{dx} = kv(x), \tag{B.20}$$

in the absence of any external forces. This work was motivated by a desire to describe the flow of groundwater and, since the motion of the solid medium is not important in problems on this scale, Darcy's law deals only with the fluid motion. It was not until the middle of the 20th century that a theory of elasticity for the porous medium, known as *poroelasticity*, was developed. This theory was first proposed by Biot [9] for an isotropic medium and later extended, also by Biot [10] to the case of an anisotropic porous medium. In both cases the fluid flow is assumed to be governed by Darcy's law.

Biot [9, 10] considered a porous solid with *porosity* (volume fraction occupied by fluid) ϕ.[1] Suppose the displacements in the solid and the fluid are defined to be $\mathbf{u}^{(\mathrm{s})}$ and $\mathbf{u}^{(\mathrm{f})}$ respectively and the the pressure in the fluid is p. On generalizing Darcy's law to 3-d the fluid pressure can be seen to be governed by the relations

$$-\frac{\partial p}{\partial x_i} + \rho^{(\mathrm{f})} f_i = k_{ij}\left(\dot{u}_j^{(\mathrm{f})} - \dot{u}_j^{(\mathrm{s})}\right), \quad i = 1, 2, 3, \tag{B.21}$$

where $\rho^{(\mathrm{f})}$ is the density of the fluid and $\mathbf{f}$ the body force. It can be shown (see, e.g., Biot [10]) that the coefficients k_{ij} have the property that $k_{ij} = k_{ji}$. The porous solid is assumed to be linearly elastic, with its deformation governed by equation (2.58), the usual

[1] Observe here that it is assumed that the pore space is connected so that fluid can flow freely within it. Any unconnected pores are considered for the purpose of this theory to make up part of the solid medium.

governing equations for such a material. However, the stress tensor used here is not the usual Cauchy stress tensor, but is defined by the constitutive equations

$$\sigma_{ij} = -\phi p \delta_{ij} + \lambda \epsilon_{kk} \delta_{ij} + 2\mu \epsilon_{ij}, \tag{B.22}$$

$$p = g(\epsilon, d), \tag{B.23}$$

where $g(\epsilon, d)$ is some function of Cauchy's infinitesimal strain tensor in the solid ϵ and the *dilation* of the fluid $d = \nabla.\mathbf{u}^{(f)}$. In general g is determined by the specification of the strain energy function W as a function of ϵ and d.[2] These equations account for the force exerted on the solid by the fluid. General solutions of these equations in terms of the applied stress (or displacement) field have been derived by Biot [11].

B.1.2.2 The Modern Theory of Mixtures

The modern theory of mixtures can be considered to begin with the work of Truesdell (see, e.g., Truesdell and Noll [84]) on inter-diffusing materials. In this theory each point in the domain of interest is simultaneously occupied by all of the constituents of the mixture and two key principles govern the behaviour of the mixture:

- The properties of the motion of the constituents completely determines the mean motion of the mixture as a whole.
- If the effects of the interactions between constituents are properly accounted for then the governing equations for the mean motion of the mixture are the same as those governing the motion of a simple medium.

This means that the governing equations of the mixture must be the usual conservation laws for mass, momentum and energy (with appropriate constitutive equations) since these quantities must be conserved across the system as a whole. They can, however, be transferred between the different constituents of the mixture. So, the quantities $m^{(\alpha)}$, $\boldsymbol{\pi}^{(\alpha)}$ and $E^{(\alpha)}$, $\alpha = 1, \ldots, N$, where N is the number of constituents, are defined to be the changes in mass, momentum and energy respectively of the αth constituent. Truesdell and Noll

[2]This is an appropriate way of determining g since, for the stress tensor that is used here, strain energy can be shown to have the form $W = \sigma_{ij}\epsilon_{ij} - \phi pd$, by a generalization of the argument used to derive equation (2.68) for the Cauchy stress tensor and Cauchy's infinitesimal strain tensor.

[84] show that these quantities obey the following relations:

$$\sum_{\alpha=1}^{N} m^{(\alpha)} = 0, \tag{B.24}$$

$$\sum_{\alpha=1}^{N} \left(\boldsymbol{\pi}^{(\alpha)} + m^{(\alpha)}\mathbf{v}^{(\alpha)}\right) = 0, \tag{B.25}$$

$$\sum_{\alpha=1}^{N} \left[E^{(\alpha)} + \boldsymbol{\pi}^{(\alpha)}.\mathbf{v}^{(\alpha)} + m^{(\alpha)}\left(E^{(\alpha)} + \tfrac{1}{2}|\mathbf{u}^{(\alpha)}|^2\right)\right] = 0, \tag{B.26}$$

where $\mathbf{v}^{(\alpha)}$ is the *diffusion velocity* of the αth constituent, defined to be the velocity of that constituent relative to the mean velocity of the mixture. Substantial work has been done in extending this theory (see, e.g., Atkin and Craine [2] or Bowen [14] for developments up to 1976 and Rajagopal and Tao [70] for later work), however disagreements remain over the correct way of imposing second law restrictions and boundary conditions for the constituents and over the way in which the constitutive relations for mass, momentum and energy transfers should be determined.

B.2 Models

In this section the theory presented in Section B.1 is extended to provide models appropriate for use in the field of soft tissue mechanics. Two classes of viscoelastic models are considered. First, in Section B.2.1 the linear theory of viscoelasticity is extended to produce quasi-linear viscoelastic models. Then, in Section B.2.2 a variety of alternative, fully nonlinear viscoelastic models are described. Finally, a variety of mixture models that have been applied to describe soft tissue are discussed in Section B.2.3.

B.2.1 Quasi-Linear Viscoelasticity

As discussed in Section A.3.1 many soft tissues have been found to display viscoelastic characteristics. This is not surprising since they consist predominantly of water but with solid structures. One possible improvement on the pseudoelastic models discussed in the previous section is to model soft tissue using viscoelastic techniques. This has the potential to give more accurate results for those materials that can be preconditioned and

also provides a framework for predicting the behaviour of soft tissues that do not exhibit pseudoelastic behaviour even after preconditioning. Since the middle of the 20th century viscoelastic models of soft tissue have been proposed. However, Fung [30] and Humphrey [39] observe that, as was the case when using elastic models, the standard linear viscoelastic models discussed in Section B.1.1 are valid only for small strains. For finite deformations a nonlinear theory must be developed. One of the earliest attempts to develop an appropriate nonlinear theory of viscoelasticity was by Fung [28] in 1972. The results of that attempt are discussed in this section.

B.2.1.1 Deriving a Constitutive Equation

Consider a viscoelastic material in uniaxial loading and suppose a step increase in elongation (of magnitude λ) is imposed on the material. Then the resulting uniaxial second Piola-Kirchoff stress T_{11} will be a function of both time and the stretch λ. Suppose it is assumed that the history of the stress response $K(\lambda, t)$ (called the *relaxation function*) is defined by

$$K(\lambda, t) = G(t)T^e_{11}(\lambda), \quad G(0) = 1, \tag{B.27}$$

where $G(t)$ is a normalized function of time called the *reduced relaxation function* and $T^e_{11}(\lambda)$ is known as the *elastic response function.* For nonlinear viscoelasticity T^e_{11} is assumed to depend nonlinearly on λ. With this definition it is possible to proceed using the superposition principle as for the general linear integral formulation of Section B.1.1.2. The increment in stress at time t due to an infinitesimal stretch $d\lambda(\tau)$ applied to a specimen in a state of stretch λ at time $\tau < t$ is

$$dT_{11}(\lambda, t) = G(t-\tau)\frac{\partial T^e_{11}[\lambda(\tau)]}{\partial \lambda} d\lambda(\tau). \tag{B.28}$$

Then, on applying the superposition principle to sum the contributions of all the infinitesimal stretches gives the integral formulation

$$T_{11}(\lambda, t) = \int_{-\infty}^{t} G(t-\tau)\frac{\partial T^e_{11}}{\partial \lambda}\frac{\partial \lambda}{\partial \tau} d\tau. \tag{B.29}$$

The physical interpretation of this integral is that the tensile stress at time t is the sum of the contributions from all of the earlier states of stress, each scaled by the same reduced

relaxation function. Note that equation (B.29) can be rewritten as

$$T_{11}(\lambda, t) = \int_{-\infty}^{t} G(t-\tau)\dot{T}^{e}_{11}[\lambda(\tau)]d\tau, \tag{B.30}$$

where the dot denotes differentiation with respect to time. Then by comparing with equation (B.17) it can be seen that this model is equivalent to the general linear model but with the stress response function taking the role assumed by the strain in the conventional theory. Hence the method is referred to as *quasi-linear viscoelasticity* (QLV) theory. The inverse of equation (B.30) can be written as

$$T^{e}_{11}[\lambda(t)] = \int_{-\infty}^{t} J(t-\tau)\dot{T}_{11}(\lambda, \tau)d\tau, \tag{B.31}$$

where $J(t)$ is the *reduced creep function* and can be defined in terms of the history of the stretch function in a manner analogous to that used to define the reduced relaxation function in terms of the stress history. Equation (B.30) can be regarded as a constitutive equation — once $T^{e}_{11}(\lambda)$ and $G(t)$ are determined it fully determines the behaviour of the material.

Since soft tissue is not, in general, 1-d it is necessary to generalize equation (B.30) to a constitutive equation relating the second Piola-Kirchoff stress tensor to Green's strain tensor for a 3-d QLV material. According to Fung [30] this can be done simply by converting to a tensor equation:

$$T_{MN}(E, t) = \int_{-\infty}^{t} G_{MNPQ}(t-\tau)\frac{\partial T^{e}_{PQ}[E(\tau)]}{\partial \tau}d\tau, \tag{B.32}$$

where G is the *reduced relaxation function tensor* and T^{e} is the *elastic stress tensor* corresponding to the strain tensor E. The condition in 1-d that $G(0) = 1$ generalizes to

$$G_{MNPQ}(0) = 1, \tag{B.33}$$

for all (M, N, P, Q) and is the reason for the terminology 'reduced'. Note that in an isotropic material G will only have two independent components. For an anisotropic material G must be determined by consideration of the material symmetry as discussed in Green and Adkins [34]. This constitutive equation in 3-d is analogous to the generalization to 3-d for the linear model given by equation (B.18).

B.2.1.2 The Elastic Response

The elastic response function $T^e_{11}(\lambda)$ (in 1-d initially) is defined in equation (B.27) to be the tensile stress instantaneously generated in the material when a stretch λ is imposed. Experimental determination of this quantity is difficult since, according to Fung [30], sudden extension of a material in this way will almost certainly give rise to stress waves. However, if it is assumed that $G(t)$ is continuously differentiable then, noting that the motion begins at $t = 0$ and $G(0) = 1$, equation (B.30) can be seen to be equivalent to

$$T_{11}(\lambda, t) = T^e_{11}[\lambda(t)] + \int_0^t T^e_{11}[\lambda((t-\tau)]\frac{\partial G(\tau)}{\partial \tau}d\tau. \tag{B.34}$$

Suppose now that the stretch is increased from 0 to λ^* monotonically in some small time interval ϵ. Then,

$$T_{11}(\lambda^*, \epsilon) = T^e_{11}(\lambda^*) + \int_0^\epsilon T^e_{11}[\lambda(\epsilon-\tau)]\frac{\partial G(\tau)}{\partial \tau}d\tau. \tag{B.35}$$

As τ increases from 0 to ϵ the integrand is continuous, so, on applying the mean value theorem for integrals,[3] it can be seen that there exists $c \in [0, \epsilon]$ such that

$$T_{11}(\lambda^*, \epsilon) = T^e_{11}(\lambda^*) + \epsilon T^e_{11}[\lambda(\epsilon-c)]\frac{\partial G}{\partial \tau}(c). \tag{B.36}$$

Then, since $T^e_{11}(c)$ and $\partial G/\partial\tau(c)$ are finite the second term will tend to zero as $\epsilon \to 0$. So, by choosing ϵ to be sufficiently small we can approximate the elastic response function as

$$T^e_{11}(\lambda^*) \approx T_{11}(\epsilon). \tag{B.37}$$

So, $T^e_{11}(\lambda)$ can be approximated experimentally by a uniaxial loading experiment with a sufficiently high rate of loading. For a material with known material parameters $T_{11}(\lambda, \epsilon)$ can be determined theoretically using equation (A.20) and hence the stress response function can be approximated by the pseudoelastic stress.

[3] THEOREM (MEAN VALUE THEOREM FOR INTEGRALS):
Suppose $g : [a, b] \to \mathbb{R}$ is continuous. Then, there exists $c \in [a, b]$ such that $\int_a^b g(x)dx = (b-a)g(c)$.

B.2.1.3 The Reduced Relaxation Function

According to Fung [30] it is customary to write the reduced relaxation function as a sum of exponential functions and identify each exponent with the rate constant of a relaxation mechanism, i.e.

$$G(t) = \left(\sum_{k=1}^{N\to\infty} a_k \exp(-\nu_k t) \right) \Big/ \left(\sum_{k=1}^{N\to\infty} a_k \right), \tag{B.38}$$

where a_k, ν_k are constants and $\nu_i \leqslant \nu_{i+1}$ for all i.[4] Note that the experimental determination of $G(\infty)$ can be difficult. Fung [30] observes that this, together with the fact that the representation of empirical data by a sum of exponentials is non-unique, makes the experimental determination of $G(t)$ unreliable. One way of dealing with this is to observe that the amplitudes a_k corresponding to each characteristic frequency ν_k can be plotted as a *hysteresis spectrum.* If N is finite then the spectrum must be *discrete.* However, a fundamental assumption of QLV theory was that hysteresis was independent of strain rate (frequency). This is equivalent to requiring the hysteresis spectrum to be continuous (see Fung [30]) suggesting that the reduced relaxation function $G(t)$ should be written in the form of an integral rather than a sum. Observe that equation (B.14) for the standard linear model can be written as

$$k(t) = \frac{1}{1+S}\left(1 + Se^{-t/\tau_\epsilon}\right), \tag{B.39}$$

where τ_ϵ corresponds to one of the ν_k's of equation (B.38) and

$$S = \frac{\tau_\sigma}{\tau_\epsilon} - 1. \tag{B.40}$$

The natural way to extend this to a reduced relaxation function with a continuous spectrum is to replace τ_ϵ with a continuous variable τ and S with the integral of a function $S(\tau)$ over the range $\tau \in [0, \infty)$, i.e.

$$G(t) = \left(1 + \int_0^\infty S(\tau) e^{-t/\tau} d\tau\right) \left(1 + \int_0^\infty S(\tau) d\tau\right)^{-1} \tag{B.41}$$

[4]This is a generalization of equation (B.14) and can be considered to be the relaxation function associated with a system of N standard linear models connected in series with one another. The kth standard linear model has relaxation rate ν_k.

One specific example is to choose

$$S(\tau) = \begin{cases} \dfrac{c}{\tau}, & \text{for } 0 < \tau_1 \leqslant \tau \leqslant \tau_2 < \infty \\ 0, & \text{otherwise} \end{cases}, \tag{B.42}$$

where τ_1, τ_2 and c are constants. This gives the reduced relaxation function

$$G(t) = \left\{1 + c\left[E_1\left(\frac{t}{\tau_2}\right) - E_1\left(\frac{t}{\tau_1}\right)\right]\right\}\left[1 + c\log\left(\frac{\tau_2}{\tau_1}\right)\right]^{-1}, \tag{B.43}$$

where $E_1(z)$ is the exponential integral function defined (for $z > 0$) by

$$E_1(z) = \int_z^\infty \frac{e^{-\tau}}{\tau} d\tau. \tag{B.44}$$

Notice that as $t \to \infty$, $E_1(t/\tau_1)$ and $E_1(t/\tau_2) \to 0$ so that $G(\infty)$ can be calculated theoretically as

$$G(\infty) = \left[1 + c\log\left(\frac{\tau_2}{\tau_1}\right)\right]^{-1}, \tag{B.45}$$

where c, τ_1, τ_2 are chosen to fit experimental data. Recall that experimental determination of $G(\infty)$ is difficult. So, a better way of determining the parameter values is to observe that when $\tau_1 \ll t \ll \tau_2$, $G(\log t)$ is nearly a straight line, with

$$\frac{dG}{d(\log t)} = \frac{c}{1 + c\log(\tau_2/\tau_1)}. \tag{B.46}$$

According to Fung [30] the reduced relaxation function given by equation (B.43) with parameters determined using equation (B.46) produces results that agree with experimental findings. A complementary reduced creep function can be determined from this reduced relaxation function using Laplace transforms (see Fung [30]). However, this does not produce such good predictions of the behaviour of real tissues. Fung [30] hypothesizes that "creep is fundamentally more nonlinear [than relaxation], and perhaps does not obey the quasi-linear hypothesis."

B.2.2 Other Viscoelastic Models

Haut and Little [36] showed that at low strain rates (0.06–0.75%s^{-1}) the results of QLV theory agree with experimental findings. However, it was later shown by Woo *et al.* [90] that at higher strain rates (up to 10%) the predictions of QLV become less accurate. According to Humphrey [39] numerous other sources have since suggested that QLV is not sufficiently general to describe other aspects of the behaviour of many soft tissues. In particular, it cannot account for a strain-dependent relaxation rate [67] or the fundamentally different short-term and long-term viscous responses (or 'memories') exhibited by many soft tissues [64]. A stressed elastic tissue returns to its previous state instantaneously after the removal of the stress. A viscoelastic tissue exhibits the effects of short-term memory if it returns (within the timescale of a reasonable set of observations) to its unstressed configuration and exhibits the effects of long-term memory if it never returns to its original configuration.[5] Several alternatives to QLV that attempt to overcome one or both of these problems are briefly described in this section.

B.2.2.1 The Single Integral Model

Schapery [44, 74, 75] derived a theory of nonlinear viscoelasticity using thermodynamics. Treating strain as the independent state variable this theory can be written in the form of a single integral. For uniaxial loading,

$$T_{11}(E_{11},t) = h_e(E_{11})G_eE_{11} + h_1(E_{11})\int_0^1 \Delta G[\rho(t)-\rho'(\tau)]\frac{d}{d\tau}[h_2(E_{11})E_{11}]d\tau, \qquad \text{(B.47)}$$

where time has been rescaled so that the deformation occurs in the interval $t \in [0,1]$, the *reduced time* ρ and *reduced time variable of integration* ρ' are defined by

$$\rho = \int_0^t \frac{dt'}{a_e[E(t')]}, \quad \rho' = \int_0^\tau \frac{dt'}{a_e[E(t')]}, \qquad \text{(B.48)}$$

G_e is the equilibrium (or final) value of the elastic modulus (equivalent to $G(\infty)$ in QLV theory) and $\Delta G(\rho)$ is the transient modulus. The terms h_e, h_1, h_2 and a_e are strain-dependent, material properties (a_e is also time- and possibly temperature-dependent) that

[5] Note that according to QLV a material can only ever exhibit long-term memory since the reduced relaxation function has an exponential form so, at best, the deformation of a material from its original configuration can only *tend towards* zero as $t \to \infty$.

result from the thermodynamic analysis. The key difference between this model and the QLV equivalent (equation (B.30)) is that the time dependence of the relaxation (ΔG) is no longer independent of strain. A complementary relation for strain as a function of stress is given in Schapery [75]. Observe that when $h_e \equiv h_1 \equiv h_2 \equiv a_e \equiv 1$ equation (B.47) reduces to the 1-d version of the linear viscoelastic constitutive equation (B.18). According to Provenzano *et al.* [67], in polymers and fibrous composite materials in isothermal settings it is common to set $h_1 \equiv e_e \equiv 1$. This method can be generalized to 3-d by replacing the various 1-d functions described here by appropriate tensor-valued functions.

It has been shown (for example by Lakes and Vanderby [43] and Provenzano *et al.* [68]) that an appropriate form for the transient modulus of soft tissue (in this case ligament) at a constant temperature is

$$\Delta G(\rho) = C\rho^n \tag{B.49}$$

where C and n are material parameters. On substituting for ΔG in equation (B.47) this gives

$$T_{11}(E_{11}, t) = h_e(E_{11})G_e E_{11} + h_1(E_{11}) \int_0^1 (\rho - \rho')^n \frac{d}{d\tau}[h_2(E_{11})E_{11}]d\tau. \tag{B.50}$$

Then, for a typical uniaxial loading experiment with $E_{11}(t) = E_0 H(t)$ where $H(t)$ is the Heaviside function and assuming that $h_1 \equiv e_e \equiv 1$ as discussed above, the uniaxial stress can be calculated to be

$$T_{11}(E_0, t) = h_e(E_0)G_e E_0 + h_2(E_0)CE_0 t^n. \tag{B.51}$$

All of the terms in this expression are known (or are constants that can be determined by curve fitting in the case) with the exception of $h_e(E_0)$ and $h_2(E_0)$. So a series of uniaxial loading experiments can be performed to determine these functions and, hence, to predict the behaviour of the material under more complex loading.

B.2.2.2 Provenzano *et al.*'s Modified Superposition Model

Recall that the constitutive equation (B.18) for linear viscoelastic was derived by the superposition of contributions from each of a series of infinitesimal deformations scaled by the relaxation function $G(t)$. In order to model the strain-dependent relaxation rate observed in deformation of soft tissue Provenzano *et al.* [67] have adapted this model by

allowing the relaxation function to depend on strain, i.e., in uniaxial loading:

$$T_{11}(E,t) = \int_0^t G[t-\tau, E_{11}(t)]\frac{dE_{11}(\tau)}{d\tau}. \tag{B.52}$$

A complementary relationship can be formulated for the strain in terms of a stress-dependent creep function. Provenzano *et al.* assume a strain-dependent power law form for the relaxation function:

$$G(t, E_{11}) = A(E_{11})t^{B(E_{11})}, \tag{B.53}$$

where $A(E_{11})$ represents the initial elastic modulus of the material as a function of strain and

$$B(E_{11}) = g(E_{11})n_0, \tag{B.54}$$

with n_0 some initial relaxation rate and $g(E_{11})$ is a function that accounts for the strain-dependent nonlinearity in the relaxation rate. The functions A and g can be determined experimentally. Typically g is approximated by a polynomial. To generalize this model to 3-d the scalar functions A and B (and, hence g and G) must be replaced by the corresponding tensor-valued functions of strain.

Provenzano *et al.* [67] compared the predictions of this model to both experimental results and the predictions of the single integral model described in the previous section. For appropriate choices of material parameter values and functions both models gave accurate predictions for the stress-relaxation of a ligament for a range of stresses. The main difference between the models was the number of parameters and unknown functions that must be determined before the model can be used — for the single integral model between two and four functions and at least 2 parameters must be determined, whereas the modified superposition model only requires the determination of two functions. This makes the single integral model much more flexible, especially in terms of response to changes in temperature, but the modified superposition model is much simpler to use and allows a more direct interpretation of the relationship between model parameters and physical behaviour (since the nonlinear behaviour is encapsulated in one term).

B.2.2.3 Pioletti and Rakotomanana's Combined Differential-Integral Model

In general the effects of the short- and long-term viscous responses have been treated separately, leading to (at least) two constitutive equations for each material. Ideally both

types of behaviour should be described by the same constitutive equation. Moreover, this should be valid for large and small deformations, for any rate of strain, and should be compatible with the laws of thermodynamics. In this section a constitutive equation of this type, proposed by Pioletti and Rakotomanana [64], is described.

Suppose it is assumed that the stress at any given point in the material is dependent both on the current strain and the strain at all previous times at that point (with reference to the undeformed coordinate system). Then, Truesdell and Noll [84] give a general constitutive equation for such a material:

$$T(E,t) = T^e[E(t)] + \mathfrak{F}_{s=0}^{\infty}\{G(t-s), E(t)\}, \tag{B.55}$$

where $\mathfrak{F}$ is a functional representing the history of $G(t-s) = E(t-s) - E(t)$ and $T^e[E(t)]$ is an 'equilibrium term' representing the instantaneous (elastic) change in the stress resulting from a deformation (this is equivalent to the elastic response function in QLV theory). Pioletti and Rakotomanana [64] split the second term of this constitutive equation into two parts according to the time scale so that

$$T(E,t) = T^e[E(t)] + \mathfrak{F}_{s=0}^{\delta}\{G(t-s), E(t)\} + \mathfrak{F}_{s=\delta}^{\infty}\{G(t-s), E(t)\}, \tag{B.56}$$

where $\delta \ll 1$. This is valid only if $\mathfrak{F}$ is linear, but this does not restrict the response of the material to be linear. The first term of this expression has already been discussed. The second term represents the almost instantaneous change in stress (i.e. short-term memory) and can be approximated by the second Piola-Kirchoff viscous stress tensor[6] $T^v[\dot{E}(t), E(t)]$ since δ is small. (see Truesdell and Noll [84]). Finally, the third term represents the effects due to the long-term memory of the tissue and is the only term dependent on the strain history. A mathematical interpretation of the fading memory principal (see, e.g., Truesdell and Noll [84]) leads to an integral form for the third term:

$$\mathfrak{F}_{s=\delta}^{\infty}\{G(t-s), E(t)\} = \int_{\delta}^{\infty} \Sigma[G(t-s), s, E]ds, \tag{B.57}$$

where Σ is a tensor-valued function of strain and time.

Pioletti and Rakotomanana [64] attempted to identify appropriate forms for each of the

[6]Recall that the stress tensor can be written as the sum of an isotropic part $-p\delta_{ij}$ and a deviatoric part d_{ij} due entirely to the motion of the fluid. The deviatoric part is sometimes referred to as the *viscous stress tensor.*

terms in equation (B.56). T^e and T^v can be determined using appropriate strain energy functions (W^e and W^v respectively). These are defined so that equation (2.70) can be used to determine T^e and T^v from W^e and W^v respectively. No acceptable strain energy function for a compressible material could be identified, but assuming incompressibility (which has already been seen to be appropriate for many soft tissues) the strain energy function of Veronda and Westmann [86] (discussed in Section A.3.2) was used, i.e.

$$W^e = \frac{\alpha}{2}\exp[\beta(I_1 - 3)] - \frac{\alpha\beta}{4}(I_2 - 3), \tag{B.58}$$

where α, β are material parameters and I_1, I_2 are the first and second strain invariants. A viscous potential has been proposed by Pioletti *et al.* [65]:

$$W^v = \frac{\eta'}{4}\,\mathrm{tr}(\dot{C})^2(I_1 - 3), \tag{B.59}$$

where η' is a material parameter. Finally, for the third term, the function Σ is simplified by separating the variables:

$$\Sigma\big(G(t-s), s, E(t)\big) = \kappa\big(G(t-s)E(t)\big)\phi(s). \tag{B.60}$$

The validity of this simplification has has been experimentally verified (see Pioletti and Rakotomanana [63]). Then, using a similar argument as in Section B.2.1.2 it can be seen that κ can be approximated by the appropriate elastic stress, i.e.

$$\kappa(G(t-s), E(t)) = T^e(E(t-s)). \tag{B.61}$$

Pioletti and Rakotomanana [64] define ϕ in terms of a function $M(s')$ according to the relationship

$$\int_\delta^{s'} \phi(s)ds = M(s') = \left(\sum_{k=1}^{3} a_k \exp\left(-\frac{s'}{\tau_k}\right)\right) \Big/ \left(\sum_{k=1}^{3} a_k\right), \tag{B.62}$$

where a_k, τ_k, $k = 1, 2, 3$ are material parameters. This is analogous to the definition of the reduced relaxation function in QLV theory (Section B.2.1.3) but experimental data has been used to show that the sum can be terminated after 3 terms.

Hence, the constitutive equation resulting from the combined differential-integral model is

$$T = T^e[E(t)] + T^v[\dot{E}(t), E(t)] + \int_\delta^\infty \Sigma[G(t-s), s, E(t)]ds, \tag{B.63}$$

where

$$T^e = -pC^{-1} + \alpha\beta\{2\exp[\beta(I_1 - 3)] - I_1\}\delta + \alpha\beta C, \tag{B.64}$$

$$T^v = \eta'(I_1 - 3)\dot{C}, \tag{B.65}$$

$$\int_\delta^\infty \Sigma[G(t-s), s, E(t)]ds = \int_\delta^\infty T^e[E(t-s)]\dot{M}(s)ds, \tag{B.66}$$

δ is the Kronecker delta tensor and C is the right Cauchy-Green deformation tensor. So, this method has a total of nine material parameters (α, β, η', a_k and τ_k, $k = 1, 2, 3$). Observe that this model assumes that the material is isotropic. This simplifies the system so that the strain energy functions can be written as functions of the strain (and rate-of-strain in the case of W^v) invariants. This assumption is not usually appropriate in soft tissue.

B.2.3 Mixture Models

The use of mixture models in the modelling of biological soft tissue is made more difficult by the complex structure of the tissue. As discussed in Section 1.4.1 the tissue of interest will, in general, consist of various different types of cell embedded in an extracelluluar matrix which may itself be made up of various different materials. Constitutive equations must be formulated for all of the material types considered and for exchanges between them. For many of these materials it is difficult to devise experiments to validate the constitutive relations. Also, the cells making up the various materials are themselves composites containing a solid cytoskeleton and large amounts of water. According to Humphrey [39] this complexity means that no complete and widely accepted mixture theory for soft tissue has yet been developed. In this section some of the mixture models that have been developed are discussed.

B.2.3.1 Linear KLM Theory

One of the earliest attempts to model soft tissue using mixture theory was the development of a *linear biphasic theory* known as *linear KLM theory* by Mow *et al.* [51] for articular cartilage. In KLM theory the soft tissue is modelled as a binary mixture of solid and fluid: as described in Section B.1.2.2, each spatial point in the domain of interest is simultaneously occupied by both solid and fluid throughout the motion with solidity (solid fraction) $\phi^{(\mathrm{s})}$ and porosity (fluid fraction) $\phi^{(\mathrm{f})}$ where

$$\phi^{(\mathrm{s})} + \phi^{(\mathrm{f})} = 1. \tag{B.67}$$

At each point the true densities $\rho_T^{(\mathrm{s})}$ of the solid and $\rho_T^{(\mathrm{f})}$ of the fluid are defined to be the respective masses per unit volume of solid and fluid respectively whilst the apparent densities $\rho^{(\mathrm{s})}$ and $\rho^{(\mathrm{f})}$ are defined to be the respective masses per unit volume of the composite.[7] Hence, the density of the mixture is given by

$$\rho = \rho^{(\mathrm{s})} + \rho^{(\mathrm{f})}. \tag{B.68}$$

Both solid and fluid are assumed to be incompressible and the solid to be linearly elastic. Initially the fluid was assumed to be inviscid although in later work this assumption was relaxed to account for a Newtonian viscous fluid — this is what will be assumed here.

On introducing the Eulerian coordinate system $\mathbf{x}$ and Lagrangian coordinate systems $\mathbf{X}^{(\mathrm{s})}$ and $\mathbf{X}^{(\mathrm{f})}$ for the solid and fluid respectively the governing equations for the two constituents can be seen (from equations (2.62) and (2.48) respectively) to be

$$\rho^{(\mathrm{s})} \frac{D^{(\mathrm{s})} v_i^{(\mathrm{s})}}{Dt} = \frac{\partial \sigma_{ij}^{(\mathrm{s})}}{\partial x_j} + \rho^{(\mathrm{s})} f_i^{(\mathrm{s})} + \pi_i^{(\mathrm{s})}, \quad i = 1, 2, 3, \tag{B.69}$$

$$\rho^{(\mathrm{f})} \frac{D^{(\mathrm{f})} v_i^{(\mathrm{f})}}{Dt} = \frac{\partial \sigma_{ij}^{(\mathrm{f})}}{\partial x_j} + \rho^{(\mathrm{f})} f_i^{(\mathrm{f})} + \pi_i^{(\mathrm{f})}, \quad i = 1, 2, 3, \tag{B.70}$$

where $D^{(\mathrm{s})}/Dt$, $D^{(\mathrm{f})}/Dt$ are time derivatives holding $\mathbf{X}^{(\mathrm{s})}$ and $\mathbf{X}^{(\mathrm{f})}$ constant respectively, $\mathbf{v}^{(\alpha)}$, $\alpha = \mathrm{s}, \mathrm{f}$, are the phase velocities and $\mathbf{f}^{(\alpha)}$ and $\boldsymbol{\pi}^{(\alpha)}$, $\alpha = \mathrm{s}, \mathrm{f}$, the external and interaction forces respectively per unit mass acting on the two phases. Since there are only two phases

[7] Note that in biphasic theory quantities with the superscript (s) refer to properties of the solid phase, those with superscript (f) to properties of the fluid phase. The overall properties of the composite will not have any superscript at all.

clearly the interaction forces must satisfy

$$\boldsymbol{\pi}^{(\mathrm{s})} = -\boldsymbol{\pi}^{(\mathrm{f})}. \tag{B.71}$$

Note that since the porosity is not necessarily constant the incompressibility condition (for *both* phases) is

$$\nabla.\left(\phi^{(\mathrm{s})}\mathbf{v}^{(\mathrm{s})} + \phi^{(\mathrm{f})}\mathbf{v}^{(\mathrm{f})}\right). \tag{B.72}$$

Assuming that the body forces are known, all that is required to complete the model is to give constituent equations for the two phases and for the interaction force between them.

For the fluid phase the constitutive equation is derived from the general constitutive equation (2.45) for an incompressible Newtonian viscous fluid. The only difference is that the pressure p must be scaled by the porosity to account for the fact that it is now acting (equally) on both solid and fluid phases (the fraction of the pressure that is not acting on the fluid phase acts on the solid phase — see below), i.e.

$$\sigma_{ij}^{(\mathrm{f})} = -\phi^{(\mathrm{f})}p\delta_{ij} + 2\mu_f V_{ij} - \tfrac{2}{3}\mu_f V_{kk}\delta_{ij} \tag{B.73}$$

where V is the strain rate tensor, μ_f the fluid viscosity and δ the Dirac delta function. For the solid phase the constitutive equation is identical to that for a general linear elastic solid (see equation (2.52)), except that the component of the fluid pressure acting on the solid must be included. Hence,

$$\sigma_{ij}^{(\mathrm{s})} = -\phi^{(\mathrm{s})}p\delta_{ij} + \lambda_s\epsilon_{kk}\delta_{ij} + 2\mu_s\epsilon_{ij}, \tag{B.74}$$

where λ_s, μ_s are the Lamé coefficients of the solid and ϵ is Cauchy's infinitesimal strain tensor. Finally, the interaction force is assumed to be a function of the viscous drag force due to the relative velocity between the two phases and of the pressure in the changing fluid volume. In the linear theory of Mow *et al.* [51] it is assumed to have the form

$$-\boldsymbol{\pi}^{(\mathrm{f})} = \boldsymbol{\pi}^{(\mathrm{s})} = p\nabla\phi^{(\mathrm{f})} + K\left(\mathbf{v}^{(\mathrm{f})} - \mathbf{v}^{(\mathrm{s})}\right), \tag{B.75}$$

where K is the *diffusive drag coefficient.*

Mow *et al.* [51] used the KLM model to simulate creep and stress relaxation in articular cartilage. These viscoelastic effects are a consequence of the movement of the fluid within

the porous medium and the consequent frictional interaction between the two phases. According to Lai *et al.* [42] the viscous drag is the main cause of viscoelastic behaviour of articular cartilage in compression. Since this is directly accounted for in the biphasic KLM theory but not in any of the viscoelastic models discussed in Sections B.2.1 and B.2.2 it can be seen that using mixture models makes the intuitive understanding of the mechanisms leading to viscoelastic effects much simpler. This is a major advantage for any model that is to be used in a practical application since it makes it easier for the model to be altered to take into account changes in the behaviour of the material. However, as has been the case with all of the linear models discussed here, the linear KLM model is only valid for small deformations. In order to deal with large strains an alternative model is required.

B.2.3.2 Extensions of the Biphasic Theory

In order to overcome the deficiencies of KLM theory, in particular with regard to large deformations, a nonlinear biphasic theory has been developed. A full discussion of the developments made before 1993 can be found in Lai *et al.* [42]. One of the most fundamental advances in the theory was its extension to finite elasticity by Holmes [38] and Mow *et al.* [52]. In this model it is assumed that the soft tissue is a binary mixture of a Newtonian viscous fluid and an elastic solid with constitutive equation

$$T_{MN} = T^{(e)}_{MN}(E_{MN}), \tag{B.76}$$

for a homogeneous region in the absence of any fluid. Since the fluid is modelled in exactly the same way as for the KLM model its governing and constitutive equations are given by equations (B.70) and (B.73).[8] The governing and constitutive equations of the solid can be derived from those of the homogeneous material in a similar way to in the previous section. On doing this it can be seen that

$$\rho^{(\mathrm{s})}\frac{D^{(\mathrm{s})}v_i^{(\mathrm{s})}}{Dt} = \frac{\partial}{\partial X_M}\left(T^{(\mathrm{s})}_{MN}F_{iN}\right) + \rho^{(\mathrm{s})}f_i^{(\mathrm{s})} + \pi_i^{(\mathrm{s})}, \quad i = 1,2,3, \tag{B.77}$$

$$T^{(\mathrm{s})}_{MN} = -\phi^{(\mathrm{s})}p\delta_{MN} + T^{(e)}_{MN}, \tag{B.78}$$

[8]Recall that for a Newtonian viscous fluid, strain does not persist over time so all of the strain tensors discussed here are identical, as are the Cauchy and second Piola-Kirchoff stress tensors. Hence, it is consistent to define the motion of the fluid using the Cauchy stress but that of the solid using the second Piola-Kirchoff stress.

where F is the deformation gradient tensor. The constitutive equation of the interaction force must also be modified to take into account the nonlinearity in the viscous drag force (observe that this does not affect the linearity of the contribution due to the changing fluid volume). Holmes [38] models the interaction force as

$$-\boldsymbol{\pi}^{(\mathrm{f})} = \boldsymbol{\pi}^{(\mathrm{s})} = p\nabla\phi^{(\mathrm{f})} + K(I_3)\big(\mathbf{v}^{(\mathrm{f})} - \mathbf{v}^{(\mathrm{s})}\big), \tag{B.79}$$

where I_3 is the third principal invariant of strain (i.e. $I_3 = \det C$, where C is the right Cauchy-Green deformation tensor). As was the case in Sections A.2 and A.3 the accuracy of the model is dependent on the determination of an accurate constitutive equation for the elastic solid in the absence of fluid. Any of the models proposed in the earlier sections sections can be used and several have been. However, according to Humphrey [39] there is no universally accepted finite elastic biphasic model.

B.2.3.3 Peskin's Fibre-Reinforced Fluid Model

An alternative way of modelling soft tissue as a mixture is to consider it to be a fibre-reinforced fluid. This model is used by Peskin and McQueen [62] to model the motion of the heart wall. The model is based on the observation that the heart wall is made up of incompressible muscle fibres which are immersed in a fluid in which they are neutrally buoyant and arranged in a non-random way. According to Peskin and McQueen [62], at any point in the heart wall there is a well-defined fibre direction which varies smoothly from point to point.

Based on this observation, the heart wall is modelled as an incompressible Newtonian viscous fluid containing an immersed system of elastic fibres. The fibres are considered to be tenuous elements cutting through the fluid without displacing any volume or adding any mass. The only contribution they make to the system is to transmit a stress acting in the direction of the fibres at any given point. Hence the composite material is incompressible, has the same density ρ and viscosity μ as the fluid and its constitutive relation can be calculated by adding this additional stress to the usual constitutive relation for an incompressible Newtonian viscous fluid (see equation (2.45)), i.e.

$$\sigma_{ij} = -p\delta_{ij} + 2\mu V_{ij} + T'\tau_i\tau_j, \tag{B.80}$$

where p is the fluid pressure, T' is the fibre tension (force per unit cross-sectional area of

composite since the fibres themselves have no cross-sectional area) and $\boldsymbol{\tau} = (\tau_1, \tau_2, \tau_3)$ is the unit tangent to the fibres. From this it can be seen that the governing equations for the fluid flow are the Navier-Stokes equations for an incompressible fluid with density ρ and viscosity μ, i.e.

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v}.\nabla\mathbf{v}\right) = -\nabla p + \mu\nabla^2\mathbf{v} + \mathbf{f}, \tag{B.81}$$

$$\nabla.\mathbf{v} = 0, \tag{B.82}$$

where $\mathbf{v}$ is the velocity field of the fluid and $\mathbf{f}$ the fibre force density acting on the fluid (i.e. force per unit volume applied by the fibres to the fluid). These equations are to be solved in Cartesian coordinates $\mathbf{x} = (x_1, x_2, x_3)$.

To determine the fibre tension from the fibre strain a Lagrangian formulation is required. To do this define curvilinear coordinates (q, r, s, t) fixed in the fibres such that:

- fixed (q, r, s) indicates a material point;
- fixed (q, r) indicates a fibre;
- s measures arc length along a fibre in the reference (not necessarily unstressed) configuration.

The motion and spatial configuration of the fibres is described by the function $\mathbf{X}(q, r, s, t)$ having the following properties:

- if (q, r, s) are held constant then the equation $\mathbf{x}(t) = \mathbf{X}(q, r, s, t)$ defines the trajectory (in Cartesian space) of the material point with coordinates (q, r, s);
- if (q, r) and t are held constant then the equation $\mathbf{x}(s) = \mathbf{X}(q, r, s, t)$ defines the curve (in Cartesian space) occupied by the fibre which has coordinates (q, r) at time t.

Next, define $T(q, r, s, t)dqdr$ to be the force transmitted by the bundle of fibres $dqdr$ (note that this is *different* from T', the force per unit cross-sectional area). Peskin and McQueen [62] assume that since the fibres are elastic, T can be modelled by a generalized Hooke's law:

$$T = \sigma\left(\left|\frac{\partial \mathbf{X}}{\partial s}\right|, q, r, s, t\right), \tag{B.83}$$

for some function σ. The first argument of the function σ measures the ratio of the distances between two nearby points on a given fibre in the current and reference configurations —

it can be considered as an approximate measure of the fibre strain.[9]

In order to calculate the force exerted on the fluid by the fibres, recall that the fibre stress was assumed to act in the fibre direction at all times. This direction is defined by the unit tangent

$$\boldsymbol{\tau} = \frac{\partial \mathbf{X}/\partial s}{|\partial \mathbf{X}/\partial s|}. \tag{B.84}$$

Consider the bundle of fibres corresponding to some region Ω of the (q, r)-plane. The fibre force on the segment $s \in [a, b]$ due from outside the segment is given by

$$\int_\Omega [(T\boldsymbol{\tau})(q, r, b, t) - (T\boldsymbol{\tau})(q, r, a, t)]dqdr = \int_\Omega \int_a^b \frac{\partial}{\partial s}(T\boldsymbol{\tau})dsdqdr. \tag{B.85}$$

Since the fibres have no mass all of this force must be transferred to the fluid occupying the region $\Omega \times [a, b]$. Since this region is arbitrary the local force density applied to the fluid by the fibres is given by

$$\mathbf{F}(q, r, s, t) = \frac{\partial}{\partial s}(T\boldsymbol{\tau}). \tag{B.86}$$

Note that this force density is defined in terms of the Lagrangian fibre coordinates whereas the fluid equations were written in terms of Cartesian coordinates and the Cartesian force density $\mathbf{F}$. Since these two forces represent the same physical quantity the following relationship must hold:

$$\int_{R(t)} \mathbf{f}(\mathbf{x}, t)d\mathbf{x} = \int_{R_0} \mathbf{F}(q, r, s, t)dqdrds, \tag{B.87}$$

where R_0 is an arbitrary region in (q, r, s)-space (the reference configuration) and $R(t)$ its equivalent in the Cartesian space at time t. The standard way to relate $\mathbf{f}$ and $\mathbf{F}$ is through the Jacobian determinant

$$J(q, r, s) = \left| \frac{\partial \mathbf{X}}{\partial(q, r, s)} \right|, \tag{B.88}$$

which is independent of t since the composite is incompressible. Then,

$$J\mathbf{f}(\mathbf{X}(q, r, s, t), t) = \mathbf{F}(q, r, s, t). \tag{B.89}$$

[9]Note that if the material is unstressed in its reference configuration then an appropriate definition of the fibre strain would be $E = |\partial \mathbf{X}/\partial s| - 1$. So, the quantity $|\partial \mathbf{X}/\partial s|$ can be more accurately considered as an approximate Cauchy-Green deformation tensor.

Alternatively, $\mathbf{f}$ can be defined in terms of the 3-dimensional Dirac delta function as

$$\mathbf{f}(\mathbf{x},t) = \int \mathbf{F}(q,r,s,t)\delta(\mathbf{x} - \mathbf{X}(q,r,s,t))dqdrds, \tag{B.90}$$

where the integral is over the whole system of fibres.

Finally, it is assumed that there is a no-slip condition between the fibres and the fluid so that the fibres move with the same velocity of the fluid, i.e.

$$\frac{\partial \mathbf{X}}{\partial t}(q,r,s,t) = \mathbf{u}(\mathbf{X}(q,r,s,t),t) = \int \mathbf{u}(\mathbf{x},t)\delta(\mathbf{x} - \mathbf{X}(q,r,s,t))d\mathbf{x}. \tag{B.91}$$

This is the equation of motion of the fibres and *not* a constraint on the fluid motion. The seven equations (B.81)–(B.84), (B.86), (B.90) and (B.91) are the governing equation of the system. A numerical method for solving the fibre-reinforced fluid model of the heart — known as the *immersed boundary method* — is described in McQueen and Peskin [49]. The key difference between this model and the linear KLM model of Section B.2.3.1 is that here the solid phase only transmits a force in one direction. Observe that this formulation of the fibre-reinforced fluid model is linear, hence it is unlikely to be valid for large displacements.

Appendix C

Methods for Solving the Equations of Plane Strain Linear Elasticity

In this appendix the methods used in Chapter 5 to determine exact solutions of the 2-d plane strain equations of linear elasticity are described. Two separate techniques are described. First, in Section C.1 the governing equations are reformulated in the form of the biharmonic equation and solutions are determined in terms of harmonic functions (functions that satisfy Laplace's equation). This is the method used to calculate Goodier's solution in Section 5.1.2. Secondly, in Section C.2 the displacements are written in complex variable form (this can be done since the plane strain problem is 2-dimensional) and techniques of complex analysis are used to reformulate the governing equation as an ODE for the complex displacement in terms of a complex position variable. This method is used to demonstrate the presence of stress singularities in the solutions of the forward problem in Section 5.3.

C.1 Solution by Plane Harmonic Functions

In this section two special sets of solutions of the equations of linear elasticity are introduced. In order to do this it is shown in Section C.1.1 that the the stress components and displacements can be written in terms of a potential function that satisfies the biharmonic equation. Then, in Section C.1.2 *harmonic functions* are defined as solutions of Laplace's equation and a 2-d subclass *plane harmonic functions of degree n* is also defined. Two

methods of constructing solutions of Navier's equation in terms of these functions are described in Section C.1.3. Finally, in Section C.1.4 these special solutions are applied to the cylindrical inclusion to derive *Goodier's solution.*

C.1.1 The Biharmonic Form

Recall from Section 2.1.3.1 that the governing equations for a linear elastic solid that is not subject to any external forces are, in terms of the stresses,

$$\frac{\partial \sigma_{ij}}{\partial x_j} = 0, \quad i = 1, 2, 3, \tag{C.1}$$

or, on terms of the displacements

$$\mu \nabla^2 \mathbf{u} + (\lambda + \mu) \nabla (\nabla . \mathbf{u}) = \mathbf{0}. \tag{C.2}$$

For plane strain (i.e. $\mathbf{u} = (u(x, y), v(x, y)^T)$ the first two components of equation (C.1) become

$$\frac{\partial \sigma_{11}}{\partial x} + \frac{\partial \sigma_{12}}{\partial y} = 0, \tag{C.3}$$

$$\frac{\partial \sigma_{12}}{\partial x} + \frac{\partial \sigma_{22}}{\partial y} = 0. \tag{C.4}$$

Hence, there exists a function $\phi(x, y)$ (unique to within terms of the form $\alpha x + \beta y + \gamma$) such that

$$\sigma_{11} = \frac{\partial^2 \phi}{\partial x^2}, \quad \sigma_{12} = -\frac{\partial^2 \phi}{\partial x \partial y}, \quad \sigma_{22} = \frac{\partial^2 \phi}{\partial y^2}. \tag{C.5}$$

Recall also that

$$\sigma_{11} = (\lambda + 2\mu) \frac{\partial u}{\partial x} + \lambda \frac{\partial v}{\partial y}, \tag{C.6}$$

$$\sigma_{22} = \lambda \frac{\partial u}{\partial x} + (\lambda + 2\mu) \frac{\partial v}{\partial y}. \tag{C.7}$$

Hence,

$$\nabla^2 \phi = \sigma_{11} + \sigma_{22} = 2(\lambda + \mu) \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) = \nabla . \mathbf{u}. \tag{C.8}$$

Now, by observing that $\nabla^2\mathbf{u} = \nabla(\nabla.\mathbf{u}) + \nabla \times (\nabla \times \mathbf{u})$, where $\nabla \times \mathbf{u}$ represents the curl of $\mathbf{u}$ and recalling that the divergence of a curl is zero (both of these vector identities can be found in, e.g., Bourne and Kendall [13]), it can be seen that taking the divergence of equation (C.2) gives

$$\nabla^2(\nabla.\mathbf{u}) = 0, \tag{C.9}$$

$$\Rightarrow \quad \nabla^4\phi = 0. \tag{C.10}$$

This equation is known as the *biharmonic equation.*

C.1.2 Harmonic Functions

Clearly any function that is a solution of Laplace's equation ($\nabla^2\phi = 0$) will also be a solution of the biharmonic equation. Such solutions are known as *harmonic functions.* One way of identifying such functions is by separation of the variables in Laplace's equation, i.e. by writing

$$\phi(x, y) = R(r)\Theta(\theta), 6 \tag{C.11}$$

where (r, θ) are the polar coordinates in the plane, i.e.

$$x = r\cos\theta, \quad y = r\sin\theta. \tag{C.12}$$

Note that although not all solutions of Laplace's equation are separable they can all be expressed as a (possibly infinite) sum of separable solutions. The Laplacian operator is written in plane polar coordinates as

$$\nabla^2 = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}, \tag{C.13}$$

and so Laplace's equation becomes

$$\frac{r(rR')'}{R} = -\frac{\Theta''}{\Theta}, \tag{C.14}$$

Then, since the left-hand-side is independent of θ and the right-hand-side independent of θ they must both be independent of r and θ, i.e. they are constants. So, for $\Theta(\theta)$,

$$\Theta'' = k\Theta, \tag{C.15}$$

Suppose further that we seek periodic solutions for Θ (so that the point represented by (r, θ) is the same as that represented by $(r, n\theta)$ for any integer n). Then this equation only has a solution if $k = n^2$, where $n \in \mathbb{Z}$ (note, however, that non-periodic solutions to Laplace's equation do exist). If $n = 0$ then

$$\Theta(\theta) = \Theta_0(\theta) = A_0, \tag{C.16}$$

where A_0 is a constant (note that $\Theta = A + B\theta$ is not possible as this is not periodic, even though it appears to satisfy Laplace's equation). If $n \neq 0$ then

$$\Theta(\theta) = \Theta_n = A_n \sin n\theta + B_n \cos n\theta, \quad n \in \mathbb{N}. \tag{C.17}$$

For $R(r)$, if $n = 0$

$$(rR')' = 0, \tag{C.18}$$

$$\Rightarrow \quad R(r) = R_0(r) = C_0 \log r + D_0. \tag{C.19}$$

And, if $n \neq 0$,

$$r(rR')' = n^2, \tag{C.20}$$

$$\Rightarrow \quad R(r) = R_n(r) = C_n r^n + D_n r^{-n}, \quad n \in \mathbb{N}. \tag{C.21}$$

Any periodic solution of Laplace's equation can be obtained by superposition of these forms, i.e. any solution can be written as a Fourier series in terms of the separable solutions:

$$\phi(r, \theta) = A_0 \log r + B_0 + \sum_{n=1}^{\infty} \left(r^n (A_n \sin n\theta + B_n \cos n\theta) + r^{-n} (A_{-n} \sin n\theta + B_{-n} \cos n\theta) \right), \tag{C.22}$$

$$= A_0 \log r + B_0 + \sum_{n=1}^{\infty} A_n \psi_n \sin n\theta + B_n \omega_n + A_{-n} \psi_{-n} + B_{-n} \omega_{-n}. \tag{C.23}$$

But, in particular for this work, the functions

$$\psi_n = r^n \sin n\theta, \quad \omega_n = r^n \cos n\theta, \quad n \in \mathbb{N} \tag{C.24}$$

are *plane harmonic functions of degree* n, with $\log r$ being a plane harmonic function of degree zero.

C.1.3 Solution of Navier's Equation

Clearly any harmonic function ϕ (solution of $\nabla^2\phi$) will also be a solution of the biharmonic equation ($\nabla^4\phi = 0$). So, one type of solution of the Navier equation (C.2) for plane strain is $\mathbf{u} = (u_r, u_\theta)$, where u_r and u_θ are the displacements in the radial and transverse directions respectively and are given by

$$u_r = \frac{\partial\phi}{\partial r}, \quad u_\theta = \frac{1}{r}\frac{\partial\phi}{\partial\theta}, \tag{C.25}$$

where ϕ is any plane harmonic function. These are called *solutions of type* ϕ.

Suppose now that ω_n is a plane harmonic function of degree n. Then consider a displacement field of the form

$$\mathbf{u} = r^2\nabla\omega_n + \alpha_n \mathbf{r}\omega_n, \tag{C.26}$$

where $\mathbf{r} = (x, y)^T$ is the position vector,

$$\alpha_n = -2n\left(\frac{3-4\nu}{n+4-4\nu}\right), \tag{C.27}$$

and, as usual, ν is the Poisson's ratio of the material. It can be shown that displacement fields of this form are solutions of the Navier equation (C.2) by direct substitution and considering the individual components separately.

The full details of verifying these solutions is omitted here, except to observe that it is

necessary to derive some properties of the plane harmonic functions of degree n. First,

$$\nabla^2(x\omega_n) = \omega_n\nabla^2 x + 2\nabla x.\nabla\omega_n + x\nabla^2\omega_2, \quad \text{by applying the chain rule,} \tag{C.28}$$

$$= 2\nabla x.\nabla\omega_n, \quad \text{since } \nabla^2 x = \nabla^2\omega_n = 0, \tag{C.29}$$

$$= 2(1,0).\left(\frac{\partial\omega_n}{\partial x}, \frac{\partial\omega_n}{\partial y}\right), \tag{C.30}$$

$$= 2\frac{\partial\omega_n}{\partial x}. \tag{C.31}$$

A similar result holds for $\nabla^2(y\omega_n)$. Next,

$$\nabla^2\left(r^m\omega_n\right) = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\left(r^{m+n}\cos n\theta\right)\right) + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\left(r^{m+n}\cos n\theta\right), \tag{C.32}$$

$$= \frac{m+n}{r}\frac{\partial}{\partial r}\left(r^{m+n}\cos n\theta\right) - \frac{n}{r^2}\frac{\partial}{\partial\theta}\left(r^{m+n}\sin n\theta\right), \tag{C.33}$$

$$= (m+n)^2 r^{m+n-2}\cos n\theta - n^2 r^{m+n-2}\cos n\theta, \tag{C.34}$$

$$= m(m+2n)r^{m-2}\omega_n. \tag{C.35}$$

In each of these cases the result is derived for the ω_n case in equation (C.24) but can also be shown to hold for ψ_n.

So, a second type of solution of the Navier equation (C.2) for plane strain is $\mathbf{u} = (u_r, u_\theta)$, where u_r and u_θ are the displacements in the radial and transverse directions respectively and are given by

$$u_r = r^2\frac{\partial\omega_n}{\partial r} + \alpha_n r\omega_n, \quad u_\theta = r\frac{\partial\omega_n}{\partial\theta}, \tag{C.36}$$

where ω_n is a plane harmonic function of degree n. These are called *solutions of type* ω.

C.1.4 Application to the Cylindrical Inclusion Problem

Recall from Section 5.1.2 that an infinite material undergoing plane strain with radially symmetric Lamé constants of the form

$$\lambda(r) = \begin{cases} \lambda_1 & \text{if } r \geqslant a, \\ \lambda_2 & \text{if } r < a, \end{cases} \tag{C.37}$$

$$\mu(r) = \begin{cases} \mu_1 & \text{if } r \geqslant a, \\ \mu_2 & \text{if } r < a, \end{cases} \tag{C.38}$$

for some $a > 0$, is said to contain a *cylindrical inclusion.* Goodier [32] proposed a solution to the equations for a material containing a cylindrical inclusion undergoing uniaxial tension by splitting the domain into two regions ($r \geqslant a$ and $r < a$) in which the Lamé constants are constant, determining solutions in terms of plane harmonic functions in each and then applying matching boundary conditions at $r = a$. This solutions method is described in this section.

In the region $r > a$ the strategy is to take a plane harmonic solution that gives rise to displacements that tend to zero at infinity and add it to the solution for uniaxial tension given in the previous section. This will give a solution that satisfies the boundary conditions as $x \to \infty$ and $y \to \infty$ and contains arbitrary constants that can be used to match with the solution in $r < a$. With this in mind ϕ and ω in this region are chosen to be

$$\phi(r,\theta) = A\omega_0 + \tfrac{1}{2}B\omega_{-2} = A\log r + \frac{B}{2r^2}\cos 2\theta, \tag{C.39}$$

$$\omega(r,\theta) = \tfrac{1}{2}(1-2\nu_1)C\omega_{-2} = \frac{C(1-2\nu_1)}{2r^2}\cos 2\theta, \tag{C.40}$$

so that the displacements and stresses due to the plane harmonic functions are given by

$$u_r = \frac{A}{r} + \left[-\frac{B}{r^3} + \frac{2C}{r}(1-\nu_1) \right] \cos 2\theta, \tag{C.41}$$

$$u_\theta = -\left[\frac{B}{r^3} + \frac{C}{r}(1-2\nu_1) \right] \sin 2\theta, \tag{C.42}$$

$$\sigma_{rr} = 2\mu_1 \left[-\frac{A}{r^2} + \left(\frac{3B}{r^4} - \frac{2C}{r^2} \right) \cos 2\theta \right], \tag{C.43}$$

$$\sigma_{\theta\theta} = 2\mu_1 \left[\frac{A}{r^2} - \frac{3B}{r^4} \cos 2\theta \right], \tag{C.44}$$

$$\sigma_{r\theta} = 2\mu_1 \left[\frac{3B}{r^4} - \frac{C}{r^2} \right] \sin 2\theta, \tag{C.45}$$

where A, B, C are arbitrary constants and ν_1 is the Poisson's ratio of the medium. The displacements given in equation (5.8) for simple uniaxial tension can be rewritten in (r, θ)-coordinates as

$$u_r = \frac{Tr}{4\mu_1}(1 - 2\nu_1 + \cos 2\theta), \tag{C.46}$$

$$u_\theta = -\frac{Tr}{4\mu_1} \sin 2\theta, \tag{C.47}$$

with corresponding stress components

$$\sigma_{rr} = \frac{T}{2}(1 + \cos 2\theta), \tag{C.48}$$

$$\sigma_{r\theta} = -\frac{T}{2} \sin 2\theta, \tag{C.49}$$

where the expression for the stress component $\sigma_{\theta\theta}$ is not given as it is not required to determine the traction on a circle $r =$ constant (see below). So, the overall displacements

and stress components in the region $r > a$ are given by

$$u_r = \frac{A}{r} + \left[-\frac{B}{r^3} + \frac{2C}{r}(1-\nu_1) \right] \cos 2\theta + \frac{Tr}{4\mu_1}(1 - 2\nu_1 + \cos 2\theta), \tag{C.50}$$

$$u_\theta = -\left[\frac{B}{r^3} + \frac{C}{r}(1-2\nu_1) \right] \sin 2\theta - \frac{Tr}{4\mu_1} \sin 2\theta, \tag{C.51}$$

$$\sigma_{rr} = 2\mu_1 \left[-\frac{A}{r^2} + \left(\frac{3B}{r^4} - \frac{2C}{r^2} \right) \cos 2\theta \right] + \frac{T}{2}(1 + \cos 2\theta), \tag{C.52}$$

$$\sigma_{r\theta} = 2\mu_1 \left[\frac{3B}{r^4} - \frac{C}{r^2} \right] \sin 2\theta - \frac{T}{2} \sin 2\theta. \tag{C.53}$$

In the region $r < a$ a solution which is continuous at $r = 0$ is sought. Hence the ϕ- and ω-type solutions are chosen to be

$$\phi(r,\theta) = \tfrac{1}{2} G\omega_2 = \tfrac{1}{2} G r^2 \cos 2\theta, \tag{C.54}$$

$$\omega(r,\theta) = F\omega_0 + \tfrac{1}{2}(3 - 2\nu_2) H\omega_2 = F \log r + \tfrac{1}{2}(3 - 2\nu_2) H r^2 \cos 2\theta, \tag{C.55}$$

so that

$$u_r = Fr + \left[Gr + 2\nu_2 H r^3 \right] \cos 2\theta, \tag{C.56}$$

$$u_\theta = -\left[Gr + (3 - 2\nu_2) H r^3 \right] \sin 2\theta, \tag{C.57}$$

$$\sigma_{rr} = 2\mu_2 \left[\frac{F}{1 - 2\nu_2} + G \cos 2\theta \right], \tag{C.58}$$

$$\sigma_{\theta\theta} = 2\mu_2 \left[\frac{F}{1 - 2\nu_2} - \left(G + 6Hr^2 \right) \cos 2\theta \right], \tag{C.59}$$

$$\sigma_{r\theta} = -2\mu \left[G + 3Hr^2 \right] \sin 2\theta, \tag{C.60}$$

where F, G, H are constants.

The constants are determined by observing that the displacements and stress components

σ_{rr} and $\sigma_{r\theta}$ must be continuous across the interface $r = a$.[1] It can then be shown that

$$A = \frac{Ta^2}{4\mu_1}\left(\frac{(1-2\nu_2)\mu_1 - (1-2\nu_1)\mu_2}{(1-2\nu_2)\mu_1 + \mu_2}\right), \tag{C.61}$$

$$B = \frac{Ta^4}{4\mu_1}\left(\frac{\mu_1 - \mu_2}{\mu_1 + (3-4\nu_1)\mu_2}\right), \tag{C.62}$$

$$C = \frac{Ta^2}{2\mu_1}\left(\frac{\mu_1 - \mu_2}{\mu_1 + (3-4\nu_1)\mu_2}\right), \tag{C.63}$$

$$F = \frac{T}{2}\left(\frac{(1-2\nu_2)(1-\nu_1)}{(1-2\nu_2)\mu_1 + \mu_2}\right), \tag{C.64}$$

$$G = T\left(\frac{(1-\nu_1)}{\mu_1 + (3-4\nu_1)\mu_2}\right), \tag{C.65}$$

$$H = 0. \tag{C.66}$$

Here, the values of A and F are determined from continuity of u_r and σ_{rr} at $\theta = \pi/4$ (so that the $\cos 2\theta$ terms vanish). This then leaves the four equations for general θ to determine B, C, G and H.

C.2 Complex Variable Methods

In this section the use of complex variable techniques to solve problems of linear elasticity is considered. In Section C.2.1 the components of stress and displacement are written in terms of a complex potential function. Then, in Section C.2.3, this formulation is used to analyze the solution of the linear elasticity equations for the *punch problem* on a half plane. In particular, a singularity in the resulting stresses is shown to exist at points where there is a switch from Dirichlet to Neumann boundary conditions.

C.2.1 The Complex Potential

Recall from equation (2.61) that the Navier equation for a material in which the Lamé constants do not depend on position and which is undergoing plane strain in equilibrium

[1]The stress component $\sigma_{\theta\theta}$ is not required to be continuous. This is because across any curve in the domain the displacements and the traction are required to be continuous. The traction across a curve $r =$ constant is given by $\sigma_{rr}\sin\theta - \sigma_{r\theta}\cos\theta$, which does not depend on $\sigma_{\theta\theta}$.

is

$$\mu\nabla^2\mathbf{u} + (\lambda+\mu)\nabla(\nabla.\mathbf{u}) = \mathbf{0}, \tag{C.67}$$

everywhere in the region R occupied by the material, where $\mathbf{u} = (u(x,y), v(x,y))$. Introduce the complex variables

$$z = x + \mathrm{i}y, \quad \bar{z} = x - \mathrm{i}y, \tag{C.68}$$

$$d = u + \mathrm{i}v, \quad \bar{d} = u - \mathrm{i}v, \tag{C.69}$$

where $\mathrm{i} = \sqrt{-1}$, d is a function of z and $\bar{z}$ and $\bar{d}$ are referred to as the *complex conjugates* of z and d respectively. Then,

$$\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)^T = J\left(\frac{\partial}{\partial z}, \frac{\partial}{\partial \bar{z}}\right)^T, \tag{C.70}$$

where J is the Jacobi matrix

$$J = \frac{\partial(z,\bar{z})}{\partial(x,y)} = \begin{pmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{pmatrix}. \tag{C.71}$$

Then, on inverting J it can be seen that

$$2\frac{\partial}{\partial z} = \frac{\partial}{\partial x} - \mathrm{i}\frac{\partial}{\partial y}, \tag{C.72}$$

$$2\frac{\partial}{\partial \bar{z}} = \frac{\partial}{\partial x} + \mathrm{i}\frac{\partial}{\partial y}. \tag{C.73}$$

Then Navier's equation can be rewritten in the form

$$2\mu\frac{\partial^2 d}{\partial z \partial \bar{z}} + (\lambda+\mu)\frac{\partial}{\partial \bar{z}}\left(\frac{\partial d}{\partial z} + \overline{\frac{\partial d}{\partial z}}\right) = 0. \tag{C.74}$$

In this expression, the real part represents the first component of the matrix equation (C.67) and the complex part the second component. This is now a differential equation for d as a complex function of z and can be analyzed using complex variable methods.

First, observe that equation (C.74) can be rewritten as

$$\frac{\partial H}{\partial \bar{z}} = 0, \tag{C.75}$$

where

$$H = 2\mu\frac{\partial d}{\partial z} + (\lambda+\mu)\left(\frac{\partial d}{\partial z} + \overline{\frac{\partial d}{\partial z}}\right) = \theta'(z), \tag{C.76}$$

where θ is a holomorphic function of z in Ω.[2] Then the real and imaginary parts of $\partial d/\partial z$ can be determined by considering equation (C.76) and its complex conjugate. It turns out that

$$4(\lambda+2\mu)\Re\left(\frac{\partial d}{\partial z}\right) = \theta'(z) + \overline{\theta'(z)}, \tag{C.77}$$

$$4\mu\mathrm{i}\Im\left(\frac{\partial d}{\partial z}\right) = \theta'(z) - \overline{\theta'(z)}, \tag{C.78}$$

where $\Re(z)$ and $\Im(z)$ denote the real and imaginary parts of z respectively, and

$$4\mu(\lambda+2\mu)\frac{\partial d}{\partial z} = (\lambda+3\mu)\theta'(z) - (\lambda+\mu)\overline{\theta'(z)}. \tag{C.79}$$

This equation can be integrated to give the displacement as

$$4\mu(\lambda+2\mu)d = (\lambda+3\mu)\theta(z) - (\lambda+\mu)z\overline{\theta'(z)} + \overline{\phi(z)}, \tag{C.80}$$

where $\phi(z)$ is an arbitrary holomorphic function of the form $\phi(z) = p(x,y) + iq(x,y)$. This is calculated by writing $\theta'(z) = f(x,y) + ig(x,y)$ and observing that

$$2\frac{\partial}{\partial z}\left(z\overline{\theta'(z)} + \overline{\phi(z)}\right) = \left(\frac{\partial}{\partial x} - \mathrm{i}\frac{\partial}{\partial y}\right)(xf(x,y) - \mathrm{i}xg(x,y) + \mathrm{i}yf(x,y) + yg(x,y) + p - iq), \tag{C.81}$$

$$\begin{aligned} &= f - \mathrm{i}g + x(f_x - \mathrm{i}g_x) + y(\mathrm{i}f_x + g_x) + p_x - q_y \\ &\quad - \mathrm{i}[x(f_y - \mathrm{i}g_y) + \mathrm{i}f + g + y(\mathrm{i}f_y + g_y) + p_y + q_x], \end{aligned} \tag{C.82}$$

$$= 2(f - ig) + (x + \mathrm{i}y)[f_x - g_y - \mathrm{i}(f_y - g_x)] + [p_x - q_y - \mathrm{i}(p_y + q_x)], \tag{C.83}$$

[2] A complex-valued function is said to be *holomorphic* in an open set Ω if it is differentiable at every point $z \in \Omega$. Further details on complex variable theory can be found in, e.g., Priestley [66].

where the subscripts denote partial differentiation. Now, the Cauchy-Riemann equations[3] can be applied to both $\theta'(z)$ and $\phi(z)$ to show that the two terms in square brackets on the right-hand-side are zero and, hence, that

$$\frac{\partial}{\partial z}\left(z\overline{\theta'(z)} + \overline{\phi(z)}\right) = \overline{\theta'(z)}, \tag{C.84}$$

which leads to equation (C.80).

Suppose now that

$$\theta(z) = \frac{4(\lambda + 2\mu)}{\lambda + \mu}\Omega(z), \tag{C.85}$$

$$\phi(z) = -4(\lambda + 2\mu)\omega(z). \tag{C.86}$$

Then the complex displacement becomes

$$\mu d = \kappa\Omega(z) - z\overline{\Omega'(z)} - \overline{\omega(z)}, \tag{C.87}$$

where

$$\kappa = \frac{\lambda + 3\mu}{\lambda + \mu}. \tag{C.88}$$

So, for plane strain, $\kappa = 3 - 4\nu$. The functions $\Omega(z)$ and $\omega(z)$ are referred to as the *complex potentials.*

Recall from Section 2.1.3.1 that for a linear elastic material the stresses are given in terms of the displacements by

$$\sigma_{ij} = \lambda\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + \mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \quad i, j = 1, 2, \tag{C.89}$$

---

[3] THEOREM (CAUCHY-RIEMANN EQUATIONS)
Suppose that $F(z) = f(x, y) + \mathrm{i}g(x, y)$ is differentiable at $z = x + \mathrm{i}y \in \Omega$. Then f and g have first-order partial derivatives at z and these satisfy the *Cauchy-Riemann equations*

$$f_x = g_y, \quad f_y = -g_x.$$

For a proof of this theorem see, e.g., Priestley [66].

where $(x_1, x_2) = (x, y)$ and $(u_1, u_2) = (u, v)$. Then, in terms of the complex displacement,

$$\sigma_{11} + \sigma_{22} = 2(\lambda + \mu)\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = 2(\lambda + \mu)\left(\frac{\partial d}{\partial z} + \overline{\frac{\partial d}{\partial z}}\right), \tag{C.90}$$

$$\sigma_{11} - \sigma_{22} + 2\mathrm{i}\sigma_{12} = 2\mu\left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} + \mathrm{i}\frac{\partial u}{\partial y} + \mathrm{i}\frac{\partial v}{\partial x}\right) = 4\mu\frac{\partial d}{\partial \bar{z}}. \tag{C.91}$$

So, the stresses can be written in terms of the complex potentials as

$$\Theta = \sigma_{11} + \sigma_{22} = 4\big(\Omega'(z) + \overline{\Omega'(z)}\big), \tag{C.92}$$

$$\Phi = \sigma_{11} - \sigma_{22} + 2\mathrm{i}\sigma_{12} = -4\big(z\overline{\Omega''(z)} + \overline{\omega'(z)}\big). \tag{C.93}$$

Hence, the complete stress and displacement fields have been expressed in terms of the complex potentials. A full theory of complex potentials can be found in, e.g., England [23].

C.2.2 Analytic Extension of the Complex Potentials

Suppose now that the Navier equation is to be solved on the half-plane $R^+ = \{(x, y) \in \mathbb{R}^2 : y > 0\}$. The stress and displacements in this region are determined entirely by the complex potentials in R^+ and their boundary conditions (applied on $y = 0$). So, the values of the complex potentials $\Omega(z)$, $\omega(z)$ in $R^- = \{(x, y) \in \mathbb{R}^2 : y < 0\}$ are arbitrary. This allows the solution to be simplified by defining the value of one of the potentials in R^- in terms of the other one in order to express the stresses and displacements in terms of a single complex potential.

Suppose that the definition of $\Omega(z)$, $z \in R^+$ is extended into R^- by setting

$$\Omega(z) = \begin{cases} \Omega(z), & \text{if } z \in R^+, \\ -z\overline{\Omega'(\bar{z})} - \overline{\omega(\bar{z})}, & \text{if } z \in R^-. \end{cases} \tag{C.94}$$

Then, in the half-plane of interest R^+, $\omega(z)$ can be written in terms of $\Omega(z)$ as

$$\omega(z) = \overline{\Omega(\bar{z})} - z\Omega'(z). \tag{C.95}$$

Then, on substituting for $\omega(z)$ in equations (C.87), (C.92) and (C.93) the displacement

and stresses in R^+ are given by

$$\mu d = \kappa\Omega(z) - +\Omega(\bar{z}) + (z - \bar{z})\overline{\Omega'(z)}, \tag{C.96}$$

$$\sigma_{11} + \sigma_{22} = 4\big(\Omega'(z) + \overline{\Omega(z)}\big), \tag{C.97}$$

$$\sigma_{22} - \mathrm{i}\sigma_{12} = 2\big(\Omega'(z) - \Omega'(\bar{z}) + (z - \bar{z})\overline{\Omega''(z)}\big), \tag{C.98}$$

where $\Omega(z)$ is holomorphic in both R^+ and R^-.

C.2.3 The Punch Problem on the Half-Plane

Suppose that a linear elastic material undergoing plane strain occupies the semi-infinite half-plane $R^+ = \{(x, y) \in \mathbb{R}^2 : y > 0\}$. The displacements are assumed to tend to zero at infinity. The line $y = 0$ is assumed to be stress free except for the section $|x| \leqslant a$, where $u = 0$, $v = \delta$. In this section a singularity in the stress is shown to exist at the points $(x, y) = (\pm a, 0)$. In order to do this first observe that, in complex variable form, the boundary conditions are

$$d = \mathrm{i}\delta, \quad \text{on } y = 0 \text{ where } |x| \leqslant a, \tag{C.99}$$

$$\sigma_{22} - \mathrm{i}\sigma_{12} = 0, \quad \text{on } y = 0 \text{ where } |x| > a. \tag{C.100}$$

The strategy to solve the mixed boundary value problem is to split the line $y = 0$, $|x| \leqslant a$ into n line segments $L_k = (a_k, b_k)$, $k = 1, \ldots, n$, where $a_i < b_i = a_{i+1}$.

On doing this, England [23] shows that the stress in the region $\{(x, y) \in R^+ : |x| \leqslant a\}$ is given by

$$\sigma_{22} - \mathrm{i}\sigma_{12} = 2\big(\Omega'^{+}(x) - \Omega'^{-}(x)\big) = 2(1 + \kappa)\Omega'^{+}(x), \tag{C.101}$$

where $\Omega'^{+/-}(x) = \lim_{(y \to 0+/-)} \Omega'(z)$. Further, it can be shown that in order to ensure that the stress and rotation tend to zero as $z \to \infty$ the derivative of the complex potential must satisfy

$$\Omega'(z) = -\frac{P}{\pi} X(z), \tag{C.102}$$

where

$$X(z) = (z + a)^{-\frac{1}{2} - \mathrm{i}\beta}(z - a)^{-\frac{1}{2} + \mathrm{i}\beta}, \tag{C.103}$$

and

$$\beta = \frac{1}{2\pi} \log |\kappa|. \tag{C.104}$$

Suppose now that $(z+a)$, $(z-a)$ are written in exponential form as

$$z - a = r_1 \mathrm{e}^{\mathrm{i}\theta_1}, \quad z + a = r_2 \mathrm{e}^{\mathrm{i}\theta_2}. \tag{C.105}$$

Then, substituting in equation (C.101) the stress components are given on $y = 0$, $|x| \leqslant a$ by

$$\sigma_{22} - \mathrm{i}\sigma_{12} = \frac{\mathrm{i}P}{\pi\sqrt{\kappa}} \frac{(1+\kappa)}{(a^2 - x^2)^{\frac{1}{2}}} \exp\left(\mathrm{i}\beta \log \left|\frac{a-x}{a+x}\right|\right). \tag{C.106}$$

Then, on taking real and imaginary parts, the stress components show the following asymptotic behaviour as $z \to \pm a$:

$$\sigma_{12} = O\big(r^{-1/2} \sin(\beta \log r)\big), \tag{C.107}$$

$$\sigma_{22} = O\big(r^{-1/2} \cos(\beta \log r)\big), \tag{C.108}$$

where $r = r_1$ close to $z = a$, $r = r_2$ close to $z = -a$. So, there are singularities in two of the stress components at the points where there is a change in the boundary condition type (stress-free for $|z| > a$, displacement conditions for $|z| \leqslant a$). Close to these singularities the stress components show rapid oscillations and their magnitudes become very large. This explains the behaviour seen in the corners of the domain for the solutions of the parameter recovery method shown in Chapter 5 (recall that there was a change in boundary condition at the corners).

Bibliography

[1] D. J. Acheson. *Elementary Fluid Dynamics.* Oxford University Press, 1990.

[2] R. J. Atkin and R. E. Craine. Continuum theories of mixtures: Basic theory and historical development. *Q. J. Mech. Appl. Math.*, 29:209–244, 1976.

[3] F. S. Azar, D. N. Metaxas, and M. D. Schnall. A deformable finite element model of the breast for predicting mechanical deformations under external perturbations. *J. Acad. Radiol.*, 8(10):965–975, 2001.

[4] F. S. Azar, D. N. Metaxas, and M. D. Schnall. Methods for modeling and predicting mechanical deformations of the breast under external perturbations. *Med. Im. Anal.*, 6:1–27, 2002.

[5] P. E. Barbone and J. C. Bamber. Quantitative elasticity imaging: What can and cannot be inferred from strain images. *Phys. Med. Biol.*, 47:2147–2164, 2002.

[6] P. E. Barbone and N. H. Gokhale. Elastic modulus imaging: On the uniqueness and nonuniqueness of the elastography inverse problem in two dimensions. *Inverse Problems*, 20:283–296, 2004.

[7] J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *Int. J. Comput. Vision*, 12:43–77, 1994.

[8] G. K. Batchelor. *An Introduction to Fluid Dynamics.* Cambridge University Press, 1967.

[9] M. A. Biot. General theory of three-dimensional consolidation. *J. App. Phys.*, 12:155–165, 1941.

[10] M. A. Biot. Theory of elasticity and consolidation for a porous anisotropic solid. *J. App. Phys.*, 26:182–185, 1955.

[11] M. A. Biot. General solutions of the equations of elasticity and consolidation for a porous material. *ASME J. Appl. Mech.*, 78:91–96, 1956.

[12] L. Boltzmann. Zur theorie der elastichen Nachwirckung. *Ann. Phys. Suppl.*, 7:624–654, 1876.

[13] D. E. Bourne and P. C. Kendall. *Vector Analysis and Cartesian Tensors.* Chapman & Hall, London, 3rd edition, 1992.

[14] R. M. Bowen. Theory of mixtures. In A. C. Eringen, editor, *Continuum Physics*, pages 1–127. Academic Press, 1976.

[15] W. E. Boyce and R. C. DiPrima. *Elementary Differential Equations and Boundary Value Problems.* Wiley, New York, 7th edition, 2001.

[16] British Heart Foundation. *heartstats.* URL: www.heartstats.org, 2006.

[17] Cancer Research UK. *CancerStats.* URL: info.cancerresearchuk.org/cancerstats, 2006.

[18] I. Céspedes, J. Ophir, H. Ponnekanti, and N. F. Maklad. Elastography: Elasticity imaging using ultrasound with application to muscle and breast. *Ultrason. Imag.*, 15:73–88, 1993.

[19] R. Christensen. *Theory of Viscoelasticity.* Academic Press, London, 2nd edition, 1982.

[20] B. D. Coleman and W. Noll. Foundations of linear viscoelasticity. *Rev. Modern Phys.*, 3:239–249, 1961.

[21] H. Darcy. *Les Fontaines Publiques de la Ville de Dijon.* Dalmont, Paris, 1856.

[22] M. M. Doyley, P. M. Meaney, and J. C. Bamber. Evaluation of an iterative reconstruction method for quantitative elastography. *Phys. Med. Biol.*, 45:1521–1540, 2000.

[23] A. H. England. *Complex Variable Methods in Elasticity.* Wiley, London, 1971.

[24] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson. *Computational Differential Equations.* Cambridge University Press, 1996.

[25] J. J. Faran. Sound scattering by solid cylinders and spheres. *J. Acoust. Soc. Am.*, 23:405–418, 1951.

[26] D. Fu, S. F. Levinson, S. M. Gracewski, and K. J. Parker. Non-invasive quantitative reconstruction of tissue elasticity using and iterative forward approach. *Phys. Med. Biol.*, 45:1495–1509, 2000.

[27] Y. C. Fung. *Foundations of Solid Mechanics.* Prentice-Hall, Englewood Cliffs, New Jersey, 1965.

[28] Y. C. Fung. Stress-strain-history relations of soft tissues in simple elongation. In Y. C. Fung, N. Perrone, and M. Anliker, editors, *Biomechanics: Its Foundations and Objectives.* Prentice-Hall, Englewood Cliffs, New Jersey, 1972.

[29] Y. C. Fung. Biorheology of soft tissues. *Biorheology*, 10:139–155, 1973.

[30] Y. C. Fung. *Biomechanics: Mechanical Properties of Living Tissues.* Springer, New York, 2nd edition, 1993.

[31] Y. C. Fung, K. Fronek, and P. Patitucci. Pseudoelasticity of arteries and the choice of its mathematical expression. *Am. J. Physiol.*, 237:H620–H631, 1979.

[32] J. N. Goodier. Concentration of stress around spherical and cylindrical inclusions and flaws. *Trans. ASME*, 55:39–44, 1933.

[33] J. C. Gore and S. Leeman. Ultrasonic backscattering from human tissue: A realistic model. *Phys. Med. Biol.*, 22(2):317–326, 1977.

[34] A. E. Green and J. E. Adkins. *Large Elastic Deformations.* Oxford University Press, 1970.

[35] J. R. Harris, M. E. Lippman, M. Morrow, and S. Hellman. *Diseases of the Breast.* Lippincott-Raven, 1996.

[36] R. C. Haut and R. W. Little. A constitutive equation for collagen fibers. *J. Biomech.*, 5:423–430, 1972.

[37] C. R. Hill, J. C. Bamber, and G. R. ter Haar, editors. *Physical Principles of Medical Ultrasonics.* Wiley, Chichester, 2nd edition, 2004.

[38] M. H. Holmes. Finite deformation of soft tissue: Analysis of a mixture model in uni-axial compression. *ASME J. Biomech. Eng.*, 108:371–381, 1986.

[39] J. D. Humphrey. Continuum biomechanics of soft biological tissues. *Proc. R. Soc. Lond. A*, 459:3–46, 2003.

[40] C. Johnson. *Numerical Solution of Partial Differential Equations by the Finite Element Method.* Cambridge, 1992.

[41] F. Kallel and M. Bertrand. Tissue elasticity reconstruction using linear perturbation method. *IEEE Trans. Med. Im.*, 15:299–313, 1996.

[42] W. M. Lai, V. C. Mow, and W. Zhu. Constitutive modeling of articular cartilage and biomacromolecular solutions. *ASME J. Biomech. Eng.*, 115:474–480, 1993.

[43] R. S. Lakes and R. Vanderby. Interrelation of creep and relaxation: A modeling approach for ligaments. *J. Biomech. Eng.*, 121:612–615, 1999.

[44] Y. C. Lou and R. A. Schapery. Viscoelastic characterization of a nonlinear fiber-reinforced plastic. *J. Composite. Mater.*, 5:208–234, 1971.

[45] A. E. H. Love. *A Treatise on the Mathematical Theory of Elasticity.* Dover, New York, 4th edition, 1927.

[46] L. E. Malvern. *Introduction to the Mechanics of a Continuous Medium.* Prentice-Hall, New Jersey, 1969.

[47] A. McCulloch, J. B. Bassigthwaighte, P. Hunter, and D. Noble. Computational biology of the heart: from structure to function. *Prog. Biophys. Mol. Biol.*, 69:153–155, 1998.

[48] J. R. McLaughlin and J.-R. Yoon. Unique identifiability of elastic parameters from time-dependent interior displacement measurement. *Inverse Problems*, 20:25–45, 2004.

[49] D. M. McQueen and C. S. Peskin. Heart simulation by an immersed boundary method with formal second-order accuracy and reduced numerical viscosity. In H. Aref and J. W. Phillips, editors, *Mechanics for a New Millennium, Proceedings of the International Conference on Theoretical and Applied Mechanics (ICTAM) 2000.* Kluwer Academic Publishers, 2001.

[50] K. W. Morton and D. F. Mayers. *Numerical SOlution of Partial Differential Equations.* Cambridge University Press, 1994.

[51] V. C. Mow, S. C. Kuei, W. M. Lai, and C. G. Armstrong. Biphasic creep and stress relaxation of articular cartilage in compression: Theory and experiments. *ASME J. Biomech. Eng.*, 102:73–84, 1980.

[52] V. C. Mow, M. K. Kwan, W. M. Lai, and M. H. Holmes. A finite deformation theory for nonlinearly permeable soft hydrated biological tissues. In G. W. Schmid-Schönbein,

S. L.-Y. Woo, and B. W. Zweifach, editors, *Frontiers in Biomechanics*, pages 153–179. Springer-Verlag, New York, 1986.

[53] J. Ockendon, S. Howison, A. Lacey, and A. Movchan. *Applied Partial Differential Equations.* Oxford University Press, 1999.

[54] R. W. Ogden. Elastic deformations of rubberlike solids. In H. Hopkins and M. Sewell, editors, *Mechanics of Solids (The Rodney Hill 60th Anniversary Volume)*, pages 499–538. Oxford, 1982.

[55] R. W. Ogden. *Nonlinear Elastic Deformation.* Dover, New York, 1997.

[56] J. Ophir, S. K. Alam, B. S. Garra, F. Kallel, E. E. Konofagou, T. Krouskop, C. R. B. Merritt, R. Righetti, R. Souchon, S. Srinivasan, and T. Varghese. Elastography: Imaging the elastic properties of soft tissue with ultrasound. *J. Med. Ultrasonics*, 29:155–171, 2002.

[57] J. Ophir, I. Céspedes, H. Ponnekanti, Y. Yazdi, and X. Li. Elastography: A quantiatative method for imaging the elasticity of biological tissues. *Ultrason. Imag.*, 14:111–134, 1991.

[58] K. J. Parker, S. R. Huang, R. A. Musulin, and R. M. Lerner. Tissue respose to mechanical vibrations for sonoelastic imaging. *Ultrasound Med. Biol.*, 16:241–246, 1990.

[59] D. J. Patel and R. N. Vaishnav. The rheology of large blood vessels. In D. H. Bergel, editor, *Cardiovascular Fluid Dynamics*, volume 2, chapter 11, pages 2–65. Academic Press, New York, 1972.

[60] P. Pathmanathan. Deformation modelling for image analysis in breast cancer. M.Sc. dissertation, University of Oxford, 2002.

[61] P.-O. Persson and G. Strang. A simple mesh generator in matlab. *SIAM Review*, 46(2):329–345, 2004.

[62] C. S. Peskin and D. M. McQueen. A three-dimensional computational method for blood flow in the heart: (I) Immersed elastic fibers in a viscous incompressible fluid. *J. Comput. Phys.*, 81:372–405, 1989.

[63] D. P. Pioletti and L. K. Rakotomanana. On the independence of time and strain effects in the stress relaxation of soft tissues. *J. Biomech.*, 33:1729–1732, 2000.

[64] D. P. Pioletti and L. R. Rakotomanana. Non-linear viscoelastic laws for soft biological tissues. *Eur. J. Mech. A*, 19:749–759, 2000.

[65] D. P. Pioletti, L. R. Rakotomanana, J. F. Benvenuti, and P. F. Leyvraz. Viscoelastic constitutive law in large deformations: Application to human knee ligaments and tendons. *J. Biomech.*, 31:753–757, 1998.

[66] H. A. Priestley. *Introduction to Complex Analysis.* Oxford University Press, 1990.

[67] P. P. Provenzano, R. S. Lakes, D. T. Corr, and R. Vanderby Jr. Application of nonlinear viscoelastic models to describe ligament behavior. *Biomech. Model. Mechanobiol.*, 1:45–57, 2002.

[68] P. P. Provenzano, R. S. Lakes, T. Keenan, and R. Vanderby. Non-linear ligament viscoelasticity. *Ann. Biomed. Eng.*, 29:908–914, 2001.

[69] K. R. Raghavan and A. E. Yagle. Forward and inverse problems in elasticity imaging of soft tissue. *IEEE Trans. Nuc. Sci.*, 41:1639–1648, 1994.

[70] K. R. Rajagopal and L. Tao. *Mechanics of Mixtures.* World Scientific, 1995.

[71] J. N. Reddy. *An Introduction to the Finite Element Method.* McGraw-Hill, Singapore, 2nd edition, 1993.

[72] R. Rivlin and D. Saunders. Large elastic deformations of isotropic materials. *Phil. Trans. Roy. Soc. Lond. A*, 243:251–288, 1951.

[73] C. S. Roy. The elastic properties of the arterial wall. *Phil. Trans. R. Soc. Lond. B*, 99:1–31, 1880.

[74] R. A. Schapery. A theory of nonlinear thermoviscoelasticity based on irreversible thermodynamics. In *Proceedings of the 5th U.S. National Conference on Applied Mechanics*, page 511, New York, 1966. ASME.

[75] R. A. Schapery. On the characterization of nonlinear viscoelastic materials. *Polymer. Eng. Sci.*, 9:295–310, 1969.

[76] J. A. Schnabel, C. Tanner, A. D. Castellano-Smith, M. O. Leach, C. Hayes, A. Degenhard, R. Hose, D. L. G. Hill, and D. J. Hawkes. Validation of non-rigid registration using finite element methods. In *Lecture Notes in Computer Science*, volume 2082, Proc. International Conference on Information Processing in Medical Imaging, pages 344–357. Springer-Verlag, Berlin, 2001.

[77] J. A. Schnabel *et al.* Validation of nonrigid image registration using finite-element methods: Application to breast MR images. *IEEE Trans. Med. Im.*, 22(2):238–247, 2003.

[78] A. Singh. An estimation-theoretic framework and a unified perspective. *CVGIP: Image Understanding*, 56:152–177, 1992.

[79] A. R. Skovoroda, S. Y. Emelianov, and M. O'Donnell. Tissue elasticity reconstruction based on ultrasonic displacement and strain images. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, 42:747–765, 1995.

[80] A. J. M. Spencer. *Continuum Mechanics.* Longman, London, 1980.

[81] E. Süli and D. F. Mayers. *An Introduction to Numerical Analysis.* Cambridge University Press, 2003.

[82] A. Ten Eyck and A. Lew. Discontinuous Galerkin methods for non-linear elasticity. *Int. J. Num. Meth. Eng.*, 67:1204–1243, 2006.

[83] A. N. Tikhonov and V. A. Arsenin. *Solution of Ill-posed Problems.* Winston & Sons, Washington, 1977.

[84] C. Truesdell and W. Noll. *The Nonlinear Field Theories of Mechanics.* Springer-Verlag, Heidelberg, 3rd edition, 2004.

[85] R. N. Vaishnav, J. T. Young, J. S. Janicki, and D. J. Patel. Nonlinear anisotropic properties of the canine aorta. *Biophys. J.*, 12:1008–1027, 1972.

[86] D. R. Veronda and R. A. Westmann. Mechanical characteristics of skin — finite deformations. *J. Biomech.*, 3:111–122, 1970.

[87] M. G. Wertheim. Memoire sur l'elasticite et la cohesion des principaux tissues du corps humain. *Ann. Chimie Phys.*, 21:385–414, 1847.

[88] R. L. R. Wesley, R. N. Vaishnav, J. C. A. Fuchs, D. J. Patel, and J. C. Greenfield Jr. Static linear and nonlinear elastic properties of normal and arterialized venous tissue in dog and man. *Circulation Res.*, 37:509–520, 1975.

[89] J. P. Whiteley, S. J. Chapman, D. J. Gavaghan, and J. M. Brady. Nonlinear modelling of breast tissue. *In press*, 2004.

[90] S. L. Y. Woo, M. A. Gomez, and W. H. Akeson. The time and history-dependent viscoelastic properties of the canine medial collateral ligaments. *ASME J. Biomech. Eng.*, 103:293–298, 1981.

[91] M. Yamakawa and T. Shiina. Tissue elasticity reconstruction based on 3-dimensional finite-element model. *Jpn. J. Appl. Phys.*, 38:3393–3398, 1999.

[92] J. Zhou and Y. C. Fung. The degree of nonlinearity and anisotropy of blood vessel elasticity. *Proc. Natl. Acad. Sci. USA*, 94:14255–14260, 1997.