Modelling of some biological materials using continuum mechanics

Cameron Luke Hall

B. App. Sc. (Hons I)
Queensland University of Technology

A thesis submitted for the degree of Doctor of Philosophy in the School of Mathematical Sciences, Faculty of Science, Queensland University of Technology.

2008

Principal Supervisor: Professor D. L. Sean McElwain

Associate Supervisor: Associate Professor Graeme J. Pettet
Associate Supervisor: Professor Zee Upton
Abstract

Continuum mechanics provides a mathematical framework for modelling the physical stresses experienced by a material. Recent studies show that physical stresses play an important role in a wide variety of biological processes, including dermal wound healing, soft tissue growth and morphogenesis. Thus, continuum mechanics is a useful mathematical tool for modelling a range of biological phenomena.

Unfortunately, classical continuum mechanics is of limited use in biomechanical problems. As cells refashion the fibres that make up a soft tissue, they sometimes alter the tissue’s fundamental mechanical structure. Advanced mathematical techniques are needed in order to accurately describe this sort of biological ‘plasticity’.

A number of such techniques have been proposed by previous researchers. However, models that incorporate biological plasticity tend to be very complicated. Furthermore, these models are often difficult to apply and/or interpret, making them of limited practical use.

One alternative approach is to ignore biological plasticity and use classical continuum mechanics. For example, most mechanochemical models of dermal wound healing assume that the skin behaves as a linear viscoelastic solid. Our analysis indicates that this assumption leads to physically unrealistic results.
In this thesis we present a novel and practical approach to modelling biological plasticity. Our principal aim is to combine the simplicity of classical linear models with the sophistication of plasticity theory. To achieve this, we perform a careful mathematical analysis of the concept of a ‘zero stress state’. This leads us to a formal definition of strain that is appropriate for materials that undergo internal remodelling.

Next, we consider the evolution of the zero stress state over time. We develop a novel theory of ‘morphoelasticity’ that can be used to describe how the zero stress state changes in response to growth and remodelling. Importantly, our work yields an intuitive and internally consistent way of modelling anisotropic growth. Furthermore, we are able to use our theory of morphoelasticity to develop evolution equations for elastic strain.

We also present some applications of our theory. For example, we show that morphoelasticity can be used to obtain a constitutive law for a Maxwell viscoelastic fluid that is valid at large deformation gradients. Similarly, we analyse a morphoelastic model of the stress-dependent growth of a tumour spheroid. This work leads to the prediction that a tumour spheroid will always be in a state of radial compression and circumferential tension. Finally, we conclude by presenting a novel mechanochemical model of dermal wound healing that takes into account the plasticity of the healing skin.

**Keywords:** asymptotic analysis, biomechanics, continuum mechanics, dermal wound healing, growth, mathematical biology, mathematical modelling, morphoelasticity, plasticity, residual stress, viscoelasticity, zero stress state
Contents

Abstract i

List of figures viii

Statement of original authorship x

Acknowledgements xi

1 Introduction 1

2 Literature Review 5

2.1 Introduction 5

2.2 Biology of dermal wound healing 7

2.2.1 Normal dermal wound healing in humans 8

2.2.2 Hypertrophic scars and keloids 16

2.2.3 In vitro studies of traction and contraction 19

2.3 Mechanochemical models of dermal wound healing 21
2.3.1 The Tranquillo-Murray model ........................................ 23

2.3.2 Other mechanochemical models ..................................... 29

2.3.3 Models of lattice contraction ........................................ 32

2.4 Mathematical approaches to growth and remodelling .................. 35

2.5 The present work in context ........................................... 43

3 Analysis of the Tranquillo-Murray wound healing model ............... 45

3.1 Introduction ........................................................................ 45

3.2 Outline of the Tranquillo-Murray model ................................ 47

  3.2.1 Description of the model ............................................. 47

  3.2.2 Preliminary mathematical analysis .................................. 54

  3.2.3 Numeric simulations ................................................... 57

3.3 Asymptotic analysis ......................................................... 74

  3.3.1 Simplifying the Tranquillo-Murray model ....................... 74

  3.3.2 Matched asymptotics and the Wright function ................. 79

3.4 Elasticity vs. viscoelasticity ............................................. 97

  3.4.1 Experimental measurements of dermal viscoelasticity .......... 97

  3.4.2 Analysis of an elastic-only model ................................. 103

3.5 Towards a new model of dermal wound healing ....................... 110

4 Zero stress state theory and the definition of strain ................... 113
4.1 Introduction ...................................................... 113
4.2 The classical theory of elastic deformation .................. 116
  4.2.1 Motions, deformations and coordinate systems ......... 116
  4.2.2 The deformation gradient tensor ...................... 119
  4.2.3 Rotations and orthogonal tensors ...................... 121
  4.2.4 The Cauchy-Green tensors ............................ 125
  4.2.5 The classical definition of strain ..................... 128
4.3 Stress, strain and the zero stress state ..................... 133
  4.3.1 Stress and plasticity .................................... 133
  4.3.2 Residual stresses and the zero stress state .......... 136
  4.3.3 The improved definition of strain .................... 146
4.4 Infinitesimal strain approximations ......................... 150
  4.4.1 The principal zero stress deformation gradient ....... 151
  4.4.2 Eulerian infinitesimal strain ........................ 155
  4.4.3 Lagrangian infinitesimal strain ..................... 157
4.5 The challenge of defining strain ........................... 158

5 Morphoelasticity ................................................. 161
  5.1 Introduction ............................................... 161
  5.2 One-dimensional morphoelasticity ........................ 164
  5.3 Three-dimensional morphoelasticity ...................... 175
5.3.1 Volumetric growth ............................................. 175
5.3.2 The pre-symmetrised growth tensor ......................... 181
5.3.3 A unique representation of the growth tensor ............... 188
5.3.4 Incorporating morphoelasticity into a model ............... 198

5.4 Evolution equations for strain .................................. 203
5.4.1 One-dimensional strain evolution ............................ 203
5.4.2 Three-dimensional strain evolution .......................... 206

5.5 Comparison with other mechanical treatments of growth .... 221
5.5.1 The Lagrangian growth equation ............................... 221
5.5.2 Rodriguez et al. and the rate of growth tensor ............. 227
5.5.3 Goriely and Ben Amar’s cumulative theory of growth ... 231

5.6 Summary of zero stress state theory and morphoelasticity .... 239

6 Applications of zero stress state theory and morphoelasticity 246

6.1 Introduction .................................................. 246
6.2 A one-dimensional Maxwell material .......................... 249
6.3 Modelling growth ............................................. 260
6.3.1 Dependence of the growth tensor on stress ............... 260
6.3.2 A model of spherically-symmetric growth .................. 267
6.4 A mechanochemical model of dermal wound healing ........ 280
6.4.1 Modelling approach ....................................... 280
<table>
<thead>
<tr>
<th>Section</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.4.2 Development of model equations</td>
</tr>
<tr>
<td>6.4.3 Initial conditions and boundary conditions</td>
</tr>
<tr>
<td>6.4.4 Nondimensionalisation</td>
</tr>
<tr>
<td>6.4.5 Preliminary observations and further work</td>
</tr>
<tr>
<td>6.5 Further applications of morphoelasticity</td>
</tr>
</tbody>
</table>

7 Conclusions and further work 318

Appendices 322

A Asymptotic analysis of a general viscoelastic wound healing model 322

B Existence and uniqueness of the zero stress state 338

B.1 Existence of the zero stress Cauchy-Green tensor 338

B.2 Uniqueness of the zero stress Cauchy-Green tensor 343

C Rodriguez et al.’s definition of Lagrangian strain 346

D Finite strain evolution and related topics 349

D.1 Preliminary results 349

D.2 Evolution of finite strain 352

D.3 Observer independence 353

Bibliography 358
# List of Figures

2.1 Illustration of the healthy skin ........................................ 8
2.2 Inflammation .............................................................. 11
2.3 Proliferation ............................................................... 12
2.4 Maturation ................................................................. 15
3.1 A narrow rectangular wound ............................................. 48
3.2 Numeric results obtained from the base Tranquillo-Murray model 59
3.3 Tranquillo-Murray model showing oscillations at large values of $x$ 61
3.4 The Tranquillo-Murray model with and without advection: cell density profiles ......................................................... 63
3.5 The Tranquillo-Murray model with and without advection: ECM displacement profiles ....................................................... 64
3.6 Tranquillo-Murray model: $\gamma = 0$ .................................. 66
3.7 Tranquillo-Murray model: $r_0 = 0$ .................................... 68
3.8 Viscous-only Tranquillo-Murray model: $\bar{E} = 0$ ............... 70
3.9 Tranquillo-Murray model: $\mu \ll 1$ .................................... 72
3.10 Tranquillo-Murray model: \( s = 0 \) .............................................. 73

3.11 Comparison of numeric results from the simplified Tranquillo-Murray model with a short-space asymptotic approximation .................. 86

3.12 Comparison of numeric results from the simplified Tranquillo-Murray model with a long-space asymptotic approximation .................. 94

3.13 ECM displacement profiles from the elastic-only Tranquillo-Murray model ................................................................. 108

4.1 Two particles within a body ......................................................... 119

4.2 The relationship between the initial configuration, the current configuration and the zero stress state. ............................ 138

5.1 A growing one-dimensional body composed of discrete cells .... 166

5.2 Comparison of the desired length and actual length of a single cell 166

5.3 A general prism as a volume element ............................................ 185

5.4 Diagram of a growing prism .................................................... 187

6.1 A Maxwell element ................................................................. 249

6.2 Comparison of the initial profiles of traction stress, elastic stress and total stress. ............................................................... 284

B.1 The curve \( C \) in a two-dimensional domain ............................ 340

B.2 Diagram showing the ‘wall’, \( G \) ............................................ 342
Statement of original authorship

The work contained in this thesis has not been previously submitted to meet the requirements for an award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

Signed:

Date:
Acknowledgements

First and foremost, I’d like to thank my supervisor, Sean McElwain, for his guidance and support over the last three years. We’ve traversed a wide range of topics in that time and I’ve enjoyed and appreciated our regular discussions. I’d especially like to thank Sean for his bemused tolerance (and patience!) as the scope of this project broadened and my investigations became more theoretical. In the end, I feel that the final product is something that we can be proud of.

I’ve also had many interesting and informative discussions with other QUT staff members, with my fellow students and with visitors to the mathematics department. Without these discussions, this thesis would look very different (for example, I would never have recognised the series definition of the Wright function on my own). I’m grateful to everyone who has lent a listening ear and a thinking brain to the problems I’ve shared.

In particular, I’d like to thank Helen Byrne and Colin Please, two of our regular visitors at QUT. Talking to Helen and Colin has greatly helped me to clarify my thoughts and bring my work into perspective. I’m especially grateful to Helen, Colin and Philip Maini for their hospitality during my visit to the UK in 2006.

Last, but by no means least, I’d like to thank my family for their unwavering love and support. Over the last few months, I’ve been particularly appreciative of the tolerance and patience shown by my friends and family. Having waited so long, you should finally be able to get ‘old Cameron’ back again.

Cameron Hall

25 August 2008; 22 January 2009

Make everything as simple as possible, but not simpler.
Albert Einstein
Chapter 1

Introduction

Many biological processes involve the mechanical interaction of cells with their tissue substrate. In dermal wound healing, for example, fibroblasts and myofibroblasts actively pull on the fibres that make up the skin, ultimately leading to wound contraction [117, 186]. As they do so, these cells respond to mechanical cues from their surrounding environment; *in vitro* and *in vivo* experiments indicate that the tensile stress experienced by the skin affects the final quality of a dermal scar [1, 171, 186].

Wound healing is not the only biological process in which mechanical interactions are significant. Recent studies of morphogenesis [43, 169], bone remodelling [84, 108], stem cell therapy [145] and tumour growth [25, 88] all demonstrate the biological importance of mechanical forces. Unfortunately, it is very difficult to measure these forces in a non-invasive and non-destructive manner. Despite significant improvements over the past fifteen years, new experimental and theoretical methods are needed in order to develop a thorough and well-grounded understanding of three-dimensional biomechanics [146].

Given the limitations of experimental techniques in biomechanics, mathemati-
cal modelling is essential.\textsuperscript{1} In particular, continuum mechanics provides a useful theoretical framework for studying the stresses that are present in biological materials. As described in the recent review by Humphrey [96], continuum mechanical models have been developed for such diverse tissues as blood vessels, cartilage, muscle and the embryonic heart.

However, continuum biomechanics is still in its infancy. In general, the classical theories of continuum mechanics are not well suited to modelling biological materials. Soft tissues, for example, are often highly anisotropic and many of them exhibit unusual viscoelastic properties. Furthermore, many tissues in vivo undergo a continual process of internal revision and mechanical restructuring. Sometimes, this leads to a situation where the mechanical behaviour of a biological material changes markedly over time.

Various mathematical approaches have been developed to address these difficulties. One promising theoretical framework for modelling mechanical restructuring is the tensorial representation of the zero stress state introduced by Rodriguez et al. [159]. However, this approach generally leads to a highly complicated system of equations that is difficult to interpret and difficult to apply computationally (see, for example, Lubarda and Hoger [113]). As a consequence, this approach has generally been limited to models of relatively simple processes (e.g. the growth of tumour spheroids [5]).

When modelling more complex biomechanical phenomena, researchers have generally been content to use classical linear viscoelasticity; see, for example, Tranquillo and Murray [189], Olsen and coworkers [136, 140] and Ferrenq et al. [58]. In doing so, they effectively ignore the complexities introduced by anisotropy and mechanical restructuring. If such effects are significant, classical linear models

\textsuperscript{1}In his doctoral thesis, Cook [33] makes the following remark:

‘Why use mathematics to study something as intrinsically complicated and ill-understood as wound healing? We would argue that mathematics must [original emphasis] be used if we hope to truly convert an understanding of the underlying mechanisms involved into a predictive science. Mathematics is required to bridge the gap between the level on which most of our knowledge is accumulating (cellular and below) and the macroscopic level of the scar itself[,] which is of primary concern to the plastic surgeon and the patient.’ – Cook (1995) [33] p3
will yield inaccurate results. However, these models have the advantage that they are considerably easier to analyse, interpret and use than the densely theoretical models that incorporate anisotropy and restructuring.

In this thesis, we seek to map out a middle ground between these two extremes. We aim to construct a framework for biomechanical modelling that accurately describes mechanical restructuring but is not unnecessarily complicated. We show that it is possible to construct models that retain the relative simplicity of linear models while incorporating a representation of the evolving zero stress state.

The thesis is structured as follows:

Chapter 2 presents a review of the biological and mathematical literature that is relevant to our investigations. We begin by describing the current understanding of dermal wound healing, emphasising recent experimental studies into the role of mechanical tension in healing. We also discuss the mechanochemical models of dermal wound healing that have been developed, outlining their strengths and deficiencies. Finally, we review the use of the zero stress state to describe biomechanical remodelling.

Chapter 3 contains a detailed analysis of the Tranquillo-Murray model of dermal wound healing [189]. This analysis yields some intriguing results. Using Tranquillo and Murray’s parameter values, we are able to show that the mechanical parts of the model can be removed without having any significant effect on the rest of the model. Furthermore, we find that viscoelastic constitutive law used by Tranquillo and Murray causes their model to predict oscillations in displacement that do not appear to be physically realistic.

Chapter 4 contains a thorough presentation of zero stress state theory. We briefly review classical solid mechanics and we present our own theoretical justification for the multiplicative decomposition of the deformation gradient. This leads naturally to a definition of strain that incorporates the changing zero stress state.
We use this to construct a novel formulation for infinitesimal strain that is valid in a number of biologically relevant situations.

Chapter 5 details the theory that we have developed to describe the evolution of the zero stress state. Following the example of Goriely and coworkers [70], we refer to this as a theory of ‘morphoelasticity’. By considering spatially-biased volumetric growth, we develop a tensorial description of the changing zero stress state. Using results obtained in Chapter 4, this leads to an evolution equation for strain that can be used in practical contexts.

Chapter 6 outlines some applications of morphoelasticity. Firstly, we demonstrate that a Maxwell material can be modelled as a morphoelastic material where the rate of growth is dependent on the stress. Next, we investigate a morphoelastic model of stress-dependent growth, obtaining some analytic results concerning the stress distribution within a growing tumour spheroid. Finally, we describe how our theory might be used to construct an improved model of dermal wound healing.

Finally, Chapter 7 summarises our achievements and outlines directions for further research. One of the major challenges that we face is developing appropriate numeric methods for the models described. We discuss the strengths and limitations of our theory and we illustrate a number of possible applications within biomechanics.
Chapter 2

Literature Review

2.1 Introduction

By its very nature, biomechanics is an interdisciplinary pursuit. Biologists, clinicians, mathematicians and engineers have all made significant contributions to the current understanding of the mechanical behaviour of biological materials. Importantly, each of these disciplines brings its own distinct strengths to the broader field of study. Advances in such diverse fields as cell biology, polymer rheology and computational mathematics can significantly impact on the development of biomechanics.

As a result, we find that the biomechanical literature covers a wide range of topics and techniques. Furthermore, there is a significant body of work outside the narrow definition of ‘biomechanics’ that is potentially relevant to our investigations. The sheer breadth of this field prevents any review from being truly comprehensive. Hence, it behooves us to focus our attention on some specific topics. In the present work, we concentrate our efforts on mechanical models of dermal wound healing and other studies of biological plasticity/morphoelasticity.
At present, there is much to be learnt about the mechanics of dermal wound healing. Recent experimental studies support the view that the importance of mechanical stresses to dermal wound healing has previously been underestimated [1]. It is now understood that the cells of the skin respond directly to mechanical cues from their surrounding environment. Interestingly, this has not previously been incorporated into mathematical models of dermal wound healing.

Similarly, there are very few mechanochemical models that account for the ‘plastic’ behaviour of the healing dermis that is observed experimentally. Mathematically, this behaviour can be modelled using the concept of a ‘zero stress state’; a theoretical description of the zero stress state was first introduced to biomechanics by Rodriguez et al. in 1994 [159]. However, modelling the evolving zero stress state leads to some practical and theoretical problems. Despite considerable advances over the last fifteen years, the present theories of morphoelasticity are generally too complicated to be incorporated into practical and interpretable models of wound healing.

Our literature review is divided into three main sections. Firstly, in Section 2.2, we describe the current biological understanding of dermal wound healing. We outline the stages of dermal wound healing and we discuss in vitro and in vivo experiments that have been used to investigate the effect of dermal mechanics on cell behaviour. On a macroscopic level, these results indicate that skin tension affects the likelihood of a wound developing into a hypertrophic scar or keloid.

Next, in Section 2.3, we review the mechanochemical models of dermal wound healing in the current literature. Interestingly, there has been very little activity in this area over the last decade. Although wound healing continues to attract considerable attention from mathematical biologists, few have attempted to include mechanical effects in their models.

In Section 2.4, we consider the tensorial representation of the zero stress state that was introduced to biomechanics by Rodriguez et al. [159]. Some recent theoretical advances in zero stress state theory are described and we outline some
mathematical models that use this theoretical framework. Much of the work described in later chapters is devoted to exploring and extending this theory.

Finally, in Section 2.5, we summarise the literature review, giving a biological and mathematical context for the work that follows.

## 2.2 Biology of dermal wound healing

Successful dermal wound healing is a complex biological process that involves an interplay between cells, bioactive compounds and mechanical stresses. After a wound to the skin occurs, the body responds quickly. Firstly, it acts to protect against infection and secondly it repairs the wound with continuous tissue.

This repaired tissue is called scar tissue and it has structural features that differ from the healthy dermis. Scars tend to be less flexible and less functional than the surrounding skin and they may stand out as being visibly different. A number of research groups are studying dermal wound healing with the ultimate aim of developing treatments that will result in scar-free healing (see, for example, Ferguson [57] and Liu et al. [109]).

In some cases, even the reduced flexibility and functionality of normal scar tissue is never achieved. Hypertrophic scars and keloid disease are both examples of overhealing, where too much scar tissue is formed. This can lead to regions of greatly reduced dermal flexibility and that are both uncomfortable and unsightly [14, 161, 178]. At the opposite extreme, chronic ulcers (particularly common amongst diabetics and the elderly) can be considered to be examples of underhealing, where wounds never close properly. Chronic wounds require continual dressings and medical attention, making them very expensive to treat [87, 120].

In this section, we present an overview of recent studies into the biology of dermal wound healing. We begin by describing normal dermal wound healing, empha-
sising the interaction of chemical and mechanical signals. Then, we discuss the relationship between overhealing disorders and dermal mechanics. Finally, we conclude by outlining some of the *in vitro* experiments that have been used to investigate the role of tissue mechanics on the microscopic level.

### 2.2.1 Normal dermal wound healing in humans

![Figure 2.1: A diagram illustrating the healthy skin.](image)

Like most other tissues, the skin is largely made up of extracellular matrix (ECM), cells and water. The dermal ECM is a fibrous material made from collagen, elastin and other biopolymers. Effectively, it acts as the ‘ground substance’ for the tissue, providing cells with a physical substrate that enables them to migrate and perform their normal functions. The ECM also contains growth factors and other agents that affect the behaviour of cells [35, 54, 188].

The healthy skin comprises of three layers: the epidermis (uppermost), the dermis (middle) and the hypodermis or subcutaneous layer (lowermost); these are
depicted in Figure 2.1. The epidermis consists mainly of epithelial cells, which form a continuous, water-resistant layer that protects the lower tissues [117, 177]. In contrast, the dermis contains a variety of cells and biological structures. For example, sweat glands, blood vessels and hair follicles can all be found in the dermis but not in the epidermis [117, 177]. The hypodermis lies below the other layers and it connects the dermis to the fleshy tissues beneath.

Wounds to the skin can be categorised in many different ways. For example, wounds can be described by their cause (e.g. abrasion, laceration or thermal burning), by the time taken to heal (acute or chronic) and by their depth [172, 177]. The depth of a wound is thought to be important to the quality of healing, regardless of the cause of the wound. Recent studies have shown that an incisional wound will only lead to a scar if the wound is sufficiently deep [52].

Wounds that affect the epidermis and not the dermis are called epidermal wounds or superficial wounds. These wounds heal by the proliferation and migration of epithelial cells to cover the site of the wound [21]. In contrast, dermal wound healing is more complex. Simultaneously with the regeneration of the damaged epidermis, fibroblasts act to repair and reconstruct the dermis [177]. Reconstructed dermal tissue is called scar tissue and it can be distinguished from healthy tissue by its physical and histological properties (see, for example, Cuttle et al.’s study of porcine hypertrophic scars [38]).

In our work, we concentrate on dermal wound healing rather than epidermal wound healing. From a clinical perspective, dermal wound healing is important because problem scars are more likely to arise from deep wounds [52, 178].\footnote{It should be noted that the processes of epidermal wound healing and dermal wound healing are mutually dependent. Recent studies indicate that chemical communication between the epidermis and the dermis during wound healing can affect scarring [17, 194].} Interestingly, the fibroblasts of the dermis are more mechanically active than epithelial cells [55]. Furthermore, the mechanical environment of a dermal wound is known to affect the quality of the final scar [1]. Thus, mechanochemical models promise to be more useful for describing dermal wound healing than epidermal wound healing.
One of the important aspects of dermal wound healing is wound contraction [117]. As a dermal wound heals, cells actively pull on the fibres of the ECM, causing the wound to become smaller. In small mammals such as mice, wound contraction is thought to be the main mechanism of wound closure [179]. Although wound contraction is less important in humans than in small mammals, some wound healing pathologies (particularly hypertrophic scars) can be related to excessive wound contraction [161, 179]. Much of the modelling work that we present is explicitly concerned with understanding wound contraction.

The process of dermal wound healing is generally divided into three overlapping phases: inflammation, proliferation and maturation (also called remodelling) [21, 29, 87, 117]. These may be described as follows:

**Inflammation**

Inflammation is the immediate response of the body to injury and it has the purpose of ‘cleaning up’ the damaged tissue and preventing infection [21]. Most wounds entail damage to blood vessels, leading to bleeding. The first step in wound healing is the formation of a clot that prevents the loss of blood and protects the body from infection. The clot also provides an initial matrix through which cells can move [21, 124].

Once the clot has formed (usually in the first few hours), white blood cells begin to move in. This process is shown in Figure 2.2. The white blood cells have two purposes: to ‘clean up’ the dead and dying tissue and to guard against infection [21, 29]. White blood cells are attracted to the wound by growth factors produced during the clotting process and by chemical signals emitted from dying

---

2Some researchers refer to haemostasis (i.e. clot formation) as a separate phase of wound healing from inflammation, giving a total of four phases (see, for example, Chin et al. [27]). Since our investigations are mainly concerned with the proliferation and maturation phases of wound healing, the distinction between haemostasis and inflammation is not important to us. Other researchers refer to even more phases of wound healing; Strecker-McGraw et al. [177] list seven overlapping phases: haemostasis, inflammation, angiogenesis, fibroplasia, contraction, epithelialisation and remodelling.
Figure 2.2: A diagram illustrating the middle stages of inflammation. Within the wound space, the dermis and epidermis have been replaced with a fibrin clot. Neutrophils (green) and monocytes/macrophages (maroon) have been recruited to the wound site, where they release chemotactic factors that will attract fibroblasts and endothelial cells.

cells [21, 29]. At the same time, blood vessels near the wound become more porous in structure and they leak plasma proteins into the surrounding tissue. These proteins attract more white blood cells to the wound [29, 117].

The initial inflammation is characterised by the presence of neutrophils and monocytes, two types of white blood cell. As the inflammation progresses, the number of neutrophils decreases and the monocytes differentiate into macrophages [21, 29]. Macrophages consume microbes and debris in the wound space; as they do so, they release growth factors and chemotactic factors into the wound [21]. These factors are important to the next stage of wound healing, proliferation.

Proliferation

The proliferative stage of wound healing begins after a few days and can continue for several weeks (or in the case of keloid disease, indefinitely) [35, 178]. During proliferation, cells from around the wound and from the underlying tissues undergo mitosis. The new cells that are produced then migrate into the wound
A diagram illustrating the proliferative stage of dermal wound healing. The fibrin clot is being replaced with a matrix composed from collagen and elastin, but the new tissue is different from the surrounding dermis. In particular, the newly synthesised fibres are generally aligned to be parallel with each other. All of the neutrophils have gone, but a large number of macrophages persist in the granulation tissue, releasing growth factors that promote the migration and proliferation of fibroblasts and endothelial cells. Some of the fibroblasts have differentiated into myofibroblasts (red), which exert greater traction stresses.

Granulation tissue is highly vascularised; visible nodules of closely-looped capillaries give granulation tissue its eponymous grainy appearance [21]. These capillaries supply necessary nutrients and oxygen to the macrophages and fibroblasts in the granulation tissue. In turn, the macrophages act to clean the wound space of debris and the fibroblasts act to reconstruct the dermal ECM [21, 35]. It has been observed that macrophages, fibroblasts and endothelial cells move into the wound together in small units, called wound modules [21]. Within a wound module, the macrophages release chemotactic factors that attract the fibroblasts and endothelial cells, while the fibroblasts lay down new ECM, providing the migrating endothelial cells with an appropriate substrate [21].
The process of endothelial cell proliferation is called angiogenesis.\(^3\) Successful angiogenesis is critical for effective healing; impaired angiogenesis can lead to the development of chronic ulcers [64]. It is interesting to note that a number of mathematical models have been developed to investigate the role of angiogenesis in dermal wound healing. In particular, Olsen and coworkers [136, 142] and Pettet and coworkers [149, 150] have made significant contributions in this area. However, there are currently no models of wound healing that incorporate both angiogenesis and mechanical effects.

The proliferation and migration of fibroblasts into the wound space is known as fibroplasia. During this process, fibroblasts are known to change their phenotype in response to cues from their environment [2, 76, 186]. The main two phenotypes observed are ‘normal’ fibroblasts and myofibroblasts. Myofibroblasts are characterised by the presence of \(\alpha\)-smooth muscle actin and they are thought to be important to wound contraction [2, 117].

Some experiments suggest that there is also an intermediate fibroblast phenotype, called the protomyofibroblast. These cells have some stress fibres, but they do not exhibit all of the histological features of a myofibroblast [45, 186]. All fibroblast phenotypes are important to dermal wound healing. Effectively, they all have the same function: to replace the damaged and destroyed fibres of the ECM [186]. However, they go about this in different ways.

‘Normal’ fibroblasts are more motile than myofibroblasts [162] and they are attracted into the wound space by chemical factors expressed by inflammatory cells [117]. However, these fibroblasts do not form adhesion-complexes with the ECM and they are unable to exert significant forces on the wound tissue [74, 186]. Thus, it is thought that protomyofibroblasts and myofibroblasts are the cell types that are principally responsible for wound contraction [76, 117, 186].\(^4\)

\(^3\)More specifically, angiogenesis refers to the growth of new blood vessels. Endothelial cell proliferation is an essential part of this process.

\(^4\)There is an ongoing debate about the relative importance of ‘normal’ fibroblasts and myofibroblasts to wound contraction [74, 186]. The general consensus appears to be that they are both important, but they have different functions. One interesting suggestion is that fibroblasts
Evidence from \textit{in vitro} experiments suggests that fibroblasts modulate into protomyofibroblasts in response to the mechanical tension that they experience in the wound space [45, 75, 76, 171]. Protomyofibroblasts express actin filaments and form focal adhesions, allowing them to exert greater forces on the surrounding tissue [74, 186]. Furthermore, some studies indicate that protomyofibroblasts synthesise collagen at a faster rate than normal fibroblasts [99].

With further mechanical stress and the presence of certain growth factors, protomyofibroblasts differentiate into myofibroblasts [74, 186]. These cells are able to exert large contractile stresses. Effectively, the presence of myofibroblasts turns the granulation tissue into a temporary contractile organ [117, 186]. The relationship between fibroblast behaviour and mechanical tension is still being studied. We describe some important experimental results in Section 2.2.3.

Fibroplasia is particularly interesting because it is a recursive process. Fibroblasts and myofibroblasts exert stresses on the ECM and they respond to stresses in the ECM [55, 186]. Other experiments have shown that fibre alignment is a similarly recursive process. Fibroblast migration is directed by fibre alignment at the same time as fibroblasts act to realign fibres [29, 168].

Such recursive relationships are well-suited to being investigated using mathematical models. Indeed, there is a growing body of mathematical research into fibre alignment by fibroblasts (see especially the model proposed by Dallon \textit{et al.} [41]). However, at this stage there are no mathematical models that investigate the interdependency of tissue stress and fibroblast behaviour.

**Maturation**

At the same time as they repair the ECM by laying down new fibres, fibroblasts make modifications to the existing structure. This process, which is called
Figure 2.4: A diagram illustrating a mature scar. Granulation tissue contains a dense population of fibroblasts, myofibroblasts, endothelial cells and macrophages; however, mature scars are almost acellular. As shown above, many of the collagen and elastin fibres in a scar, are aligned with one another, perpendicular to the wound boundary.

maturation, continues until a mature scar is formed [21, 29]. A mature scar is depicted in Figure 2.4. Maturation begins at the same time as the formation of the granulation tissue and then continues for months or even years [21].

As noted earlier, the granulation tissue formed during wound healing contains a high density of cells [21, 29]. As the granulation tissue matures into scar tissue, many of these cells undergo apoptosis (programmed cell death). Thus, the highly vascular granulation tissue is replaced with scar tissue that is almost avascular [21, 29]. Interestingly, it has been observed that scar tissue contains fibroblasts but not myofibroblasts [74]. It is not yet clear whether the myofibroblasts present in granulation tissue differentiate back into fibroblasts or whether they are removed by apoptosis [76, 186].

One of the most important features of the maturation process is the continual revision of the ECM [21, 29, 168]. In the early stages of maturation, fibroblasts

---

5The signal that triggers apoptosis in wound maturation is not yet fully understood [1, 72]. However, there is evidence that the Akt pathway is important and it has been shown that hypertrophic scars are more severe in mice that lack the p53 gene, which promotes apoptosis [1]. Also, some studies have indicated that the apoptosis of fibroblasts and endothelial cells is related to the interaction of interferons and TGF-β [72, 133].
break down the provisional fibrin matrix and replace it with a collagen-elastin matrix [21, 168]. As maturation progresses, fibroblasts continue to modify this matrix; they degrade the existing collagen and elastin fibres at the same time as they excrete new ones. As a result, the mechanical properties of the scar can change significantly over the course of maturation [21, 168].

Another interesting aspect of scar formation is the fact that the mature collagen fibres tend to be aligned with each other. In contrast, the collagen fibres within uninjured tissue are randomly aligned [168]. The dominant direction of fibre alignment in a scar is affected by the shape of the original wound and the direction of tension in the surrounding skin [33]. When fibres are strongly aligned, the physical properties of the scar tissue are observably different from the physical properties of the surrounding tissue. For example, scar tissue is visibly different from the healthy dermis and it is often mechanically weaker [33, 168].

2.2.2 Hypertrophic scars and keloids

Hypertrophic scars and keloids are both examples of pathological scarring. In these disorders, the body overresponds to a dermal injury, producing too many fibroblasts and too much ECM during the healing response [161, 178]. Hypertrophic scars and keloids are much less flexible than normal scar tissue. Furthermore, they are unsightly; these scars often have a raised profile and an irregular texture [14, 161, 178].

Despite some similarities, hypertrophic scars and keloids are quite different. Generally, hypertrophic scars regress spontaneously over time and they rarely extend far outside the edges of the original wound. In contrast, keloids invade the surrounding tissue aggressively [161, 178]. Keloid disease is characterised by the fact that keloids continually grow outside the original confines of the wound space and rarely regress without treatment [14, 161, 178]. This condition is hereditary and susceptible people can develop keloids from very small wounds, including insect
bites and body piercings [14, 161].

The process by which a pathological scar develops is similar to the process of normal dermal wound healing but overexaggerated. In pathological scars, the inflammatory response is greater and it often continues for an unusually long time (particularly in the case of a keloid). Apoptosis is downregulated and the fibroblasts in pathological scars continue to proliferate and synthesise ECM for much longer than normal, possibly as a result of the extended inflammation [72, 133, 161, 178]. In a keloid the inflammation and proliferation stages never truly end, leading to a large, persistent growth [161].

Studies of pathological scarring indicate that mechanical and chemical factors are both important. For example, the relative concentrations of the isoforms of transforming growth factor-\(\beta\) (TGF-\(\beta\)) can be used to predict the severity of scarring [57, 124]. Continued production of TGF-\(\beta_1\) and TGF-\(\beta_2\) is associated with serious and pathological scarring. In contrast, elevated levels of TGF-\(\beta_3\) and decreased levels of the other isoforms are observed in the scarless wound healing exhibited by embryos [57, 109, 124].

Recent in vivo studies show that mechanical tension is also important. Aarabi and coworkers [1, 19] performed experiments in which they used a novel device to apply a tensile load to a healing murine wound. They found that loading the wound during the inflammatory phase led to wound dehiscence, preventing healing. However, if the wound was loaded during the proliferative phase, a hypertrophic scar developed. Like human hypertrophic scars, these murine scars were characterised by epidermal thickening, aligned collagen fibres and the presence of collagen ‘whorls’ [1].

Aarabi and coworkers used their murine model to investigate the effects of increased mechanical tension on the cellular level. Immunohistochemical comparison of loaded wounds with controls indicated that the cellular proliferation was

---

\(^6\)Interestingly, a genetic study failed to find a significant relationship between keloid disease and TGF-\(\beta_2\) polymorphisms, despite the fact that keloid disease is characterised by elevated TGF-\(\beta_2\) levels [15]. Further research is needed in this area.
not substantially changed by loading. Instead, Aarabi and coworkers found that mechanical stress leads to a significant downregulation of apoptosis. Interestingly, this effect was found to be permanent; if the tension was relaxed after two weeks of loading, apoptosis remained low and the scar continued to show hypertrophic features.

These experiments are consistent with the clinical observation that hypertrophic scars are most likely to occur in places where the skin is under greatest tension (e.g. the torso and back) [161, 178]. Furthermore, other studies indicate that pressure therapy is very effective for treating hypertrophic scars [115, 161]. Based on these results, we can safely conclude that pathological scarring is neither purely chemical nor purely mechanical in origin. Instead, it is the (faulty) interplay of chemical and mechanical signals that is important.

We are still in the process of understanding the many signals that contribute to the formation of pathological scars. For example, Bellemare et al. [17] recently investigated the role of epidermal-dermal communication in hypertrophic scars. In their experiments, skin-like structures were formed by combining epidermal cells and dermal cells obtained from different patients. Interestingly, they found that fibroblasts from healthy tissue proliferated most when they were cultured with epidermal cells from hypertrophic scars. This suggests the existence of signalling between the epidermis and the dermis that affects the quality of scarring.\(^7\)

Hypertrophic scars and keloids are complex disorders that are still poorly understood. One tool that can be used to investigate the pathogenesis of these disorders is mathematical modelling. Any mathematical model of pathological scarring needs to take into account the combination of chemical, cellular and mechanical effects that are involved. As we will see, none of the existing models of dermal wound healing are adequate for this task.

\(^7\)Wearing and Sherratt [193] developed a mathematical model of wound healing that incorporates chemical communication from the dermis to the epidermis. Similarly, Cruywagen and Murray [37] have constructed a mechanochemical model of pattern formation in the skin that includes mechanical interactions in both directions (epidermis to dermis and dermis to epidermis). However, to the best of our knowledge, no model of wound healing has modelled the effect of chemical signals from the epidermis to the dermis.
2.2.3 In vitro studies of traction and contraction

In the body, fibroblasts and myofibroblasts exist in a three-dimensional network of fibres. Various in vitro experimental techniques have been developed to study the behaviour of these cells in a similar environment. In particular, experiments on fibroblast-populated collagen lattices (FPCLs) have allowed researchers to investigate the interaction between fibroblasts and their mechanical environment.\(^8\)

In the 1980s, Grinnell and coworkers pioneered the use of FPCLs to study the traction forces exerted by fibroblasts [77, 80, 81, 82].\(^9\) They (and other groups) demonstrated that floating lattices seeded with fibroblasts contract significantly when cultured in serum; in some experiments, the lattices contracted to a tenth of their original size [80, 81, 117, 186]. In other experiments, lattices were tethered or mechanically loaded. In these cases, little contraction was observed and isometric tension developed [76, 79, 183, 186].

Interestingly, fibroblasts in tethered lattices are found to be morphologically different from fibroblasts in untethered lattices. Fibroblasts in untethered lattices have a dendritic appearance, while fibroblasts in tethered lattices have a stellate appearance [74, 157]. More importantly, only the fibroblasts in tethered lattices exhibit focal adhesions and stress fibres [74, 186]. Effectively, the mechanical tension in the tethered lattices causes the fibroblasts to differentiate into protomyofibroblasts, or even myofibroblasts.

Since protomyofibroblasts are only observed in tethered lattices, it is quite surprising that the contraction of untethered lattices is so dramatic. Interestingly, the undifferentiated fibroblasts in floating lattices manage to generate tractional forces without shortening in the way that muscle cells do [75, 117]. Instead,

---

\(^8\)There have also been a number of in vivo experiments on fibroblast mechanobiology (see, for example, Hinz et al. [90]). However, FPCL experiments have yielded a large quantity of data that cannot be obtained easily from in vivo experiments.

\(^9\)The contraction of collagen lattices by fibroblasts was originally described by Bell et al. [16] in 1979. However, Grinnell and coworkers were the first to make extensive use of FPCLs to study fibroblast traction.
floating collagen lattices contract because the fibroblasts physically rearrange the existing fibres [75, 81, 117].

Interestingly, this leads to a situation where much of the lattice contraction persists even when the cells are removed. Guidry and Grinnell [81] performed experiments where a floating FPCL was allowed to contract and the fibroblasts were killed after a certain length of time. After the fibroblasts were killed, the lattices expanded. However, they did not return to their original size. Thus, the rearrangement of fibres has a permanent effect on the collagen lattice that survives the removal of the fibroblasts.

Some experimenters modified the classic FPCL contraction experiment by using a cell force monitor (CFM) to measure the total force exerted by the cells [53, 59, 60, 119]. In a CFM, one end of the FPCL is held fixed while the other end is connected to a long needle of known stiffness that stands perpendicular to the lattice. As the lattice contracts, the needle bends and its deflection is measured using strain gauges. The resistance of the needle to bending exerts a calculatable force on the collagen lattice. This is used as a measure of the forces being exerted by the cells in the lattice [53, 59].

Like Guidry and Grinnell [81], Marenzana et al. [119] performed FPCL experiments in which the fibroblasts were killed after a certain time. However, their experiments used a CFM setup instead of a floating lattice. They found that the needle remains bent, even after the cells have been removed. Effectively, the ‘plastic’ effects of fibre rearrangement alter the needle deflection that corresponds to a state of mechanical equilibrium. This indicates that care needs to be taken when interpreting the results of CFM experiments. Most experimenters have assumed that the deflection of the needle is directly proportional to the traction

---

10 In one series of experiments, Guidry and Grinnell [81] used radiolabelled collagen lattices to investigate the possibility that fibroblasts in floating lattices degrade the existing lattice and synthesise new collagen. Their results indicated that collagen synthesis and degradation were both very limited; the main mechanism of lattice contraction is fibre rearrangement.

11 Interestingly, this indicates that it is inappropriate to model a contracting FPCL using a standard Kelvin-Voigt viscoelastic model. Such a model would predict that the lattice always returns to its original size when the fibroblasts are killed.
force exerted by the cells [53, 59, 199]. Assuming that the traction stress is summative with the elastic stress, Marenzana et al.’s results indicate that only part of the observed deflection is due to active traction.

A variety of other experiments have been performed using FPCLs. For example, Petroll and coworkers [148, 162, 163] embedded microscopic spheres into collagen lattices and used the displacement of the spheres to measure the mechanics of the fibroblast-lattice interaction on a microscopic scale. Interestingly, they found that stationary fibroblasts exert greater tractional forces than moving fibroblasts [163].

The experimental results obtained by Grinnell, Marenzana, Petroll and others illustrate the fact that the mechanical behaviour of fibroblasts is subtle and complex. We now know that the traction forces exerted by fibroblasts can lead to significant ‘plastic’ changes to the ECM. Similarly, it has been demonstrated that fibroblasts adjust their mechanical behaviour according to their environment and mechanical tension affects the differentiation of fibroblasts into myofibroblasts. As we will see, mathematical models can be used to further investigate these features of fibroblast behaviour.

2.3 Mechanochemical models of dermal wound healing

As described in the reviews by Sherratt and Dallon [168] and Murray et al. [130], mathematicians have modelled dermal wound healing in a variety of different ways. The present literature includes cellular automata (CA) models [41], ordinary differential equation models [30, 31], reaction-diffusion models [150] and mechanochemical models [140].

We are particularly interested in elucidating the role of mechanical stresses in
dermal wound healing. Although it is possible to incorporate stress and strain into CA models, most mechanical models assume that all species of interest can be represented using continua. Thus, we will focus on mechanochemical models that are expressed as systems of partial differential equations.

Mechanical/mechanochemical modelling was first developed by Murray, Oster and Harris to investigate morphogenesis and other problems in embryology [131, 132, 144]. In particular, these researchers were interested in the hypothesis that mechanical stresses govern pattern formation. Since the development of the Murray-Oster-Harris model in the 1980s, a number of other mechanochemical models of pattern formation have been developed [37, 134]. However, there has been very little activity in this area during the past decade.

In 1992, Tranquillo and Murray [189] extended the work of Murray and Oster to construct a mechanochemical model of dermal wound healing. We outline the Tranquillo-Murray model of dermal wound healing in Section 2.3.1, and we will analyse it in detail in Chapter 3. The Tranquillo-Murray model forms the foundation of all later mechanochemical models of wound healing.

In basic structure, mechanochemical models are very similar to reaction-diffusion models. The only significant difference is that mechanochemical models require additional equations for stress, strain and displacement. In many cases, these are combined to yield one ‘unusual’ equation that is coupled to a conventional reaction-diffusion system. Unfortunately, this ‘unusual’ equation is not of the classical reaction-diffusion form. As such, its inclusion is problematic because it reduces the scope for using mathematical techniques of analysis and it often leads to computational difficulties.

---

12See, for example, the mathematical description of bone remodelling developed by Kawamura et al. [98]. Although it is beyond the scope of the present work, it would be interesting (and novel) to pursue a hybrid mechanical-CA model of dermal wound healing. A possible starting point would be the model developed by Marquez et al. [123] to determine the strain field in a body containing discrete cells that apply tractional forces.

13Another difference is that most mechanochemical models include advection. However, in Chapter 3 we demonstrate that advection is actually unnecessary in a wound healing model.
The strength of mechanochemical models is that they can be used to explore aspects of wound healing that cannot be analysed in any other way. For example, Cook [33] used a mechanochemical model to explore the relationship between the shape of a wound and the shape of the resulting scar. Since reaction-diffusion models do not take the displacement of the ECM into account, it would be impossible to do this without a mechanochemical model.

There are many other aspects of wound healing that demand a mechanochemical model. One particularly interesting problem is pressure bandaging. Since the 1970s, it has been known that applying pressure to a healing wound helps to prevent hypertrophic scarring. However, the reason for the success of pressure bandages is not yet understood [34, 115]. A two-dimensional or three-dimensional model of the mechanical environment of the healing wound would be an important contribution to this study.

2.3.1 The Tranquillo-Murray model

The first and most influential mechanochemical model of dermal wound healing is the model developed by Tranquillo and Murray [189]. Here, we summarise some of its important features. In its simplest form, the Tranquillo-Murray model contains three species: the fibroblast density \( n \), the ECM density \( \rho \) and the ECM displacement \( u \). For a detailed description of the base Tranquillo-Murray model, see Section 3.2.

The equations used for the evolving fibroblast and ECM densities are effectively conventional reaction-diffusion equations. The only difference is that Tranquillo and Murray incorporate passive advection with the moving ECM. However, the equation used to define the ECM displacement is more interesting. Tranquillo and Murray assume that the ECM behaves as a linear viscoelastic solid (i.e. a}

14Tranquillo and Murray [190] also wrote a second paper with almost identical content. Although we generally refer to Tranquillo and Murray [189], either paper could be used as a reference.
Kelvin-Voigt material). In a classical Kelvin-Voigt material, the total stress is expressed as the sum of an elastic stress and a viscous stress. In the Tranquillo-Murray model, a third stress is added: the traction stress.\(^\text{15}\)

The traction stress represents the effect of fibroblasts pulling on the ECM that surrounds them. The traction stress is taken to be zero when no fibroblasts are present and it increases with increasing fibroblast density.\(^\text{16}\) As a result, the regions of the dermis with the greatest cell density tend to undergo contraction.

This has some undesired effects. Initially, the cell density is assumed to be zero inside the wound but positive elsewhere. This means that the forces experienced by the skin initially lead to wound retraction, not wound contraction. In fact, it is possible to show that the base Tranquillo-Murray model can never lead to wound contraction. This is unrealistic, so Tranquillo and Murray proposed a number of extensions to their base model to overcome this problem.

### Extensions of the Tranquillo-Murray model

Specifically, Tranquillo and Murray found that wound contraction can be simulated if an inflammatory mediator is introduced to the model. By making the cell behaviour dependent on the mediator concentration, it is possible to construct models where the tractional stresses inside the wound are ultimately greater than the tractional stresses in the healthy skin. When this is the case, the wound tends to contract rather than retract.

\(^{15}\)This approach to modelling cellular traction forces was adopted from the Murray-Oster-Harris model of morphogenesis [131, 132, 144].

\(^{16}\)Interestingly, Tranquillo and Murray do not use a strictly increasing function to model the dependence of traction stress on fibroblast density. However, it is reasonable to expect that increasing the number of fibroblasts will generally lead to an increase in traction stress. This can be reconciled with the Tranquillo-Murray model by noting that the fibroblast density is subject to physiological constraints; it is physically impossible for the number of fibroblasts in a confined space to increase without bound. Hence, we generally assume that traction stress is an increasing function of fibroblast density throughout the domain of realistic values of fibroblast density.
In the extended Tranquillo-Murray models, the inflammatory mediator is represented using an imposed steady-state concentration profile. We note that this is less realistic than proposing a concentration profile that evolves over time. However, the simplicity of Tranquillo and Murray’s approach permits detailed analysis of the spatially nonhomogeneous steady state. As we will discuss, evolution of the mediator concentration was later introduced by Cook [33] and Olsen et al. [140].

Tranquillo and Murray considered three different effects that an inflammatory mediator might have. Firstly, they considered a model in which the traction stress exerted by the cells is dependent on the mediator concentration (the ‘traction variation’ model). The traction variation model was proposed as a way of representing the effect of myofibroblasts without needing to introduce a second fibroblast species. Importantly, we note that the mediator concentration is assumed to be greater inside the wound than outside it. Thus, the traction variation model effectively augments the traction stress within the wound space, enabling contraction to occur. At equilibrium, the cell density is spatially homogeneous, but the traction stress (and hence the displacement profile) is not.

As an alternative to traction variation, Tranquillo and Murray proposed a ‘growth variation’ model in which the carrying capacity of fibroblasts is dependent on the mediator concentration. In this case, we find that the equilibrium cell density is greatest at the centre of the wound, as is the traction stress. Thus, the growth variation model can also be used to obtain wound contraction.

Thirdly, Tranquillo and Murray proposed a model in which fibroblasts respond to the mediator gradient by chemotaxis. There is a growing body of experimental evidence that supports the view that chemotaxis is important in wound healing (see, for example, Schneider and Haugh [164]). Thus, chemotaxis is a physically

---

17This is not strictly true. As noted above, the expression for traction stress used by Tranquillo and Murray is not a strictly increasing function of cell density. Hence, it is theoretically possible to encounter a situation where the cell density attains its maximum at the wound centre but the traction stress does not. In fact, Tranquillo and Murray include an example of this in their paper, but they do not comment on the fact that this behaviour may be unphysical. In general, the parameters of the growth extension model should be chosen so as to avoid a situation where the increased cell density causes a decrease in traction stress.
realistic way of modelling the response of fibroblasts to an inflammatory mediator. Like the growth variation model, the chemotaxis model leads to an increased concentration of cells inside the wound space and a corresponding increase in traction stress.

All three of these extended models predict wound contraction but they yield subtly different results. Tranquillo and Murray analysed these results, comparing them with the experimental data obtained by McGrath and Simon [126]. Interestingly, they found that the model predictions depend strongly on the function used to represent the response of the cells to the mediator. However, they concluded that there are insufficient experimental data to make any conclusions about whether traction variation, growth variation or chemotaxis is most important.

As well as investigating the effect of an inflammatory mediator, Tranquillo and Murray considered extending their model to include ECM biosynthesis. In this extended model, ECM is synthesised at a rate that is proportional to the net increase of fibroblast density. Tranquillo and Murray point out that this may not be physically realistic; in 1992, very little was known about the rate of ECM turnover. In effect, Tranquillo and Murray used their model to demonstrate one possible approach that could be used to model ECM biosynthesis.

Unfortunately, Tranquillo and Murray’s approach to ECM biosynthesis suffers from a number of deficiencies. Most notably, the Tranquillo-Murray model implies that there are no restrictions on the ECM density profile at equilibrium. In contrast, the later models of dermal wound healing developed by Cook [33] and Olsen and coworkers [136, 140] assume that the ECM density tends to a specified equilibrium value over time. Tranquillo and Murray did not analyse their model of ECM synthesis in detail and we will not consider it further in this thesis.
Some criticisms

Tranquillo and Murray were the first researchers to develop a mechanochemical model of wound healing. As might be expected, their pioneering work suffers from a number of flaws. For example, the Tranquillo-Murray model fails to take into account a number of phenomena that are very important to the mechanics of wound healing. On the other hand, the model includes mathematical representations of some processes that are almost completely irrelevant. In many cases, the restrictive assumptions underlying the Tranquillo-Murray model prevent it from accurately describing the mechanical behaviour of a healing wound. As a result, it is difficult to meaningfully compare the predictions of the model with experimental data.

In their original paper, Tranquillo and Murray outline a number of criticisms of their own model. In particular, they comment on the fact that their model is severely limited by the linear constitutive law that they use to describe the mechanical behaviour of the dermis. Tranquillo and Murray point out that linear viscoelasticity is generally only valid when the displacement gradient is small. However, McGrath and Simon [126] demonstrated experimentally that the displacement of the wound boundary (and hence the displacement gradient) can sometimes be very large. Effectively, the Tranquillo-Murray model can only be used to describe the early stages of fibroplasia (i.e. when contraction is still small).

Although Tranquillo and Murray make some use of McGrath and Simon’s experimental results, a true comparison is not possible. For example, we note that McGrath and Simon observed that wound contraction leads to an exponential decrease in wound area, tending to some limiting size. Mathematically, this

---

18 It is interesting to note that most experimental studies of wound contraction (including McGrath and Simon [126]) depend on data obtained from mice and other small animals. However, such animals heal differently from humans [179]; although human wound contraction is significant, it is not as dramatic as McGrath and Simon’s data would suggest. Linear (visco)elasticity is not ideal, but it is less of a problem than Tranquillo and Murray thought.

19 The exponential wound contraction is preceded by a brief period of wound retraction [126,
can be expressed in the form

\[ A(t) = A_f + (A_0 - A_f) e^{-kt}, \]  

(2.1)

where \( A(t) \) is the wound area at time \( t \), \( A_0 \) is the initial wound area, \( A_f \) is the final wound area and \( k \) is a constant describing the rate of contraction. In one well-reported experiment, McGrath and Simon found that the final wound area was less than one quarter of the original wound area.

Tranquillo and Murray attempted to use McGrath and Simon’s results by comparing their model predictions with an exponential curve of the form given in equation (2.1). However, Tranquillo and Murray were limited by their use of linear viscoelasticity; they could only consider situations in which the final wound area was very close to the original wound area. Despite being able to show that their extended models predict exponential wound contraction, Tranquillo and Murray were not able to produce results that are quantitatively similar to the results obtained by McGrath and Simon. It is unclear as to whether the Tranquillo-Murray model would still predict an exponential rate of contraction if the ratio between the initial wound area and the final wound was more realistic.

Another problem with the Tranquillo-Murray model is that it does not include any mechanism that could account for the permanence of wound contraction. Tranquillo and Murray assume that the dermal ECM behaves as a Kelvin-Voigt viscoelastic solid. In any such model, the displacement will universally tend to zero if the traction stresses are spatially homogeneous. However, this is inconsistent with the clinical observation that mature scars remain contracted despite the apoptosis of most of the fibroblasts and all of the myofibroblasts [46]. In order to accurately model the permanence of wound contraction, the constitutive law for the ECM would need to include some representation of tissue plasticity.

Further difficulties arise from the fact that the Tranquillo-Murray model does not take into account the fact that fibroblasts respond to mechanical changes in their

\[ \text{Reference 189}. \] The extended Tranquillo-Murray models exhibited both retraction and contraction.
environment. Instead, fibroblasts are passively advected with the moving ECM but are otherwise insensitive to the stress and strain that they experience.\textsuperscript{20} This means that there are a number of interesting problems (such as the effectiveness of pressure bandages for preventing hypertrophic scars) that cannot be investigated using the Tranquillo-Murray model.

Tranquillo and Murray mention a number of other perceived deficiencies in their model. In addition to the issues described above, they discuss the fact that their model does not account for the mechanical anisotropy of the wound environment. Similarly, they acknowledge that their representation of the inflammatory mediator could be improved and they suggest that orthokinesis (the tendency of fibroblasts and other cells to follow each other as they migrate) may need to be included in an expanded model of fibroblast movement. Many of these issues have been addressed in later models.

\subsection*{2.3.2 Other mechanochemical models}

The Tranquillo-Murray model was a pioneering attempt to incorporate mechanical effects into a reaction-diffusion model of dermal wound healing. Importantly, Tranquillo and Murray demonstrated that a purely ‘mechanical’ model is not sufficient; in order to simulate wound contraction, they found it necessary to construct a ‘mechanochemical’ model. Without an inflammatory mediator, the Tranquillo-Murray model predicts wound retraction but not wound contraction.

In the years that followed the publication of Tranquillo and Murray’s work, a number of extended models were proposed. In particular, Cook [33] and Olsen [136] constructed mechanochemical models of wound healing during their doctoral studies.\textsuperscript{21} Cook and Olsen worked independently and they focused on dif-

\textsuperscript{20}This leads to a number of problems. In Chapter 3 we demonstrate that the effects of advective terms are negligible compared to the other terms of the Tranquillo-Murray model. Effectively, it is possible to decouple the mechanical equation from the rest of the Tranquillo-Murray model without significantly changing the model predictions.

\textsuperscript{21}Olsen and coworkers have written several papers that describe their mechanochemical mod-
ferent aspects of wound healing. Cook concentrated on the effects of mechanical anisotropy while Olsen’s model emphasises the role of growth factors.

After more than a decade, Cook’s mechanochemical model remains unsurpassed in the current literature. Unlike his contemporaries, Cook accounted for the fibrous nature of the ECM by constructing a tensorial description of fibre alignment. This tensor is used in two different ways: fibre alignment is assumed to affect the migration of cells as well as the anisotropic response of the ECM to applied stresses.

Cook’s model also takes into account the plasticity of the dermal ECM; Cook was one of the first researchers to use a multiplicative decomposition of the deformation gradient to describe the mechanical remodelling of a biological tissue. One of the strengths of Cook’s model is that it is able to describe the permanence of wound contracture by modelling the evolving residual strain. We will use a similar approach to ECM plasticity in the one-dimensional wound healing model that we propose in Section 6.4.

Interestingly, Cook chose to omit viscosity from his model of dermal mechanics. Cook contended that much of the ‘viscous’ behaviour observed in dermal wound healing is actually related to matrix turnover. As we discuss in Section 3.4, this proposal is borne out by recent experimental results, which indicate that the viscous behaviour of the dermis is negligible on the time scale of wound healing. Hence, it is more appropriate to use a morphoelastic model to describe dermal mechanics than it is to use a viscoelastic model.

Although Cook’s approach to dermal mechanics is very thorough, there are a few problems with his model. Most notably, Cook decided to omit the inflammatory elements of wound healing [137, 138, 139, 140, 141]. Unfortunately, Cook’s work is largely unpublished; it is necessary to consult Cook’s thesis [33] for a detailed description of his models. An outline of Cook’s work can be found in Murray’s Mathematical Biology [129].

22In Chapter 5 of his thesis, Cook [33] develops a comprehensive theory that can be used to relate stress and strain in a fibrous tissue. This theory incorporates a model of fibre undulation as well as a model of fibre alignment.

23See Section 2.4 for the history of this approach to plasticity.
mediator that Tranquillo and Murray concluded was necessary. In order to avoid a situation where wound contraction never occurs, Cook assumed that the traction stress is a biphasic function of fibroblast cell density. This function was defined so that the fibroblasts exert no stress when they are at their carrying capacity. In contrast, in vitro experiments indicate that the traction stress is a monotone increasing function of cell density (see, for example, Tamariz and Grinnell [183]).

Unlike Cook, Olsen retained the linear viscoelastic model of dermal mechanics proposed by Tranquillo and Murray. However, Olsen developed a more realistic model of the cell-chemical interactions that take place during wound healing. We recall that Tranquillo and Murray proposed models in which the concentration of inflammatory mediator is specified at every point. In contrast, Olsen’s model includes a freely-diffusing growth factor that is produced and eliminated at rates that are dependent on the local cell density. This growth factor affects the behaviour of cells in a variety of different ways, ranging from encouraging proliferation to governing the rate of collagen synthesis.

Another interesting aspect of Olsen’s work is that he includes fibroblasts and myofibroblasts as separate model species. In his models, fibroblasts differentiate into myofibroblasts at a rate that depends on the local growth factor concentration but not on the local stress. Myofibroblast behaviour is quite different from fibroblast behaviour in Olsen’s model; although fibroblasts are treated as motile cells, Olsen assumes that myofibroblasts are fixed to the underlying ECM (i.e. they are passively advected but are otherwise stationary). Furthermore, the myofibroblast contribution to the total traction stress takes a different algebraic form from the fibroblast contribution.

However, Olsen’s model also suffers from some flaws. In particular, we note that Olsen’s model includes a large number of unknown parameters that have not been measured experimentally (or cannot be measured experimentally). It is difficult

\[24\] In this sense, Cook’s model is ‘mechanical’ rather than ‘mechanochemical’. However, we will continue to use the term ‘mechanochemical’ to refer to all models that combine a system of reaction-diffusion equations with a mechanical description of the ECM.
to justify the complicated expressions that are used throughout the models when so few experimental data exist. Furthermore, Olsen fails to make any significant use of the ‘mechanical’ aspects of his models. As we discuss in Chapter 3, the force-balance equation of the Tranquillo-Murray model can be decoupled from the other equations without having any significant effect on the model predictions. Since Olsen’s model is also limited to small displacements and Olsen assumes that cells are unaffected by their mechanical environment, we expect that it will decouple in a similar manner.

The Cook model and the Olsen model are complementary. The strengths of Cook’s mechanical model are balanced by the strengths of Olsen’s growth factor model. We will incorporate aspects of both models in our own model of dermal wound healing.

Apart from Cook [33] and Olsen [136, 137, 138, 139, 140, 141], very little has been published on the mechanochemical modelling of dermal wound healing. The only other significant work in this field is the visco-elasto-plastic model proposed by Tracqui et al. [187]. This model is notable because it includes two different species of ECM: a provisional fibrin matrix and a mature collagenous matrix. As fibroblasts invade the wound space, they degrade the fibrin clot and replace it with mature tissue. Also, Tracqui et al. account for the changing mechanical structure of the healing wound using a strain-hardening approach that is analogous to classical plasticity. Although this approach can be used in wound healing without any problems, we prefer to use Cook’s multiplicative decomposition because it can be interpreted more easily and it can be extended to other biomechanical problems like soft tissue growth.

2.3.3 Models of lattice contraction

Despite the decade-long lack of progress on the mechanochemical modelling of dermal wound healing, related modelling techniques have been used to describe
the contraction of fibroblast populated collagen lattices (FPCLs). In this section, we briefly outline some models of lattice contraction that are similar to the models of dermal wound healing described above.

The earliest model of FPCL contraction was developed by Moon and Tranquillo [128], who adapted the Tranquillo-Murray model of dermal wound healing to describe the contraction of a collagen microsphere. Despite the fact that the deformations predicted are significant, this model uses the same linear constitutive law as the Tranquillo-Murray model. Furthermore, the Moon-Tranquillo model is unable to account for the permanent contraction observed experimentally (see, for example, Guidry and Grinnell [82, 81]).

Barocas et al. [11] modified the Moon-Tranquillo model by replacing the Kelvin-Voigt constitutive law with a Maxwell constitutive law. Like the Kelvin-Voigt model of viscoelasticity, the Maxwell model is only valid for small displacement gradients. However, a Maxwell model can be used to account for the permanence of matrix contraction. Thus, the Barocas et al. model of FPCL contraction is a significant improvement on the earlier Moon-Tranquillo model.

Barocas and Tranquillo [12, 13] also developed a biphasic model of collagen lattice contraction. In this model, the interaction between the fibrous lattice and the permeating fluid medium is described using the theory of mixtures. Similarly to Cook [33], Barocas and Tranquillo use a tensor to describe the distribution of fibre orientations. The Barocas-Tranquillo model of lattice contraction is a very advanced model that takes into account a variety of important effects, including the partial expansion of collagen lattices after cell traction stresses are removed.

Since the development of the Barocas-Tranquillo model, Barocas and coworkers have used the theory of mixtures and other mathematical techniques to describe a variety of experiments on collagen lattices [23, 24, 66, 103, 166]. Most of these models are very different in structure from the earlier mechanochemical models.

25 Interestingly, it is possible to show that the Maxwell constitutive law is a special case of the theory of morphoelasticity that we develop in Chapter 5. See Section 6.2 for more details.
of wound healing. However, a number of common features are retained, such as the inclusion of fibroblast traction as an additive contribution to the total stress.

Another interesting model of FPCL contraction is the model developed by Ferrenq et al. [58]. This model is very similar to the Moon-Tranquillo model but it contains some interesting subtleties. As with the Moon-Tranquillo model, a linear Kelvin-Voigt constitutive law is used; however, Ferrenq et al. are careful to ensure that their model is restricted to situations in which the displacement gradient is small. One unusual feature of this model is that the cell traction stress is explicitly dependent on time. This is used to explain the experimentally observed lag between the introduction of cells and the commencement of contraction.

Recently, Ramtani [155, 156] developed a model of FPCL contraction that incorporates a repulsive interaction force between cells. This represents the inhibition of gap junction formation that occurs at large cell densities. It would be interesting to compare the relative importance of cell-cell repulsion and contact-inhibition of cell traction stress.

Marquez, Zahalak and coworkers [121, 123, 122, 154, 198] have also made a recent contribution to the mathematical modelling of lattice contraction by fibroblasts. Like Barocas and coworkers, this group has used its mechanical theories to model a variety of rheological experiments that are performed on collagen lattices. However, their modelling approach is very different from that introduced by Tranquillo and Murray. Most interestingly, they model the mechanical effect of cells by considering individual cells and using Eshelby’s solution to describe the local strain fields.

Although the mechanochemical modelling of dermal wound healing has apparently stagnated, the study of collagen lattice mechanics has remained an active field. Many of the innovations introduced in models of FPCL contraction could potentially be incorporated into models of dermal wound healing. In particular, it would be interesting to develop a model of wound healing that uses the biphasic theory proposed by Barocas and coworkers and/or the cellular strain fields.
described by Marquez, Zahalak and coworkers.

2.4 Mathematical approaches to growth and remodelling

From a mechanical perspective, the dermis is a complex and anisotropic material that exhibits viscous, elastic and plastic behaviours [62, 171]. This means that it is very difficult to construct a mechanochemical model of dermal wound healing that accounts for the subtleties of dermal mechanics. As described above, recent advances in modelling FPCL contraction have yielded practical results that can be applied to wound healing models. Moreover, there is a wide range of relevant mathematical techniques that have been developed to describe the complex mechanical behaviour of biological materials. In this section, we describe some general biomechanical theories that are relevant to our further investigations.

One of the pioneers of biomechanical modelling was Fung [61, 62], who developed a generalised theory of biological elasticity. Although more advanced models exist, Fung elasticity remains an important concept in biomechanics. Fung also introduced the idea of an ‘incompatible zero stress state’ as a way to describe residual strain. We cover this in greater detail in Chapter 4.\(^{26}\)

Since Fung’s early work, considerable advances have been made in the field of biomechanics. However, many of these advances have yet to impact on the mechanochemical modelling of dermal wound healing. As we describe above, the most mechanically sophisticated model of dermal wound healing is the model developed by Cook [33]. Amongst his other achievements, Cook incorporated a tensorial representation of the evolving residual strain into his model.

\(^{26}\)Although Fung was the first to consider the concept of a zero stress state, it is interesting to note that he did not develop the multiplicative decomposition of the deformation gradient that we now use to quantify the zero stress state. For further details of Fung’s early work on viscoelasticity and residual strain, see Fung [61, 62] and the recent review of biomechanical theories by Humphrey [96].
However, Cook was not the first to develop this formulation. In the late 1960s, Lee [107] proposed a multiplicative decomposition of the deformation gradient tensor that is mathematically equivalent to the model used by Cook. In Lee’s formulation, the deformation from the initial state to the current state is separated into a plastic deformation, $F_p$, and an elastic deformation, $F_e$. Expressed mathematically, this decomposition takes the form

$$F = F_e F_p,$$

where $F$ is the deformation gradient tensor and $F_e$ and $F_p$ are also second-order tensors. The plastic deformation gradient, $F_p$, is defined in terms of a deformation from the initial configuration of the body to a configuration where it is unstressed, while the elastic deformation gradient, $F_e$, is defined using a deformation from the zero stress state to the current state. Thus, $F_p$ can be used to represent residual strain (or in biological models, the net growth) and $F_e$ can be associated with elastic stress and related concepts such as work energy.

As described above, Lee’s formulation depends on the concept of a ‘zero stress state’. This idea is central to much of the work that follows (especially in Chapters 4 and 5) and it deserves further introduction. Essentially, the zero stress state is a locally-defined representation of the deformation required to relieve all elastic stresses at a given point on the interior of a solid body. As discussed in Chapter 4, this leads naturally to a representation of the zero stress state using a hypothetical deformation gradient.

It must be emphasised that the zero stress state is always defined locally rather than globally. Thus, equation (2.2) is applied independently at every point in the body being considered. Although we will refer to ‘the zero stress state of a body’ throughout this thesis, this is difficult to define unambiguously. The reason for this is that it is not generally possible to construct a global configuration of a body in which all points in the body are simultaneously stress-free. Hence, any definition of the zero stress state of a body based on a global stress-relieving configuration will be ineffective.
For any given point, however, we expect that it is always possible to find a family of global configurations of the body where the point of interest is at zero stress. As discussed in Section 4.3.2, any configuration from this family can be used to define a local tensorial representation of the zero stress state around the given point. This gives us a clearly-defined way of constructing the local zero stress state at every point in the body. Thus, ‘the zero stress state of a body’ can be thought of as a simple way of referring to all of these local zero stress states simultaneously. Despite the possible ambiguity involved, this leads to a clean presentation of the work that follows.

In addition to the difficulties that arise because of the impossibility of defining a global stress-relieving configuration, further problems are introduced when trying to uniquely define a representation of the zero stress state. As described in Section 4.3.2, following a deformation by a global rotation leaves the stress distribution throughout a body unchanged. This has significant consequences for the common practice of using a local deformation gradient to represent the zero stress state. Specifically, we find that any representation of the zero stress state as a Lagrangian deformation gradient can be premultiplied by a proper orthogonal tensor to obtain an equally valid representation of the zero stress state. The effects of this are described in more detail in Chapter 4.

There are four ways of dealing with this problem of non-uniqueness. One approach would be to use local Cauchy-Green tensors instead of local deformation gradient tensors to represent the zero stress state. However, this would cause us to lose the useful mathematical properties of Lee’s multiplicative decomposition of the deformation gradient. Equivalently, a second approach for imposing uniqueness would be to represent the zero stress state using an element of the quotient ring $\mathbb{T}/\mathbb{P}$, where $\mathbb{T}$ is the set of second-order tensors and $\mathbb{P}$ is the set of proper orthogonal second-order tensors. Although this approach leads to some mathematically elegant definitions, it also makes it more difficult to construct...
evolution equations for morphoelasticity. It will not be considered further.

A third approach is to construct equations that are valid regardless of the choice of representation for the zero stress state. Although this approach leads to the most useful and easily interpretable models, it also requires us to be very careful when changing coordinate systems. We will make extensive use of this approach in the work that follows. Lastly, it is possible to choose a unique representation of the zero stress state by introducing some arbitrary additional constraints to ensure uniqueness. As described in Section 4.4.1, there is a physically-motivated way of doing this that leads to a number of useful results. We will also use this 'principal zero stress deformation gradient' for developing definitions of infinitesimal strain.

The problems caused by the impossibility of constructing a global stress-relieving deformation and the further problems surrounding the issue of the uniqueness mean that using Lee’s multiplicative decomposition of the deformation gradient is not always straightforward. For example, Xiao et al. [196] point out that these problems combine to make it very difficult to define the rate of morphoelastic deformation unambiguously. Great care is needed throughout the mathematical developments in Chapter 5 to ensure that these issues are dealt with appropriately. Despite these issues, however, the fact that the zero stress state is a physically intuitive concept with obvious applications in biomechanical research means that it is desirable for us to address these issues directly rather than attempting to find an alternative way of describing morphoelastic growth.

Lee’s original work was developed with the intention of modelling the elastoplastic behaviour of engineering materials at large strains. To this day, most of the theoretical work relating to the multiplicative decomposition of the deformation gradient has focused on engineering applications (for further details, see especially Lubarda [111] and Xiao et al. [196]). A full review of the use of Lee’s multiplicative decomposition in elastoplasticity theory is beyond the scope of this thesis.

The multiplicative decomposition of the deformation gradient was first introduced
to biomechanics by Rodriguez et al. in 1994 [159]. Similarly to Cook, Rodriguez 
et al. used the multiplicative decomposition of the deformation gradient as a 
framework for investigating the residual stresses that arise in biological tissues. Additionally, they proposed a time-dependent model of growth to investigate how the stresses in a soft tissue evolve over time.

Soft tissue growth is a good ‘testing ground’ for biomechanical models because of the growing body of experimental data that can be compared with the predictions of mathematical models. In particular, various researchers have studied the mechanical behaviour of growing tumour spheroids [8, 88, 160]. Recently, mathematicians have begun to use these experimental results to inform their models [5, 9, 10, 114].

Interestingly, there are a number of issues surrounding soft tissue growth that are yet to be resolved. For example, it has been observed that residual stresses develop during growth, but very little is known about the fundamental mechanisms of stress generation [96]. Advances in mathematical modelling have the potential to clarify many of the questions surrounding the mechanics of growth as improved theoretical frameworks are developed. Ultimately, mathematical models are essential for developing a flexible, quantitative and predictive theory of biological growth.

In Chapters 4 and 5, we will investigate the mathematical model of growth developed by Rodriguez et al. [159] and extended by various modellers. As we will see, there are a number of difficulties with Rodriguez et al.’s approach to modelling the evolving zero stress state. Their definition of elastic strain is problematic and we find that there are fundamental inconsistencies in their approach to stress-dependent growth. However, these are relatively minor issues; in our in-

---

28Interestingly, it appears that Rodriguez et al. and Cook developed their models of biological plasticity independently of the existing engineering literature (and each other). The first reference to Lee’s paper in the biomechanical literature appeared as recently as 2002 [113].

29Soft tissue growth is not the only biomechanical application of Lee’s multiplicative decomposition of the deformation gradient. For example, a number of researchers have used this approach to describe the mechanical stresses within blood vessels and the development of aneurysms [95, 96, 192].
vestigations of growth and remodelling, we essentially adopt the same structure as Rodriguez et al.’s model.

In developing their theory of stress-dependent growth, Rodriguez et al. built on the earlier models proposed by Skalak and coworkers [173, 174]. However, Skalak and coworkers focused most of their attention on modelling surface growth \textit{(i.e.} growth by accretion that leaves the zero stress state relatively unchanged). Although some forms of growth can be effectively modelled in this manner, we will concentrate on bulk changes to the zero stress state rather than surface accretion.

Rodriguez et al. also drew on theoretical studies of the zero stress state by Hoger and coworkers [91, 92, 93]. Hoger remains one of the leading theoreticians in the field of continuum biomechanics. She and her coworkers have published a number of important papers that deal with the multiplicative decomposition of the deformation gradient and its application to biological remodelling (see especially [26, 94, 102, 113, 175]).

Some of the important results obtained by this group are as follows:

- Hoger [91, 92] developed a formal constitutive theory for residually-stressed hyperelastic materials. This theory was proposed in two different forms; a general theory for arbitrary deformations [91] and a simpler theory that can be used when the apparent strain is small \textit{(i.e.} when the initial state is close to the current state) [92].

- Skalak \textit{et al.} [175] showed how residual stresses can arise when the zero stress state is incompatible with any real deformation. They demonstrated the use of a number of different measures of incompatibility, emphasising the need for care when considering multiply connected bodies \textit{(e.g.} blood vessels and other tubes).

- Chen and Hoger [26] constructed a firm theoretical basis for zero stress state theory by developing the concept of an equivalence transformation. They also demonstrated how their theory could be used to construct a model of
a growing spherical shell.

- Klisch, Van Dyke and Hoger [102] used the work energy principle and other thermodynamic arguments to construct a model of the volumetric growth of a spherical shell. Similarly to the later work by Goriely and Ben Amar [70], Klisch et al. model growth by considering accumulated discrete changes to the zero stress state.\(^{30}\)

- Lubarda and Hoger [113] give a very detailed theoretical investigation of zero stress state theory applied to growth. Returning to thermodynamic principles, they derive a general constitutive framework for modelling the changing stresses and strains within a growing material.

- Hoger et al. [94] extended their previous work on small deformations to develop a mathematical framework that can be used when the apparent strain is small, but the initial condition is residually stressed. This forms the theoretical foundation for the recent model of cumulative growth developed by Goriely and Ben Amar [70].

Interestingly, the approach taken by Hoger and coworkers is distinctly different from the approach that we will take in Chapters 4 and 5. Chen and Hoger [26] and Hoger et al. [94] focus on the case where the current stress distribution is known and small changes are made to the current configuration. Lubarda and Hoger [113] use formal strain energy arguments to develop a general constitutive law for a hyperelastic material with a changing zero stress state.

In contrast, our investigations will avoid the concepts of ‘strain energy’ and ‘response function’\(^{31}\) altogether. Indeed, a thermodynamic approach to biological morphoelasticity may not be appropriate because biological tissues are not always energetically passive. This unusual situation is a result of the fact that cells can

\(^{30}\)A particularly interesting feature of this work is that Klisch et al. investigate a few forms of directed growth. However, they do not consider the case where the rate of growth is stress-dependent.

\(^{31}\)The response function was introduced by Hoger and coworkers to describe how the local stress changes in response to a reversible deformation away from the current state. For more details, see especially Chen and Hoger [26].
use energy gained from metabolism to actively work on the mechanical structure of the tissue. Thus, strain energy is only a useful concept on time-scales short enough for cellular activity to be negligible. Although it will be appropriate to use classical theories of elasticity (linear or nonlinear) to instantaneously relate stress and strain, we will approach morphoelasticity from the perspective of modelling the time-evolution of strain rather than considering energetic changes.

Hoger and coworkers are not the only biomechanical researchers to have investigated the zero stress state. In recent years, Ambrosi and Guana [3], Doblaré and García-Aznar [47], Goriely and Ben Amar [18, 69, 70], Loret and Simões [110] and Menzel [127] have all explored the use of an evolving zero stress state to model biological growth. In general, these models use a thermodynamic approach that is similar to the approach of Lubarda and Hoger [113]. However, some of the methods developed by Goriely and Ben Amar are similar to ours. We will consider their approach to morphoelasticity in greater detail in Section 5.5.3.

Mathematical modellers have also used zero stress state theory to describe other biological phenomena. Most notably, there is a significant body of work on cardiovascular remodelling (see, for example, the work of Taber and coworkers [181, 182]). We will not discuss this work in detail, but we note that there are many similarities between models of growth and models of vascular remodelling.

Lastly, it is interesting to note that it is possible to model the mechanics of growth and remodelling without introducing a multiplicative decomposition of the deformation gradient. Drozdov and Khanina [51], for example, proposed a model of volumetric growth in which growth is directly related to mechanical work. This approach enabled Drozdov and Khanina to avoid the mathematical challenges involved in clearly defining the zero stress state and the way in which it changes. However, their model is very complex and difficult to interpret.

Although care is needed when using a tensorial representation of the zero stress state, this method has a number of advantages. Most importantly, the zero stress state is a physically intuitive concept and, unlike strain energy, it can be
directly measured experimentally (albeit using methods that destroy the body being studied). Thus, we will exclusively focus on representations of growth and remodelling that use a tensorial description of the zero stress state.

2.5 The present work in context

The mechanical interaction of cells with the ECM is a biologically important process that can be modelled mathematically. In dermal wound healing, for example, it is known that fibroblasts and myofibroblasts exert traction stresses on the fibres of the ECM, leading to wound contraction [45, 55, 186]. Furthermore, fibroblasts and myofibroblasts respond to their mechanical environment by changing their behaviour; \textit{in vitro} experiments such as those performed by Grinnell and coworkers [75, 78, 79, 183] indicate that mechanical signals influence the differentiation of fibroblasts into myofibroblasts.

These changes at the cellular level can have clinically significant consequences. Aarabi \textit{et al.} [1] demonstrated that mechanical loading can cause murine wounds to develop into hypertrophic scars. Similarly, pressure bandages are known to reduce hypertrophic scarring, possibly as a result of changes to the stress distribution within the healing wound [34, 115].

Some mathematical models of dermal wound healing have been developed that take into account the mechanical response of the dermis. Most notably, Tranquillo and Murray [189] pioneered the use of ‘mechanochemical’ models to describe wound healing. However, the Tranquillo-Murray model is unable to replicate important aspects of dermal mechanics and many of the subsequent mechanochemical models are similarly flawed; for example, none of the present models can account for the response of fibroblasts to mechanical signals.

One of the difficulties with developing a mechanochemical model of dermal wound healing is that the dermis exhibits complex, time-varying mechanical behaviour
Cook [33] made several significant advances towards dealing with these problems; he used a tensorial description of fibre alignment to account for mechanical anisotropy and he introduced the concept of an evolving ‘zero stress state’ to model wound plasticity. Interestingly, Cook was not the first to use an evolving zero stress state in this manner. Lee [107] originally proposed an equivalent theory to describe metal plasticity and Rodriguez et al. [159] used the same approach to describe soft tissue growth.

Since Cook proposed his model of dermal wound healing, significant advances have been made in zero stress state theory (see, for example, Lubarda and Hoger [113]). However, these recent insights have not yet been used in a wound healing model. One reason for this is that the present theories are too complex to incorporate into a practical model.

Given the recent lack of progress in modelling the mechanics of wound healing, it seems appropriate to revisit earlier mechanochemical models like the Tranquillo-Murray model. In Chapter 3, we analyse the Tranquillo-Murray model in detail; the results of our analysis will be used to guide the model of wound healing that we propose in Chapter 6.

However, the focus of this thesis is on developing a simple and interpretable theory of morphoelasticity that formalises and unifies previous approaches to biological plasticity. In Chapters 4 and 5 we use zero stress state theory to construct a model of the evolving strain experienced by a growing or contracting tissue. Together with an appropriate model of tissue anisotropy, this could be used to develop an advanced mechanochemical model of dermal wound healing. Moreover, our theory can be applied to a variety of problems, some of which are described in Chapter 6.
Chapter 3

Analysis of the
Tranquillo-Murray model of
dermal wound healing

3.1 Introduction

In the past decade, much progress has been made in mathematically modelling
the mechanics of biological tissues. However, these advances have not been re-
lected in improved mechanical models of dermal wound healing. As described
in the previous chapter, the models constructed in the mid-1990s by Cook [33],
Olsen et al. [136, 137, 138, 139, 140, 141] and Tracqui et al. [187] have not yet
been extended or modified in the light of improved mathematical frameworks for
biomechanics.

Thus, it remains pertinent to analyse the assumptions that underlie the early
mechanochemical models of dermal wound healing. In particular, it is useful to
analyse the 1992 Tranquillo-Murray model [189]; this was the first mechanochem-
ical model of dermal wound healing, and more recent models have generally retained the same fundamental structure.\footnote{An exception to this is Cook’s model \cite{33}, in which the constitutive assumptions were substantially reworked from first principles.} Furthermore, the Tranquillo-Murray model is closely related to mechanical models of other biological phenomena, such as the contraction of collagen lattices \cite{128, 155, 156} and mesenchymal morphogenesis \cite{37}.

Interestingly, although Tranquillo and Murray \cite{189} (and later Olsen \textit{et al.} \cite{137, 138, 140, 141}) performed numeric simulations and steady state analysis of the models that they had constructed, their published work contains little other mathematical analysis. Specifically, there has not yet been any detailed mathematical analysis of the time-dependent behaviour of models that use the Tranquillo-Murray paradigm.

In this chapter, we present our own mathematical analysis of the Tranquillo-Murray model, focusing on the effects of the constitutive assumptions underlying the model. We begin in Section 3.2 with a description of the Tranquillo-Murray model as it was originally presented. In this section, we outline the equations of the Tranquillo-Murray model, replicate some of Tranquillo and Murray’s original results and perform some preliminary mathematical analysis of the model.

Our investigation continues in Section 3.3, where we use perturbation methods to analyse a caricature Tranquillo-Murray model. These techniques enable us to explain oscillations in the displacement profile that are observed in numeric simulations. Displacement oscillations do not seem to be physically realistic but we are able to show that they are an unavoidable consequence of using a viscoelastic model.

Interestingly, displacement oscillations can be avoided by removing the viscous term from the force balance equation. In Section 3.4, we discuss the fact that experimental results suggest that the dermis behaves as an elastic solid on the time scale of wound healing. Thus, it is appropriate for us to use a purely elastic
model (or an elastoplastic model) instead of a viscoelastic model. Unfortunately, removing viscosity introduces some numeric and analytic difficulties, which we discuss in Section 3.4.2.

Finally, in Section 3.5 we conclude this chapter by summarising the results of our analysis and discussing how these results will influence the model of dermal wound healing that we propose in Chapter 6.

### 3.2 Outline of the Tranquillo-Murray model

#### 3.2.1 Description of the model

The base model proposed by Tranquillo and Murray consists of three species: the fibroblast population density \( n \), the ECM density \( \rho \) and the ECM displacement \( u \). Since \( n \) and \( \rho \) are both representative of densities, it is assumed that they satisfy continuity equations. The third equation in the Tranquillo-Murray model is a force-balance equation; this is used to describe the evolution of \( u \).

Tranquillo and Murray present their model in a general form that could potentially be used to describe a wound of any shape. However, their paper focuses on the implementation of this model in a one-dimensional Cartesian coordinate system.\(^3\) That is, Tranquillo and Murray concentrate on modelling a narrow rectangular wound with long, parallel boundaries (as shown in Figure 3.1). In our analysis, we will also restrict ourselves to considering a rectangular wound.

After nondimensionalisation of their Cartesian model, Tranquillo and Murray

\(^2\)For simplicity, the fibroblast population density is often referred to as the cell density.

\(^3\)Tranquillo and Murray also state a system of equations that could be used to model a wound with radial symmetry. However, they do not present any simulations or analysis of this model. We will not consider radially-symmetric wounds in this chapter.
Figure 3.1: Tranquillo and Murray focused on modelling narrow rectangular wounds like incisions. The figure on the left represents the view of a narrow rectangular wound from above, while the figure on the right illustrates a cross-section of the wound. For a rectangular wound of this form, it is appropriate to use the perpendicular distance from the central axis of the wound as a spatial coordinate.

obtained the following system of equations:

\[ \frac{\partial n}{\partial t} + \frac{\partial}{\partial x} \left( n \frac{\partial u}{\partial t} \right) = \frac{\partial^2 n}{\partial x^2} + r_0 n \left( 1 - n \right), \]  
\( (3.1) \)

\[ \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x} \left( \rho \frac{\partial u}{\partial t} \right) = 0, \]  
\( (3.2) \)

\[ \mu \frac{\partial^3 u}{\partial x^2 \partial t} + \frac{\partial^2 u}{\partial x^2} + \frac{\partial}{\partial x} \left( \frac{\tau_0 \rho n}{1 + \gamma n^2} \right) = s \rho u. \]  
\( (3.3) \)

In these equations, \( x \) is a dimensionless measure of the distance from the wound centre and \( t \) is a dimensionless measure of time elapsed since wounding. The variable \( x \) is nondimensionalised so that the wound boundary is located at \( x = 1 \). The physical interpretations of the constant parameters (\( r_0, \mu, \tau_0, \gamma \) and \( s \)) are discussed later in this section.

Equations (3.1) to (3.3) are solved subject to initial conditions

\[ n(x, 0) = \begin{cases} 0 & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}, \]  
\( (3.4) \)

\[ \rho(x, 0) = 1, \]  
\( (3.5) \)

\[ u(x, 0) = 0; \]  
\( (3.6) \)

\(^{4}\text{These equations are labelled (8a) to (8c) in Tranquillo and Murray [189].}\)
and boundary conditions

\begin{align}
\frac{\partial}{\partial x}(0, t) &= 0 \\
\frac{\partial}{\partial x}(\infty, t) &= 0, \\
\rho(\infty, t) &= 0, \\
\frac{\partial}{\partial x}(\infty, t) &= 0.
\end{align}

Throughout the rest of this section, we discuss the dimensionless equations used in the Cartesian implementation of the Tranquillo-Murray model. For further details (e.g. for a description of the two-dimensional model), see Tranquillo and Murray [189].

**Mass conservation equations**

Equations (3.1) and (3.2) are continuity equations that describe the evolution of the fibroblast and ECM densities. Importantly, we note that passive advection is incorporated into the total flux. This is achieved by using expressions of the form

\[ \frac{\partial \phi}{\partial t} + \frac{\partial}{\partial x}(\phi v), \]

where \( \phi \) is \( n \) or \( \rho \) and \( v \) represents the velocity of the underlying substrate \( (i.e. \text{the ECM}) \). For small displacement gradients, \( v \) can be approximated by \( \frac{\partial u}{\partial t} \), yielding the left hand sides of equations (3.1) and (3.2).

As we see from equation (3.2), the ECM is passively advected but there is no ECM synthesis or degradation \( (i.e. \text{the mass of ECM is conserved}) \). In contrast, equation (3.1) shows that the fibroblast density evolves due to the random motion of the cells \( (\text{represented by Fickian-style diffusion}) \) and by proliferation and death \( (\text{represented by a logistic growth term}) \). Thus, (3.1) is effectively a modified form of Fisher’s equation that includes advection.
It should be noted that the fibroblast density has been nondimensionalised to have a carrying capacity of unity and that time has been nondimensionalised according to the characteristic time of fibroblast ‘diffusion’. Hence, the only dimensionless parameter in equation (3.1) is $r_0$, which represents the maximum rate of fibroblast proliferation relative to the rate of random movement.

Initially, it is assumed that there are no fibroblasts inside the wound but they are at their carrying capacity outside the wound; this yields condition (3.4). In contrast, the initial ECM density is assumed to be uniform. After nondimensionalisation, this yields condition (3.5).

The Tranquillo-Murray model is defined on a semi-infinite domain with boundary conditions applied at $x = 0$ (i.e. the centre of the wound) and in the limit as $x$ tends to infinity. Model symmetry implies that there is no fibroblast flux at $x = 0$, which yields condition (3.7a). Furthermore, the fibroblast population density and the ECM density are assumed to be undisturbed far from the wound; that is, $\frac{\partial n}{\partial x}$ and $\frac{\partial \rho}{\partial x}$ tend to 0 as $x$ grows large. There are alternatives to these boundary conditions that we discuss in Section 3.2.2.

**ECM mechanics**

Neglecting inertial effects and using a Kelvin-Voigt viscoelastic model, the mechanical response of the ECM is represented using an equation of the form

$$\nabla \cdot (\sigma_{\text{viscous}} + \sigma_{\text{elastic}} + \psi) = b,$$

where $\sigma_{\text{viscous}}$ is the viscous contribution to the stress tensor, $\sigma_{\text{elastic}}$ is the elastic contribution to the stress tensor, $\psi$ is the cell traction stress and $b$ is a vector representing the body force.

---

$^5$ The limit $x \to \infty$ is physically interpreted as representing an arbitrary point far away from the wound.
In one Cartesian dimension, we can effectively replace tensors and vectors with scalar equivalents. This yields the following:
\[
\frac{\partial}{\partial x} (\sigma_{\text{viscous}} + \sigma_{\text{elastic}} + \psi) = b. \tag{3.10}
\]

Using a Newtonian constitutive law for the viscous stress, we find that
\[
\sigma_{\text{viscous}} = \mu \frac{\partial v}{\partial x},
\]
where \(v\) is the velocity and \(\mu\) is the effective coefficient of viscosity.\(^6\) In the case where the displacement gradient is small (i.e., \(\frac{\partial u}{\partial x} \rightarrow 0\)), we can approximate velocity by \(\frac{\partial u}{\partial t}\). Hence,
\[
\sigma_{\text{viscous}} = \mu \frac{\partial^2 u}{\partial x \partial t}.
\]

Similarly, we can use a Hookean constitutive law to specify the elastic stress. This yields
\[
\sigma_{\text{elastic}} = k \varepsilon,
\]
where \(\varepsilon\) is the strain and \(k\) is the effective coefficient of elasticity. In the case where the displacement gradient is small, we can approximate strain by \(\frac{\partial u}{\partial x}\). This yields
\[
\sigma_{\text{elastic}} = k \frac{\partial u}{\partial x}.
\]

By using Newtonian viscosity and Hookean elasticity, Tranquillo and Murray effectively assume that the ECM behaves isotropically. However, it is well known that this is not the case; Cook [33] and later models have considered the mechanical anisotropy of the human skin in greater detail.

\(^6\)The effective coefficient of viscosity can be expressed in terms of the bulk viscosity and the shear viscosity. Similarly, the effective coefficient of elasticity can be expressed in terms of Young’s modulus and Poisson’s ratio.
The cell traction term used by Tranquillo and Murray takes the form

$$\psi = \frac{\tau_0 \rho n}{1 + \gamma n^2}, \quad (3.11)$$

where $\tau_0$ and $\gamma$ are constants.

Interestingly, $\psi$ is not a monotone increasing function of cell density. Instead, the traction stress is assumed to decrease at very high fibroblast densities. This is to reflect the impact of contact inhibition on fibroblast traction; fibroblasts reduce the stress that they exert when they can feel the presence of other fibroblasts. The constant $\gamma$ is a measure of this contact inhibition while $\tau_0$ represents the ‘strength’ of the fibroblasts (i.e. the maximum force that they are able to exert on the ECM). Cell traction is also assumed to be proportional to the ECM density. This represents the fact that increasing the ECM density increases the number of sites to which cells can attach.\(^7\)

The other term in equation (3.10) that we need to consider is the body force, $b$. Tranquillo and Murray use a body force to model the tethering of the dermis to the hypodermis and subdermal tissues. In their model, this tethering force is taken to be directly proportional to the displacement and to the local ECM density; denser parts of the dermis are assumed to have more subdermal attachments and thus experience a greater restoring force.

Substituting equation (3.11) and an appropriate expression for body force into equation (3.10), we obtain

$$\mu \frac{\partial^3 u}{\partial x^2 \partial t} + k \frac{\partial^2 u}{\partial x^2} + \frac{\partial}{\partial x} \left( \frac{\tau_0 \rho n}{1 + \gamma n^2} \right) = s \rho u, \quad (3.12)$$

where $s$ represents the strength of the subdermal attachments and all other terms are defined as before. Equation (3.12) can be divided through by $k$ and non-di-
dimensionalised. After redefining the parameters appropriately, this yields equation (3.3) as given above.

In Section 3.3, we also consider the case where \( k \) is equal to zero. In this situation, it is no longer appropriate to divide through by \( k \) before nondimensionalising and equation (3.3) can no longer be used. Instead, we divide by \( \mu \) to obtain the following equation:

\[
\frac{\partial^3 u}{\partial x^2 \partial t} + \bar{E} \frac{\partial^2 u}{\partial x^2} + \frac{\partial}{\partial x} \left( \frac{\bar{\tau} \rho n}{1 + \gamma n^2} \right) = \bar{s} \rho u,
\]

(3.13)

where \( \bar{E} \), \( \bar{\tau} \) and \( \bar{s} \) are all dimensionless. These dimensionless parameters can be related to the dimensionless parameters of equation (3.3) as follows:

\[
\mu = \frac{1}{\bar{E}} = \frac{\tau_0}{\bar{\tau}} = \frac{s}{\bar{s}}.
\]

Initially, it is assumed that displacement is zero everywhere (i.e. no contraction or retraction of the wound boundary has occurred). This is a sensible simplification but it is inconsistent with some experimental observations.\(^8\) Importantly, early wound retraction would appear to be much faster than the deposition of the initial matrix. This means that an initial displacement of zero is inconsistent with the assumption that the initial density is spatially uniform. By the time a sturdy preliminary matrix has formed, we would expect that a certain amount of wound boundary retraction has already occurred.

Tranquillo and Murray also stipulate that the displacement, \( u \), at the wound centre is always zero. This follows from the physical symmetry of the problem; a nonzero displacement at \( x = 0 \) must correspond to either a tear or a superposition of points, neither of which is physically realistic.

\(^8\)McGrath and Simon [126] found that the wound area increased by up to 40% in the very early stages of murine wound healing. Interestingly, this means that the ‘initial area’ that they use in their exponential model of wound contraction is actually significantly larger than the size of the excised region. In a mechanical model, we should ideally account for this early expansion by using a nonzero initial condition for displacement.
Lastly, it is assumed that the mechanical effect of the wound is limited to the nearby dermis. Thus, the boundary condition given for the displacement at \( x = \infty \) takes the same form as the boundary conditions on \( n \) and \( \rho \):

\[ u_x(\infty, t) = 0. \]

### 3.2.2 Preliminary mathematical analysis

Having given an outline of the base Tranquillo-Murray model, we will now critically analyse some of the model’s mathematical features. The analysis in this section will form the basis for the closer investigation of the model given in Section 3.3.

#### Boundary conditions

Firstly, let us consider the right-hand boundary conditions stated by Tranquillo and Murray:

\[ n_x(\infty, t) = 0, \quad \rho_x(\infty, t) = 0, \quad u_x(\infty, t) = 0. \]  

(3.14)

These conditions effectively state that the dependent variables of the model are bounded and tend to some limit far from the wound. However, it is possible to show that these boundary conditions are actually equivalent to stronger boundary conditions (*i.e.* conditions on \( n \), \( \rho \) and \( u \) rather than on their derivatives).

In order to prove this, we need the following result:

‘If \( f(x) \) converges as \( x \to \infty \) and \( \lim_{x \to \infty} f'(x) \) exists, then \( \lim_{x \to \infty} f'(x) = 0. \)’

We can easily prove this by contradiction using the formal definition of the derivative (not given here).
Since equation (3.3) involves $\frac{\partial^3 u}{\partial x^2 \partial t}$, it follows that $\frac{\partial^2 u}{\partial x^2}$ exists and is well-behaved for all values of $x$ and $t$. Furthermore, condition (3.9) indicates that $u_x(\infty, t) = 0$. That is, as $x$ tends to infinity, we find that $\frac{\partial u}{\partial x}$ converges and the limit of $\frac{\partial^2 u}{\partial x^2}$ exists. Hence,

$$
\lim_{x \to \infty} \frac{\partial^2 u}{\partial x^2} = 0, 
$$

and differentiating with respect to $t$ yields

$$
\lim_{x \to \infty} \frac{\partial^3 u}{\partial x^2 \partial t} = 0.
$$

Having obtained equations (3.15) and (3.16), consider what happens to (3.3) as $x$ tends to infinity:

$$
\lim_{x \to \infty} \mu \frac{\partial^3 u}{\partial x^2 \partial t} + \lim_{x \to \infty} \frac{\partial^2 u}{\partial x^2} + \lim_{x \to \infty} \frac{\partial}{\partial x} \left( \frac{\tau_0 \rho n}{1 + \gamma n^2} \right) = \lim_{x \to \infty} s \rho u,
$$

which implies that

$$
0 + \lim_{x \to \infty} \left[ \frac{\partial}{\partial \rho} \left( \frac{\tau_0 \rho n}{1 + \gamma n^2} \right) \frac{\partial \rho}{\partial x} + \frac{\partial}{\partial n} \left( \frac{\tau_0 \rho n}{1 + \gamma n^2} \right) \frac{\partial n}{\partial x} \right] = \lim_{x \to \infty} s \rho u.
$$

Now, we recall that $n_x(\infty, t)$ and $\rho_x(\infty, t)$ both converge to zero and the partial derivatives of $\frac{\tau_0 \rho n}{1 + \gamma n^2}$ are bounded. Thus,

$$
0 = \lim_{x \to \infty} s \rho u, 
$$

Assuming the parameter $s$ to be nonzero, equation (3.17) implies that either $\rho(\infty, t) = 0$ or $u(\infty, t) = 0$. The first of these possibilities is physically unrealistic and cannot arise in the Tranquillo-Murray model because there is no degradation of the ECM and the initial ECM density is nonzero. Thus, we conclude that the boundary conditions listed in (3.14) imply that

$$
u(\infty, t) = 0, 
$$
provided that $s$ is nonzero.

By a similar process, we can use equations (3.1) and (3.2) to show that

$$\frac{d}{dt} [n(\infty, t)] = r_0 n(\infty, t) (1 - n(\infty, t))$$

and

$$\frac{d}{dt} [\rho(\infty, t)] = 0.$$

Using the initial conditions on $n$ and $\rho$ stated in equations (3.4) and (3.5), this yields

$$n(\infty, t) = \rho(\infty, t) = 1.$$

**Density as a simple function of displacement**

Another interesting point to note is that it is possible to use equation (3.2) to express $\rho$ as an explicit function of $u$. To demonstrate this, consider the case where we use the true velocity, $v$, in place of the approximate velocity, $\frac{\partial u}{\partial t}$.\(^9\)

Substituting $v$ into equation (3.2) and rearranging, we find that

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x} (\rho v) = \frac{D \rho}{Dt} + \rho \frac{\partial v}{\partial x} = 0.$$ \( \text{(3.19)} \)

By converting this equation to Lagrangian coordinates and integrating, it is possible to show that

$$\rho(x, t) = \rho(x, 0) \left( 1 - \frac{\partial u}{\partial x} \right).$$ \( \text{(3.20)} \)

As noted by Tranquillo and Murray, this is also the approximate solution obtained

\(^9\)The true velocity is the material derivative of displacement:

$$v = \frac{D u}{Dt} = \frac{\partial u}{\partial t} + v \frac{\partial u}{\partial x}.$$
from linearising equation (3.2).

This means that Tranquillo and Murray’s approximate solution is actually exact. Their linearisation has the same effect as replacing the approximate velocity with the true velocity. In effect, equation (3.20) is valid when the displacement gradient is large, even when (3.2) is not.

The disadvantage of (3.20) is that it cannot easily be modified to take into account an evolving ECM density. Thus, (3.20) is not practical for wound healing models that are more complicated than the Tranquillo-Murray model.

3.2.3 Numeric simulations

The Tranquillo-Murray model of dermal wound healing is highly nonlinear and cannot be solved exactly; thus, it is necessary to obtain solutions using numeric methods. Normally, this would not be cause for any concern. However, the Tranquillo-Murray model is unusual. Although the cell density and ECM density equations are reaction-diffusion equations (and hence parabolic), the force balance equation is not.

Indeed, we note that the force balance equation includes both a third-order mixed derivative, \( \frac{\partial^3 u}{\partial x^3} \), and a term proportional to \( u \). This makes it highly difficult to determine whether the force balance equation is parabolic, elliptic or hyperbolic. There is no simple way of determining the characteristic surfaces associated with this equation, and it may even have a different classification for different domains of \( x \) and \( t \). Thus, we cannot be confident that a standard parabolic-elliptic solver will give us accurate numeric results.

Despite these concerns, we attempted to use parabolic-elliptic solvers to obtain numeric solutions of the Tranquillo-Murray model. We were unable to run the model using Matlab’s pdepe routine [184], but NAG’s Fortran-based D03PCF
package [185] worked successfully. All of the numeric simulations described in this chapter were performed using D03PCF.

Unless otherwise noted, an evenly spaced mesh with a spatial step size of 0.05 was used in all of the simulations that follow. Halving the spatial step size did not cause any noticeable changes to the solutions obtained; this mesh independence suggests that the spatial step size is sufficiently small to avoid errors and instabilities related to the coarseness of the mesh. The size of the time step was determined by internal D03PCF routines. As described in the captions of the figures that follow, selected time steps are used to illustrate the evolution of the solutions over time.

The Tranquillo-Murray model of wound healing is defined on a semi-infinite spatial domain that must be truncated at some point for numerical simulations. In most of the simulations shown, we follow the example of the original Tranquillo-Murray paper and apply the right hand boundary conditions at $x_{\text{max}} = 10$. Exceptions to this rule can be found in Figure 3.3 (where $x_{\text{max}} = 20$) and in Figure 3.12 (where $x_{\text{max}} = 40$). In both of these cases, an extended domain is used so that the behaviour of displacement at large values of $x$ can be investigated.

Although the numeric results that we obtained are consistent with our qualitative and quantitative expectations of the Tranquillo-Murray model, it should be noted that we cannot be completely confident of their accuracy without further knowledge about the mathematical nature of the force balance equation. We would suggest that further research into the development of appropriate numeric methods is necessary before mechanochemical modelling of biological processes can come into its own as a practical technique.

**Standard Tranquillo-Murray model**

Using the numeric approach described above, we were able to replicate Tranquillo and Murray’s original results for their base model, shown in Figure 3.2 (equivalent
Figure 3.2: Numeric results obtained from the base Tranquillo-Murray model using the parameter values given in the original paper: $r_0 = 1$, $\mu = 1$, $s = 1$, $\tau_0 = 0.5$ and $\gamma = 1$. Initial conditions are plotted as dashed lines. Profiles at $t = 0.5$ are plotted as very thick lines and profiles at $t = 1.0$ are plotted as thick lines. Further profiles representing time increasing in steps of 0.5 up to $t = 4.0$ are also shown as lines of normal thickness.
to Figure 2 of Tranquillo and Murray [189]). One of the well-noted features of this model is that the displacement of the wound boundary is always positive. This happens because there are always more cells outside the wound than inside the wound and the cell traction stress is a monotone increasing function of the cell density.\(^\text{10}\) As a result, the net force due to cell traction acts in the positive \(x\) direction and the resultant displacements are always positive.

This is at odds with clinical observations of dermal wound contraction (see, for example, Majno [117]). After a period of wound retraction, we would expect the ECM displacement at the wound boundary to become negative (corresponding to contraction). As noted in Section 2.3.1, Tranquillo and Murray (and later Olsen) proposed a number of model extensions that use a chemical mediator to drive contraction.

Since the model extensions necessary for introducing wound contraction are well documented (see especially Olsen [136]), we will not discuss them further in this chapter. Although wound contraction is important when constructing a mechanochemical wound healing model, our focus is on other mechanical features of the Tranquillo-Murray model that have not yet been investigated.

Having commented on the fact that \(u\) is positive in the vicinity of the wound, it is interesting to note that Figure 3.2 shows \(u\) becoming negative at large values of \(x\). We recall from Section 3.2.2 that the boundary condition \(u_x(\infty, t) = 0\) is equivalent to \(u(\infty, t) = 0\). Thus, if we were solving on a sufficiently large spatial domain, we would expect to find that \(u\) converges to zero at large \(x\). Instead, we see that \(u\) ‘overshoots’ zero and reaches a negative value at the end of the \(x\) domain.

We might suspect that this is an artifact introduced by the numeric scheme used.

\(^{10}\) As we increase \(n\), we actually find that Tranquillo and Murray’s expression for cell traction stress reaches a maximum and then tends to zero as \(n \to \infty\). However, since \(n\) never gets much larger than 1 in model simulations, it is possible to choose the parameter \(\gamma\) so that the traction stress is a monotone increasing function of cell density within the physiological range of \(n\) values. Since it is reasonable to expect that more cells leads to a greater stress, we will assume that the model parameters have been chosen to ensure that this is always the case.
to solve for $u$; as noted above, we cannot be sure that the solution obtained for $u$ using a parabolic-elliptic solver is correct. However, we find that there is no significant change to our numeric results if we use a finer spatial mesh or a larger spatial domain. In fact, if we use a sufficiently large spatial domain and fine spatial mesh, we can see that the solution for $u$ oscillates through space around $u = 0$. This is illustrated in Figure 3.3.$^{11}$

![Figure 3.3](image)

Figure 3.3: Simulation of the Tranquillo-Murray model using the same parameters as Figure 3.2, but considering an extended spatial domain. As before, the initial condition is shown as a dashed line, the profile at $t = 0.5$ is shown as a very thick line, the profile at $t = 1.0$ is shown as a thick line and subsequent profiles (normal thickness) represent time increasing in steps of 0.5 up to $t = 4.0$.

Interestingly, we find that these oscillations persist when different parameter values are used and when the model is extended to include an inflammatory mediator (not shown). This suggests that the ‘displacement oscillations’ are not numeric artifacts but are actually inherent to the model used. Section 3.3 is devoted to demonstrating that this is indeed the case and in Appendix A, we show that all models using the Tranquillo-Murray paradigm exhibit displacement oscillations.

$^{11}$Interestingly, the same qualitative behaviour is observed if the boundary condition $u(x_{\text{max}}, t) = 0$ is applied numerically. Displacement oscillations are still observed with this boundary condition, although the results are quantitatively different as $x_{\text{max}}$ is approached.
Tranquillo-Murray model without advection

Another interesting observation from Figure 3.2 is that \( u \) is consistently small; using the same parameters as Tranquillo and Murray, we find that \( u \) is never any larger than 0.035. Since \( n \) and \( \rho \) are comparatively large, this suggests that the advection terms in equations (3.1) and (3.2) will be much smaller than the other terms. Thus, let us compare Tranquillo and Murray’s original system of equations with the simplified system obtained when advection is removed.

That is, we wish to solve the system

\[
\frac{\partial n}{\partial t} = \frac{\partial^2 n}{\partial x^2} + r_0 n (1 - n), \quad (3.21)
\]

\[
\frac{\partial \rho}{\partial t} = 0, \quad (3.22)
\]

\[
\mu \frac{\partial^3 u}{\partial x^2 \partial t} + \frac{\partial^2 u}{\partial x^2} + \frac{\partial}{\partial x} \left( \frac{\tau_0 \rho n}{1 + \gamma n^2} \right) = s \rho u; \quad (3.23)
\]

subject to the same boundary conditions and initial conditions that were used for the original problem (i.e. equations (3.4) to (3.9)).

Even without solving this system numerically we can see that omitting advection leads to some significant simplifications. Most importantly, we note that equations (3.21) and (3.22) are completely decoupled from (3.23) and from each other. Thus, it is possible to solve for \( n \) and \( \rho \) completely before we attempt to solve for \( u \).

Interestingly, the absence of ECM synthesis and degradation from the base Tranquillo-Murray model means that \( \rho \) does not change over time in the simplified model. Given the initial condition on \( \rho \) from (3.5), this yields \( \rho(x, t) \equiv 1 \). This is very close to the \( \rho \) solution given in Figure 3.2; in the full model, \( \rho \) is never

\[\text{Note that it is not necessary to have a boundary condition for } \rho \text{ in this case, as there are no terms of the form } \frac{\partial \rho}{\partial x}. \text{ However, the solution that we obtain for } \rho \text{ in the simplified system is still consistent with the } \rho \text{ boundary condition given in (3.8).}\]
any smaller than 0.95 or any larger than 1.02.

Similarly, \( n \) can be solved independently of \( \rho \) and \( u \) in this case and we note that the \( n \) equation in the simplified model is identical to Fisher’s equation. Although we cannot obtain a closed-form solution for Fisher’s equation under these conditions, we can compare the numeric solution of Fisher’s equation with the numeric solution of \( n \) from the full model.\(^{13}\) This is shown in Figure 3.4, which demonstrates that the two solutions are very similar.

![Figure 3.4: Cell density profiles from the standard Tranquillo-Murray model (shown as continuous lines) compared with the equivalent model with advection omitted (shown as dotted lines). The same parameter values are used as for Figure 3.2. As before, the initial condition is shown as a dashed line, the profiles corresponding to \( t = 0.5 \) are shown as a very thick lines, the profiles corresponding to \( t = 1.0 \) are shown as thick lines and subsequent profiles (normal thickness) represent time increasing in steps of 0.5 up to \( t = 4.0 \). Note that there is very little difference between the two sets of outputs.](image)

Lastly, we note that substituting \( \rho(x, t) = 1 \) into equation (3.23) yields

\[
\mu \frac{\partial^3 u}{\partial x^2 \partial t} + \frac{\partial^2 u}{\partial x^2} + \frac{\partial}{\partial x}\left(\frac{\tau_0 n}{1 + \gamma n^2}\right) = s u. \tag{3.24}
\]

\(^{13}\)Closed-form solutions of Fisher’s equation can be found in certain circumstances (see, for example, Guo and Chen [83]). However, the techniques used to obtain these solutions are only applicable when the wave is travelling over an infinite domain (i.e., unaffected by boundary conditions). In contrast, the model of wound healing considered above involves a Fisher-like wave that is strongly affected by the no-flux boundary condition at \( x = 0 \). Hence, analytic techniques are of limited use.
Using the Fisher’s equation solution for $n$, we can solve this numerically to obtain the results shown in Figure 3.5. As before, we find that there is good agreement between the full model and the simplified model.

Figure 3.5: ECM displacement profiles from the standard Tranquillo-Murray model (shown as continuous lines) compared with the equivalent model with advection omitted (shown as dotted lines). The same parameter values are used as for Figure 3.2. As before, the initial condition is shown as a dashed line, the profiles corresponding to $t = 0.5$ are shown as very thick lines, the profiles corresponding to $t = 1.0$ are shown as thick lines and subsequent profiles (normal thickness) represent time increasing in steps of 0.5 up to $t = 4.0$. Again, there is very little difference between the two sets of outputs.

This is a very interesting result. By performing further simulations, it is possible to show that removing advection has a relatively small effect for a wide range of parameter values. Moreover, removing advection leads to significant simplifications; solving the Tranquillo-Murray model without advection (or equivalently, solving one of the Olsen et al. models [136, 137, 138, 139, 140, 141] without advection) can be separated into two steps. Firstly, we solve a system of reaction-diffusion equations to find the evolving cell density, ECM density etc. and then we use the solutions obtained in the first step to solve for displacement. Thus, removing advection allows us to turn a mechanochemical model into a much more conventional problem.

Although this has a few computational advantages, it leads us to question the purpose of constructing a mechanochemical model. If the equation for ECM
mechanics is an ‘optional add-on’ as it appears to be, why should it not be removed altogether?

There are two possible answers to this. Firstly, we may be interested in tracking the displacements of particles in the skin for their own sakes. That is, we wish to use a mechanochemical model to describe the displacement of the dermis but we are not interested in the effect of tissue mechanics on other aspects of wound healing. This is effectively the approach taken by Tranquillo and Murray and later by Olsen et al.; by making advection the only form of feedback from the mechanical equation into the continuity equations, they eliminate the possibility of any real two-way interaction.

Alternatively, we may wish to construct a model in which mechanical signals have a real effect on the behaviour of cells and other species. As described in Chapter 2, there is an increasing body of evidence that indicates that fibroblasts respond to their mechanical environment. In order to model this, we need to construct a system of equations that is more strongly coupled. This is the approach that we take in the wound healing model described in Chapter 6.

Other modifications

In order to gain a better appreciation for the Tranquillo-Murray model, it is instructive to run simulations using a range of different parameter values. However, it is impossible to consider every combination of parameter values. In this section, we illustrate the effects of changing the parameter values by considering the results obtained when one or more of the model parameters are taken to be zero.

For example, consider the case where $\gamma = 0$. This corresponds to a situation where there is no inhibition of cell traction stress at higher densities; instead, $\psi$ is directly proportional to $n$. Thus, $\gamma = 0$ increases the traction stress exerted by cells and we would expect this to result in larger displacements. As shown in Figure 3.6, this is indeed what we observe.
Figure 3.6: Numeric results obtained from the base Tranquillo-Murray model in the case where $\gamma = 0$ but all other parameter values are the same as for Figure 3.2. As before, the initial conditions are shown as dashed lines, the profiles corresponding to $t = 0.5$ are shown as very thick lines, the profiles corresponding to $t = 1.0$ are shown as thick lines and subsequent profiles (normal thickness) represent time increasing in steps of 0.5 up to $t = 4.0$. 
One side effect of this is that the velocity is initially large but it decreases rapidly. Numerically, this leads to small errors in the estimation of velocity that would normally have little effect. However, the evolution of density depends only on advection (i.e. only on velocity). Thus, errors in velocity have a cumulative effect on the density profile, causing it to ‘kink’ at long times instead of settling back to a uniform profile. These numeric errors can be reduced by using equation (3.20) to eliminate $\rho$ from the system of equations (not shown here).

Apart from the larger displacements and associated effects, there appear to be no significant qualitative differences between the full model and the $\gamma = 0$ model. Although it cannot be seen clearly from Figure 3.6, it is possible to show that the displacement oscillations at large $x$ are retained.

Next, consider the results obtained when $\tau_0 = 0$. In this case, there is no traction stress and thus no displacement of the ECM and no advection. As a result, we obtain the same results for $n$ and $\rho$ as for the zero advection case described earlier.

When $r_0 = 0$ we find that there is no cellular proliferation, only random motion and advection. As shown in Figure 3.7, this leaves the $\rho$ and $u$ profiles practically unchanged. Instead, the most significant difference between the $r_0 = 0$ model and the full model is that it takes longer for the cell density to become spatially homogeneous. Thus, the imbalance of cell traction stresses persists for a longer period of time, causing the displacement to be slightly larger.

One particularly interesting case to investigate is the ‘viscous-only’ Tranquillo-Murray model (i.e. the model obtained when elastic effects are removed). In order to express this algebraically we need to rescale equation (3.3) to obtain equation (3.13) and let $\bar{E} = 0$. This yields the results shown in Figure 3.8.

Interestingly, we find that there are clearly visible differences between the viscous-only model and the full model. In particular, we note that displacements obtained

\footnote{If we run the simulation to very long times, we also find that the equilibrium value of $n$ is slightly less than 1. However, since $n(x, \infty) \to 1$ on a truly semi-infinite domain, we find that can reduce the significance of this effect by using a wider domain.}
Figure 3.7: Numeric results obtained from the base Tranquillo-Murray model in the case where $r_0 = 0$ but all other parameter values are the same as for Figure 3.2. As before, the initial conditions are shown as dashed lines, the profiles corresponding to $t = 0.5$ are shown as very thick lines, the profiles corresponding to $t = 1.0$ are shown as thick lines and subsequent profiles (normal thickness) represent time increasing in steps of 0.5 up to $t = 4.0$. 
in the viscous-only model are larger and the displacement oscillations are much more pronounced.\textsuperscript{15} This suggests that the oscillations in displacement are related to the viscous nature of the skin. When we analyse displacement oscillations in Section 3.3, we will use a viscous-only caricature model in order to emphasise this behaviour.

Next, consider the case where \( \mu \), the coefficient of viscosity, is small. As noted by Tranquillo and Murray, \( \mu \to 0 \) is a singular limit of the model that leads to computational difficulties. Since \( \mu \) multiplies the highest time derivative in equation (3.3), we find that taking \( \mu \) to be small leads to boundary layer effects (i.e. there is an initial transient solution).

When \( \mu = 0 \), we find that our standard simulation method fails to produce any output.\textsuperscript{16} However, it is possible to take \( \mu \) to be very small and still obtain output. One example of this is shown in Figure 3.9, where \( \mu \) is taken to be 0.0001. This figure shows clear evidence of numeric instabilities. There is a large perturbation in the \( \rho \) profile at \( x = 1 \) and the \( u \) profile is clearly kinked, even at early times.

These problems are a result of the numeric method failing to deal appropriately with the initial transient. In particular, we find that the velocities at early times are very large but rapidly changing. By focusing our attention on short times, we find that \( u \) reaches a maximum value close to 0.12 before \( t = 0.0005 \). Even with adaptive time-stepping, it is not surprising that such rapid changes cause problems for our numeric methods.

Interestingly, however, the displacement oscillations in Figure 3.9 seem to have been dramatically reduced. Thus, we hypothesise that oscillations arise as a result of the viscous contribution to the dermal constitutive law. As the ‘dermal

\textsuperscript{15}We even see some small oscillations in the ECM density as a direct result of the oscillations in displacement.

\textsuperscript{16}Setting \( \mu = 0 \) corresponds to the case where the constitutive law is purely elastic rather than viscoelastic. As we will see, this is an important case to consider; the biological literature indicates that the skin effectively behaves as an elastic solid on the time scale of wound healing [50, 105]. However, \( \mu = 0 \) also leads to some mathematical problems. These will be discussed further in Section 3.4.2.
Figure 3.8: Numeric results obtained from the viscous-only Tranquillo-Murray model. That is, equation (3.13) replaces equation (3.3) and parameter values of $r_0 = 1$, $\bar{E} = 0$, $\bar{s} = 1$, $\bar{\tau} = 0.5$ and $\gamma = 1$ are used. As before, the initial conditions are shown as dashed lines, the profiles corresponding to $t = 0.5$ are shown as very thick lines, the profiles corresponding to $t = 1.0$ are shown as thick lines and subsequent profiles (normal thickness) represent time increasing in steps of 0.5 up to $t = 4.0$. 
viscosity’ approaches zero, we find that the displacement oscillations disappear.

Lastly, consider the case where there is no subdermal tethering (i.e. where \( s = 0 \)). This is shown in Figure 3.10. Interestingly, the untethered simulations behave very differently from the full model. Instead of finding that \( u \) oscillates around 0 at large values of \( x \), we find that \( u \) increases monotonically with \( x \), tending to a nonzero value as \( x \) approaches infinity. Over time, elastic stress causes the entire displacement profile to tend towards zero, but it would still appear that \( u \) attains its maximum as \( x \to \infty \).\(^{17}\)

We recall from Section 3.2.2 that \( u_x(\infty, t) = 0 \) implies that \( u(\infty, 0) \), but only on the condition that \( s \neq 0 \). Without a tethering force to dampen the mechanical effects of the wound, there is no reason for displacement to tend towards zero as \( x \) grows large. This is not physically realistic; a wound should not have long-range mechanical effects of the magnitude shown in Figure 3.10. Thus, some sort of tethering term appears to be a necessary part of a mechanical model of dermal wound healing.

Further numeric simulations can be used to illustrate other features of the Tranquillo-Murray model. However, much useful information can be obtained simply by considering the results given above. Most notably, we find that \( \mu \to 0 \) and \( s \to 0 \) are both singular limits of the model in which the qualitative behaviour of displacement changes significantly. In contrast, we find that the effect of advection is often negligible and that the removal of advection leads to a decoupled model that yields very similar results to the full Tranquillo-Murray model. One intriguing result of the simulations is that displacement oscillations are observed for a wide variety of different parameter values. This phenomenon is investigated further in the following section.

\(^{17}\)In Figure 3.10, we see that the ECM displacement and ECM density profiles at \( t = 0.5 \) are very similar to the profiles at \( t = 1.0 \). This is a result of the fact that the displacement increases rapidly at early times, reaching a maximum at around \( t = 0.7 \) and then entering a steady decline. Thus, the \( t = 0.5 \) profile occurs during the early increase while the \( t = 1 \) profile occurs during the decline.
Figure 3.9: Numeric results obtained from the base Tranquillo-Murray model in the case where $\mu = 0.0001$ but all other parameter values are the same as for Figure 3.2. As before, the initial conditions are shown as dashed lines, the profiles corresponding to $t = 0.5$ are shown as very thick lines, the profiles corresponding to $t = 1.0$ are shown as thick lines and subsequent profiles (normal thickness) represent time increasing in steps of 0.5 up to $t = 4.0$. Note that a significant amount of numeric instability is observed as the singular limit of $\mu = 0$ is approached.
Figure 3.10: Numeric results obtained from the base Tranquillo-Murray model in the case where $s = 0$ but all other parameter values are the same as for Figure 3.2. As before, the initial conditions are shown as dashed lines, the profiles corresponding to $t = 0.5$ are shown as very thick lines, the profiles corresponding to $t = 1.0$ are shown as thick lines and subsequent profiles (normal thickness) represent time increasing in steps of 0.5 up to $t = 4.0$. 
3.3 Asymptotic analysis

3.3.1 Simplifying the Tranquillo-Murray model

One feature of the Tranquillo-Murray model that we would like to explain mathematically is the oscillatory behaviour observed in the displacement profile at large values of $x$. Our numeric results suggest that these displacement oscillations are a result of the three-way interaction between the viscous stress, the cell traction stress and the tethering force. In this section, we use the techniques of asymptotic analysis to demonstrate that this interaction does indeed lead to displacement oscillations.

Unfortunately, the Tranquillo-Murray model is too complex to analyse in its present form. Thus, we will focus our attention on the analysis of a caricature model. Using the results that we obtain from considering a caricature model, we are then able to generalise our analysis to a large class of mechanochemical models. In particular, we are able to show that all mechanochemical models using the Tranquillo-Murray paradigm (including the extended Tranquillo-Murray models and the Olsen models) will exhibit displacement oscillations. The asymptotic analysis of a general model is presented in Appendix A.

A secondary aim of this analysis is to explain the small displacements observed in our numeric simulations. We recall that using the same parameter values as Tranquillo and Murray yielded simulations in which $u$ was consistently less than 0.035. This suggests the existence of an alternative nondimensionalisation for displacement that would be more appropriate than the original scaling. However, the parameter values used by Tranquillo and Murray were all of order unity. Thus, any improved nondimensionalisation is likely to be an unusual function of the model parameters (i.e. not a simple product or quotient).

As noted above, we suspect that the displacement oscillations are a result of the interaction of viscous effects with the traction stress and the restoring force.
Thus, we propose a simplified model in which $\tau_0$ and $s$ are nonzero but all other nonlinear terms are removed.\(^{18}\) That is, we use the alternative rescaling of the force balance equation (i.e. equation (3.13)) and we set $\gamma = r_0 = \bar{E} = 0$. Omitting the bars from altered dimensionless parameters, we obtain the following system:

\[
\begin{align*}
\frac{\partial n}{\partial t} + \frac{\partial}{\partial x} \left( n \frac{\partial u}{\partial t} \right) &= \frac{\partial^2 n}{\partial x^2}, \\
\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x} \left( \rho \frac{\partial u}{\partial t} \right) &= 0, \\
\frac{\partial^3 u}{\partial x^2 \partial t} + \frac{\partial}{\partial x} \left( \tau \rho n \right) &= s \rho u;
\end{align*}
\]

with boundary and initial conditions given by equations (3.4) to (3.9).

Importantly, we recall from Section 3.2.3 that setting $\gamma$, $r_0$ and/or $\bar{E}$ to be zero does not significantly affect the qualitative results obtained from the Tranquillo-Murray model. Furthermore, we find that taking any of these three parameters to zero (individually or together) has the effect of marginally increasing displacement. Thus, the maximum displacement obtained from a simplified model where $\gamma = r_0 = \bar{E} = 0$ can be thought of as an upper bound on displacement for the full model.\(^{19}\) If this is demonstrably small, it may give us an explanation of the small displacements in the full model, and possible even a sensible rescaling for $u$.

Interestingly, our caricature model contains only two dimensionless parameters, $\tau$ and $s$. In the original Tranquillo-Murray model, these parameters were assigned values of 0.5 and 1 respectively.\(^{20}\) Since neither of these values are small or large,

\(^{18}\)For convenience of notation, we use $\tau$ in place of $\tau_0$ throughout this section.

\(^{19}\)Although this is instinctively correct, it is unfortunately impossible to show that the $u$ solution of the full model is bounded by the $u$ solution of the simplified model in all cases. For example, although it can be shown that decreasing $\gamma$ from 1 to 0 increases the cell traction stress, it is not readily obvious that this necessarily leads to greater displacements. When decreasing $r_0$ or $\bar{E}$, even more difficulties are encountered; for instance, decreasing $r_0$ does not uniformly increase the traction stress. Hence, any upper bound obtained from the simple model is not actually rigorous.

\(^{20}\)Since $\mu = 1$ in the Tranquillo-Murray model, the rescaling of the force balance equation does not affect the values of $\tau$ and $s$. 
Analysis of the Tranquillo-Murray wound healing model

it would seem inappropriate to use either of these parameters as the basis of an asymptotic expansion. However, the fact that displacement is consistently observed to be small when these parameters are used suggests that asymptotic methods will work.

Furthermore, we observe from our numeric simulations that the maximum displacement is roughly proportional to $\tau$ (both in the caricature model and in the full model). Also, this relationship becomes more precisely linear as $\tau$ gets smaller; we ultimately find that $\tau = 0$ corresponds to $u(x, t) \equiv 0$. Although these numeric results cannot be used to rigorously define a limit, this suggests that

$$\lim_{{\tau \to 0}} \frac{\max (u(x, t))}{\tau} = \kappa,$$

(3.28)

where max represents the maximum over all values of $x$ and $t$ and $\kappa$ is an unspecified constant that depends on the model parameters. Using the formal definition of asymptotic order, this implies that $u \sim O(\tau)$ as $\tau \to 0$. Thus, proposing an expansion in small $\tau$ is consistent with the small displacements observed in our numeric simulations.\(^{21}\)

Therefore, we propose that $n$, $\rho$ and $u$ can be represented by asymptotic series of the form

$$n(x, t; \tau) \sim n_0(x, t) + \tau n_1(x, t) + \ldots,$$

$$\rho(x, t; \tau) \sim \rho_0(x, t) + \tau \rho_1(x, t) + \ldots,$$

$$u(x, t; \tau) \sim u_0(x, t) + \tau u_1(x, t) + \ldots;$$

where the functions $n_i(x, t)$, $\rho_i(x, t)$ and $u_i(x, t)$ are all of order unity and $\tau$ tends to zero.\(^{22}\)

\(^{21}\)Interestingly, it can be shown that the perturbation solution that we obtain using the assumption of small $\tau$ is unreasonably good; accurate results can still be obtained when $\tau$ is larger than 1. It would appear that the accuracy of the perturbation solution depends more on the size of $u$ than the size of $\tau$.

\(^{22}\)Since our numeric results suggest that $u \sim O(\tau)$, we anticipate that $u_0$ will be identically zero. However, it is not necessary to assume this at this stage.
Substituting into equations (3.25) to (3.27), we find that

\[
\frac{\partial n_0}{\partial t} + \tau \frac{\partial n_1}{\partial t} + \ldots + \frac{\partial}{\partial x} \left[ (n_0 + \tau n_1 + \ldots) \left( \frac{\partial u_0}{\partial t} + \tau \frac{\partial u_1}{\partial t} + \ldots \right) \right] = \frac{\partial^2 n_0}{\partial x^2} + \tau \frac{\partial^2 n_1}{\partial x^2} + \ldots, \tag{3.29}
\]

\[
\frac{\partial \rho_0}{\partial t} + \tau \frac{\partial \rho_1}{\partial t} + \ldots + \frac{\partial}{\partial x} \left[ (\rho_0 + \tau \rho_1 + \ldots) \left( \frac{\partial u_0}{\partial t} + \tau \frac{\partial u_1}{\partial t} + \ldots \right) \right] = 0, \tag{3.30}
\]

\[
\frac{\partial^3 u_0}{\partial x^2 \partial t} + \tau \frac{\partial^3 u_1}{\partial x^2 \partial t} + \ldots + \tau \left( \frac{\partial n_0}{\partial x} \rho_0 + n_0 \frac{\partial \rho_0}{\partial x} + \ldots \right) = s (\rho_0 + \tau \rho_1 + \ldots) (u_0 + \tau u_1 + \ldots). \tag{3.31}
\]

Collecting only the terms of order unity, this yields

\[
\frac{\partial n_0}{\partial t} + \frac{\partial}{\partial x} \left( n_0 \frac{\partial u_0}{\partial t} \right) = \frac{\partial^2 n_0}{\partial x^2}, \tag{3.32}
\]

\[
\frac{\partial \rho_0}{\partial t} + \frac{\partial}{\partial x} \left( \rho_0 \frac{\partial u_0}{\partial t} \right) = 0, \tag{3.33}
\]

\[
\frac{\partial^3 u_0}{\partial x^2 \partial t} = s \rho_0 u_0. \tag{3.34}
\]

The initial and boundary conditions that we use can be obtained by substituting the asymptotic series for \( n \), \( \rho \) and \( u \) into conditions (3.4) to (3.9). In particular, we note that conditions (3.6) and (3.9) infer that

\[
u_0(x, 0) = 0, \quad u_0(0, t) = 0, \quad \lim_{x \to \infty} \frac{\partial u_0}{\partial x} = 0.\]

These conditions are consistent with \( u_0(x, t) \equiv 0 \), a trivial solution of equation (3.34). Since (3.34) should have a unique solution when appropriate initial and boundary conditions are specified, we conclude that \( u_0 \) is identically zero (as expected).
Now, \( u_0 \equiv 0 \) allows us to neglect the advection terms in equations (3.32) and (3.33).\(^{23}\) Thus, we find that

\[
\frac{\partial n_0}{\partial t} = \frac{\partial^2 n_0}{\partial x^2}, \quad (3.35)
\]

\[
\frac{\partial \rho_0}{\partial t} = 0. \quad (3.36)
\]

Since \( \rho_0(x, 0) \equiv 1 \) from condition (3.5), equation (3.36) has the trivial solution \( \rho_0 \equiv 1 \). Also, we note that equation (3.35) is a diffusion equation with initial and boundary conditions from (3.4) and (3.7) as follows:

\[
n_0(x, 0) = \begin{cases} 
0 & 0 < x < 1, \\
1 & x > 1; 
\end{cases} \quad n_{0x}(0, t) = 0; \quad n_{0x}(\infty, t) = 0.
\]

We can use the method of Green’s functions to solve for \( n_0 \), yielding

\[
n_0(x, t) = 1 + \frac{1}{2} \left( \text{erf} \left( \frac{x - 1}{2 \sqrt{t}} \right) - \text{erf} \left( \frac{x + 1}{2 \sqrt{t}} \right) \right). \quad (3.37)
\]

We now have the leading order solutions for \( n \) and \( \rho \). In order to construct the first nonzero term in the expansion for \( u \), we need to consider the equation obtained from the order \( \tau \) terms of (3.31). That is,

\[
\frac{\partial^3 u_1}{\partial x^2 \partial t} + \rho_0 \frac{\partial n_0}{\partial x} + n_0 \frac{\partial \rho_0}{\partial x} = s \rho_0 u_1 + s \rho_1 u_0.
\]

Now, we recall that \( u_0 \equiv 0, \rho_0 \equiv 1 \) and \( n_0 \) is given by equation (3.37). Using this information, the above equation becomes

\[
\frac{\partial^3 u_1}{\partial x^2 \partial t} - s u_1 = \frac{-1}{2 \sqrt{\pi t}} \left[ e^{-\frac{(x-1)^2}{4t}} - e^{-\frac{(x+1)^2}{4t}} \right]. \quad (3.38)
\]

\(^{23}\)Interestingly, this indicates that we could have used a perturbation in small \( \tau \) to justify the ‘no advection’ simplification described in Section 3.2.3 (see equations (3.21) to (3.23)).
Although this equation is linear, it is relatively intractable. For example, taking a Laplace transform or a Fourier transform leads to functions that are impossible to invert using standard tables. While numeric inversion would then be possible, it would not be as informative as a closed-form solution, or even an approximate closed-form solution. Moreover, we recall that we are trying to justify the claim that the observed displacement oscillations in the full model are not numeric artifacts. If oscillations were observed in a solution for \( u_1(x, t) \) that was obtained using a numeric inversion method, we would not be able to confidently state that this was not a result of further numeric instability. Instead, we need to use a second perturbation, this time in small \( s \), in order to obtain meaningful results.

3.3.2 Matched asymptotics and the Wright function

We now consider the problem of finding an asymptotic solution to equation (3.38), given above. Since \( u_1 \) is the leading order term in the expansion of displacement, we should be able to use this solution to explain the origin of the displacement oscillations observed far from the wound.

In order to avoid the confusion involved with having multiple subscripts representing the terms in different asymptotic series, we will rename \( u_1 \) as \( y \) for this section. Thus, the equation that we wish to solve is

\[
\frac{\partial^3 y}{\partial x^2 \partial t} - s y = \frac{-1}{2\sqrt{\pi t}} \left[ e^{-\frac{(x-1)^2}{4t}} - e^{-\frac{(x+1)^2}{4t}} \right].
\]  

(3.39)

Similarly to before, we will solve (3.39) subject to the initial and boundary conditions given below:

\[
y(x, 0) = 0, \tag{3.40}
\]

\[
y(0, t) = 0, \tag{3.41}
\]

\[
\lim_{x \to \infty} \frac{\partial y}{\partial x} = 0. \tag{3.42}
\]
Since $s$ is nonzero, we recall that (3.42) also implies the stronger boundary condition,

$$y(\infty, t) = 0.$$  \hspace{1cm} (3.43)

One surprisingly successful approach that can be used in this case is to propose an asymptotic expansion in small $s$. As with our small $\tau$ perturbation, we will find that the approximate solution that we obtain is unreasonably good; we are able to obtain a fairly good agreement between the numeric $s = 1$ results and our asymptotic results. However, unlike the previous problem where we exploited the fact that $u \sim O(\tau)$, we have no physical motivation for considering small values of $s$. Furthermore, the $s = 0$ model behaves qualitatively differently to the model with nonzero $s$.\textsuperscript{24} Ultimately, this means that we need singular perturbation techniques, in this case matched asymptotics, in order to construct a uniformly valid solution.

Before we consider the need for matched asymptotics, let us naively propose a regular asymptotic expansion for $y$ in powers of the tethering parameter, $s$. This yields

$$y(x, t; s) \sim y_0(x, t) + s y_1(x, t) + \ldots,$$

where $y_i$ are all of order unity. Substituting into equation (3.39), this yields

$$\frac{\partial^3 y_0}{\partial x^2 \partial t} + s \frac{\partial^3 y_1}{\partial x^2 \partial t} + \ldots - s \left( y_0 + s y_1 + \ldots \right) = \frac{-1}{2 \sqrt{\pi t}} \left[ e^{-\frac{(x-1)^2}{4t}} - e^{-\frac{(x+1)^2}{4t}} \right].$$

Collecting the terms of order unity, we find that

$$\frac{\partial^3 y_0}{\partial x^2 \partial t} = \frac{-1}{2 \sqrt{\pi t}} \left[ e^{-\frac{(x-1)^2}{4t}} - e^{-\frac{(x+1)^2}{4t}} \right],$$

\hspace{1cm} (3.44)

\textsuperscript{24}Specifically, we recall that $s \neq 0$ implies a stronger boundary condition at infinity than $s = 0$. Hence, the behaviour of $y$ observed at large $x$ is very different in the case where $s$ is zero compared with the case where $s$ is small but nonzero.
while the higher order terms yield equations of the form

\[ \frac{\partial^2 y_i}{\partial x^2 \partial t} = y_{i-1}(x, t), \quad (3.45) \]

where \( i = 1, 2, 3, \ldots \)

Equations (3.44) and (3.45) suggest that it is possible to obtain successive terms in the asymptotic series by repeated integration. For example, we can integrate equation (3.44) twice with respect to \( x \) and once with respect to \( t \) to obtain the result that

\[ y_0(x, t) = \frac{1}{2}(y^*(x + 1, t) - y^*(x - 1, t)) + C_1(t) x + C_2(t) + C_3(x), \quad (3.46) \]

where \( y^*(x, t) \) is defined by

\[ y^*(x, t) = (x t + \frac{1}{6} x^3) \text{ erf}\left(\frac{x}{2\sqrt{t}}\right) + (\frac{4}{3} t^2 + \frac{1}{3} x^2 t) \frac{1}{\sqrt{\pi} t} e^{-\frac{x^2}{4t}}, \quad (3.47) \]

and \( C_1(t) \), \( C_2(t) \) and \( C_3(x) \) are analogous to constants of integration. We will now use conditions (3.40) through to (3.42) to determine these three functions.

Without loss of generality, let \( C_1(0) = 0 \) and \( C_2(0) = 0.25 \) 25 Now, consider the boundary condition given in (3.41):

\[ 0 = y_0(0, t) = \frac{1}{2}(y^*(1, t) - y^*(-1, t)) + C_2(t) + C_3(0). \quad (3.48) \]

Interestingly, we can see from the definition of \( y^*(x, t) \) in equation (3.47) that

\[ y^*(-x, t) = y^*(x, t); \]

that is, for any choice of \( t \) we find that \( y^* \) is an even function of \( x \). Thus, (3.48) becomes

\[ C_2(t) = -C_3(0), \]

25Effectively, this means that all of the behaviour of \( y_0 \) at \( t = 0 \) is absorbed into \( C_3(x) \). Making the requirement that \( C_1(0) = C_2(0) = 0 \) allows us to avoid carrying redundant constants of integration through our algebra.
meaning that \( C_2 \) is constant for all time. Since \( C_2(0) = 0 \), it follows that \( C_2(t) \equiv 0 \) and hence \( C_3(0) = 0 \).

Next, consider the initial condition given in (3.40). Substituting in the definition of \( y_0 \) from (3.46), we find that

\[
0 = y(x, 0) = \frac{1}{2}(y^*(x + 1, 0) - y^*(x - 1, 0)) + C_3(x). \tag{3.49}
\]

Furthermore, we note that

\[
\lim_{t \to 0} \frac{1}{\sqrt{\pi t}} e^{-x^2/4t} = 2 \delta(x),
\]

where \( \delta(x) \) is the Dirac delta function. Thus, it follows that

\[
y^*(x, 0) = \begin{cases} 
-\frac{1}{6} x^3, & x < 0, \\
\frac{1}{6} x^3, & x > 0;
\end{cases}
\]

and we find that equation (3.49) becomes

\[
C_3(x) = \begin{cases} 
\frac{-1}{2} \left( \frac{(x+1)^3}{6} + \frac{(x-1)^3}{6} \right), & x < -1, \\
\frac{-1}{2} \left( \frac{(x+1)^3}{6} - \frac{(x-1)^3}{6} \right), & -1 < x < 1, \\
\frac{-1}{2} \left( \frac{(x+1)^3}{6} - \frac{(x-1)^3}{6} \right), & x > 1;
\end{cases}
\]

or, equivalently,

\[
C_3(x) = \begin{cases} 
\frac{1}{2} \left( x^2 + \frac{1}{3} \right), & x < -1, \\
\frac{1}{2} \left( - \frac{1}{3} x^3 - x \right), & -1 < x < 1, \\
\frac{1}{2} \left( - x^2 - \frac{1}{3} \right), & x > 1.
\end{cases} \tag{3.50}
\]

Lastly, consider the boundary condition given in equation (3.42). This takes the
form

$$0 = \lim_{x \to \infty} \frac{\partial y_0}{\partial x} = \lim_{x \to \infty} \left( \frac{1}{2} \left( y_x^*(x + 1, t) - y_x^*(x - 1, t) \right) + C_1(t) + C_3''(x) \right), \quad (3.51)$$

where the subscript $x$ represents differentiation. Now, differentiating equation (3.47) yields

$$y_x^*(x, t) = \left( t + \frac{1}{2} x^2 \right) \operatorname{erf} \left( \frac{x}{2 \sqrt{t}} \right) + (xt + \frac{1}{6} x^3) \frac{1}{\sqrt{\pi \ell}} e^{-\frac{x^2}{4t}}$$

$$+ \frac{2}{3} xt \frac{1}{\sqrt{\pi \ell}} e^{-\frac{x^2}{4t}} - \frac{x}{2t} \left( \frac{1}{3} t^2 + \frac{1}{3} x^2 \right) \frac{1}{\sqrt{\pi \ell}} e^{-\frac{x^2}{4t}}.$$

Fortunately, it is possible to simplify this in the case where $x$ is large. Since $e^{-\frac{x^2}{4t}}$ becomes transcendentally small as $x$ approaches infinity, we find that

$$\frac{\partial y_x^*}{\partial x} \sim \left( t + \frac{1}{2} x^2 \right) \operatorname{erf} \left( \frac{x}{2 \sqrt{t}} \right),$$

for large $x$.

From equation (3.50), we note that $C_3''(x) = -x$ for all $x > 1$. Using this result and the expression for $\frac{\partial y_x^*}{\partial x}$ given above, we find that equation (3.51) becomes

$$C_1(t) = \lim_{x \to \infty} \left( \frac{1}{2} \left( t + \frac{1}{2} (x + 1)^2 \right) \operatorname{erf} \left( \frac{x + 1}{2 \sqrt{t}} \right) - \frac{1}{2} \left( t + \frac{1}{2} (x - 1)^2 \right) \operatorname{erf} \left( \frac{x - 1}{2 \sqrt{t}} \right) - x \right). \quad (3.52)$$

Now, let $a$ and $b$ be real numbers, and let $p(x)$ be a polynomial. By repeated use of l'Hôpital's rule, it is possible to show that

$$\lim_{x \to \infty} p(x) \left( 1 - \operatorname{erf}(ax + b) \right) = 0,$$

whenever $a > 0$. That is, $\operatorname{erf}(x)$ approaches 1 faster than any polynomial ap-
approaches infinity. Using the algebra of limits, it follows that

\[ \lim_{{x \to \infty}} p(x) \operatorname{erf}(ax + b) = \lim_{{x \to \infty}} p(x). \]

Thus, equation (3.52) yields

\[ C_1(t) = \lim_{{x \to \infty}} \left( \frac{1}{2} t + \frac{1}{4} (x + 1)^2 - \frac{1}{2} t - \frac{1}{3} (x - 1)^2 - x \right) = 0. \]

Importantly, we note that we were able to cancel all of the terms that diverge as \( x \) tends to infinity. If this were not the case, we would have had to conclude that it is impossible to find a solution to (3.44) that is consistent with the stated initial and boundary conditions.

For example, consider what would happen if we attempted to determine \( C_1(t) \) using the alternative boundary condition, (3.43), instead of condition (3.42). In this case, we would require that

\[ 0 = \lim_{{x \to \infty}} y_0(x, t) = \frac{1}{2} \left( y^*(x + 1, t) - y^*(x - 1, t) \right) + C_1(t) x + C_3(x), \quad (3.53) \]

where \( C_3(x) \) is defined by equation (3.50).

Using the same approach described earlier, we find that

\[ y^*(x, t) \sim x t + \frac{1}{6} x^3 \]

as \( x \) grows large. Thus, equation (3.53) becomes

\[ 0 = \lim_{{x \to \infty}} \left( \frac{1}{2} (x + 1) t + \frac{1}{12} (x + 1)^3 - \frac{1}{2} (x - 1) t - \frac{1}{12} (x - 1)^3 + C_1(t) x - \frac{1}{2} x^2 - \frac{1}{6} \right), \]

which yields

\[ 0 = \lim_{{x \to \infty}} (t + C_1(t) x). \]

This is impossible to satisfy. The only way to make this limit converge is to have
\[ C_1(t) \equiv 0, \text{ but this leads to} \]

\[ \lim_{x \to \infty} y_0(x, t) = t, \quad (3.54) \]

not

\[ \lim_{x \to \infty} y_0(x, t) = 0. \]

This is a very interesting result. Although we were able to use condition (3.42) to find \( y_0(x, t) \), the solution that we obtained does not satisfy the stronger condition given in (3.43). This makes sense, as we have already shown that condition (3.43) only follows from condition (3.42) when \( s \) is nonzero. However, the fact that our leading order perturbation approximation does not satisfy the condition (3.43) is problematic for two reasons. Firstly, we note that the leading-order asymptotic solution does not compare well with the numeric solution when \( x \) is large (this is illustrated in Figure 3.11). Not only does the asymptotic solution fail to tend to zero, but we also see that it does not display the oscillatory behaviour that we are interested in.

Secondly, the fact that \( y_0 \) does not tend to zero in the limit as \( x \to \infty \) means that no further corrections in the regular perturbation expansion can be obtained. By combining equation (3.45) with conditions (3.40), (3.41) and (3.42), we obtain the following integral expression for \( y_1(x, t) \):

\[ y_1(x, t) = \int_0^t \int_0^x \int_{x^*}^{\infty} y_0(x^{**}, t^*) \, dx^{**} \, dx^* \, dt^*. \]

A necessary condition for the convergence of this improper integral is

\[ \lim_{x^{**} \to \infty} y_0(x^{**}, t) = 0; \]

hence, it follows that \( y_1(x, t) \) cannot be obtained using this technique when \( t > 0 \).

Thus, we are unable to obtain a uniformly valid approximation for \( y \) using regular perturbation techniques, despite the fact that there is nothing in equation (3.39)
Figure 3.11: Comparison of $y(x, t)$ from the numeric solution of equation (3.39) with the short space asymptotic approximation given by $y_0(x, t)$. The numeric solution, $y(x, t)$ is shown using continuous profiles while the asymptotic approximation, $y_0(x, t)$, is shown using dotted profiles. Note that $y_0(x, t)$ can be found by substituting the expressions obtained for $y^*(x, t)$, $C_1(t)$, $C_2(t)$ and $C_3(x)$ into equation (3.46). The parameter value used is $s = 1$.

As previously, the initial condition is shown as a dashed line. The profiles corresponding to $t = 0.05$ are shown as very thick lines, the profiles corresponding to $t = 0.1$ are shown as thick lines and subsequent profiles (normal thickness) represent time increasing in steps of 0.1 up to $t = 0.4$. 

\[ y(x, t) \times 10^{-2} \]
that would immediately lead us to suspect singular behaviour. Specifically, we observe that the approximation that we obtained breaks down at large values of $x$ and that we are unable to find the first correction (or any subsequent corrections) using regular perturbation methods.

This situation is analogous to Whitehead’s paradox, a classic problem in which regular perturbation methods fail to yield uniformly valid solutions for slow viscous flow around a cylinder or sphere [135, 197]. This difficulty was resolved by Proudman and Pearson [153], who demonstrated that a regular asymptotic expansion in small Reynolds number becomes invalid far away from the cylinder or sphere. In order to describe the flow in this region, it is necessary to rescale the problem, ultimately leading to the solution originally proposed by Oseen.

Effectively, the problem of slow flow around a body involves a ‘boundary layer at infinity’. Like a normal boundary layer problem, we are unable to apply all of the boundary conditions unless we take into consideration the multiple scales inherent in the problem. Boundary layers are typically associated with problems where the highest derivative is multiplied by a small parameter, but this is not actually necessary. As described by Ockendon and Ockendon [135], boundary layer behaviour can also occur in problems where boundary conditions are applied at either end of a large range, as in the problem of slow flow around a cylinder or sphere.

This is precisely what is happening in our problem. Even though there is no small parameter multiplying the highest derivative, we note that $x$ is defined on an infinite domain. Furthermore, the first term of our regular asymptotic expansion fails to satisfy the ‘strong’ boundary condition as $x \to \infty$. Thus, we are

---

\[26\text{In the case of a cylinder, it is impossible to use regular asymptotic methods to construct a leading order solution that satisfies the required boundary condition at infinity. In the case of a sphere, it is possible to satisfy the boundary condition at infinity but the first correction does not converge. Our problem has aspects of both of these cases. As with the problem of slow flow around a cylinder, we find that the leading order perturbation approximation does not match the actual solution at large values of } x. \text{ However, we can still satisfy the ‘weak’ boundary condition (3.42) so, in a sense, regular perturbation methods can be used to obtain a leading order perturbation approximation. The problem is then that the first correction does not converge (as in the case of slow flow around a sphere).}\]
justified in using boundary layer methods (i.e. matched asymptotic expansions) to attempt a uniformly valid approximation of the solution to (3.39).

The first step in doing this is to identify the rescaling associated with the boundary layer. Since the boundary layer behaviour appears at large values of $x$, we propose that the rescaling will take the form

$$x = \hat{X} X,$$

where $X \sim O(1)$ in the region of interest and $\hat{X}$ is large. Substituting this into equation (3.39), we find that

$$\frac{1}{X^2} \frac{\partial^3 y}{\partial X^2 \partial t} - sy = \frac{-1}{2\sqrt{\pi} t} \left[ e^{-\frac{(\hat{X} - 1)^2}{4t}} - e^{-\frac{(\hat{X} + 1)^2}{4t}} \right].$$

Since $\hat{X}$ is large, we find that the right hand side of this equation is transcendentally small. Hence, the distinguished limit occurs when the two terms on the left hand side of the equation are in balance. By the principle of least degeneracy, this means that $\hat{X} = s^{-\frac{1}{2}}$ will be an appropriate rescaling.\(^\text{27}\)

Using $Y(X, t)$ to represent the solution for $y$ in the rescaled region and neglecting transcendentally small terms, we find that rescaling yields

$$\frac{\partial^3 Y}{\partial X^2 \partial t} - Y = 0. \quad (3.55)$$

Equations (3.44) and (3.45) are correctly scaled for the case where $x \sim O(1)$, while equation (3.55) is valid where $x \sim O(s^{-\frac{1}{2}})$. We will refer to $y(x, t)$ as the inner solution and $Y(X, t)$ as the outer solution.

Interestingly, the fact that the boundary layer behaviour occurs when $x \sim O(s^{-\frac{1}{2}})$ suggests that we should use expansions in $\sqrt{s}$ rather than $s$. That is, we formally

\(^\text{27}\)At this stage, we are assuming that $t \sim O(1)$. As we will see, we encounter another change to the terms of the dominant balance at large times.
propose that
\[ y(x, t; s) \sim y_0(x, t) + y_1(x, t) \sqrt{s} + y_2(x, t) s + \ldots, \]
\[ Y(X, t; s) \sim Y_0(x, t) + Y_1(X, t) \sqrt{s} + Y_2(X, t) s + \ldots; \]
where \( y_i(x, t) \) and \( Y_i(x, t) \) are all of order unity. In the work that follows, we will only consider the leading order problem. Thus, the scale needed for the correction terms does not actually make any significant difference to our working.

When we come to match our inner and outer solutions, we will use Van Dyke’s matching condition.\(^{28}\) This states that the \( m \)-term outer limit of the \( n \)-term inner solution must be equal to the \( n \)-term outer limit of the \( m \)-term inner expansion. At leading order, this simplifies to Prandtl’s matching condition; that is,
\[ \lim_{x \to \infty} y_0(x, t) = \lim_{X \to 0} Y_0(X, t) = y_m(t), \]
where \( y_m(t) \) is the matching function, which is needed when we construct a uniformly valid composite solution. We will use (3.56) to supply the ‘missing’ boundary conditions that we need to uniquely determine \( y_0 \) and \( Y_0 \).

Now, collecting the leading order terms of equation (3.39) yields
\[ \frac{\partial^2 y_0}{\partial x^2 \partial t} = \frac{-1}{2\sqrt{\pi} t} \left[ e^{-\frac{(x-1)^2}{4t}} - e^{-\frac{(x+1)^2}{4t}} \right]. \]
Applying the left hand boundary condition and the initial condition (but not the right hand boundary condition), we can use our earlier working to show that this has a solution of
\[ y_0(x, t) = \frac{1}{2} \left( y^*(x + 1, t) - y^*(x - 1, t) \right) + C_1(t) x + C_3(x), \]
\(^{28}\)It is well known that Van Dyke’s matching condition fails in situations where the order of the correction terms is related to the logarithm of a function of the small parameter. However, since it is reasonable to expect that an expansion in \( \sqrt{s} \) will be appropriate for our problem, we do not expect this difficulty to arise.
where \( y^*(x, t) \) is defined in equation (3.47) and \( C_3(x) \) is defined in equation (3.50), but \( C_1(t) \) is unspecified.

From our earlier work, we recall that \( y_0(x, t) \) will be unbounded as \( x \to \infty \) unless \( C_1(t) \equiv 0 \). Thus, condition (3.56) effectively requires that \( C_1(t) \equiv 0 \). In this case, we can use equation (3.54) to show that

\[
y_m(t) = \lim_{x \to \infty} y_0(x, t) = t.
\]

Now, the leading order terms of equation (3.55) give the result that

\[
\frac{\partial^3 Y_0}{\partial X^2 \partial t} - Y_0 = 0.
\]  

(3.58)

Furthermore, we can use (3.40), (3.42) and (3.56) to obtain appropriate initial and boundary conditions as follows:

\[
Y_0(X, 0) = 0, \quad (3.59)
\]

\[
Y_0(0, t) = y_m(t) = t, \quad (3.60)
\]

\[
\lim_{X \to \infty} Y_0(X, t) = 0. \quad (3.61)
\]

We now wish to solve this system.

Firstly, we note that equation (3.58) is linear.\(^{29}\) This means that classical integral transform methods may lead to useful results. For example, consider what happens when we take the Laplace transform of equation (3.58) with respect to \( t \). Using \( \tilde{t} \) to denote the transformed variable, we find that

\[
\tilde{t} \frac{\partial^2 \tilde{Y}}{\partial X^2} - Y_0(X, 0) - \tilde{Y}(X, \tilde{t}) = 0,
\]

\(^{29}\) As noted earlier, equation (3.39) was also linear. However, we found that using standard integral transforms on (3.39) did not lead to interpretable results. Thus, it is necessary to use the perturbation approach that we have taken.
where $\mathcal{L}\{Y_0(X, \tilde{t})\} = \tilde{Y}(X, \tilde{t})$. Since $Y_0(X, 0) \equiv 0$, it follows that

$$\frac{\partial^2 \tilde{Y}}{\partial X^2} - \frac{1}{\tilde{t}} \tilde{Y}(X, \tilde{t}) = 0.$$  

This is effectively an ordinary differential equation in $X$, and we can obtain boundary conditions by transforming conditions (3.60) and (3.61) as follows:

$$\tilde{Y}(0, \tilde{t}) = \frac{1}{\tilde{t}^2},$$

$$\tilde{Y}(\infty, \tilde{t}) = 0.$$  

Thus, we find that solving the differential equation yields

$$\tilde{Y}(X, \tilde{t}) = \tilde{t}^{-2} \exp\left(-X \tilde{t}^{-\frac{1}{2}}\right).$$  

(3.62)

We now need to invert this in order to obtain a solution for $Y_0(X, t)$. Unfortunately, we were unable to find anything similar to $\tilde{Y}$ on standard tables of Laplace transforms. Thus, it would seem that we need to use the Bromwich integral to proceed further. Alternatively, we can expand the exponential function as a Maclaurin series and then invert term by term. That is, we note that

$$\tilde{Y}(X, \tilde{t}) = \tilde{t}^{-2} \exp\left(-X \tilde{t}^{-\frac{1}{2}}\right)$$

$$= \tilde{t}^{-2} \left(1 - X \tilde{t}^{-\frac{1}{2}} + \frac{1}{2} X^2 \tilde{t}^{-1} - \ldots\right)$$

$$= \tilde{t}^{-2} - X \tilde{t}^{-\frac{5}{2}} + \frac{1}{2} X^2 \tilde{t}^{-3} - \ldots.$$
Since $\mathcal{L}^{-1}(\hat{f}^r) = \frac{1}{\Gamma(r)} t^{r-1}$, we ultimately find that

$$Y_0(X, t) = t - \frac{X t^{3/2}}{\Gamma(3/2)} + \frac{X^2 t^2}{\Gamma(3) 2!} - \cdots$$

$$= t \left(1 - \frac{X \sqrt{t}}{\Gamma(3/2)} + \frac{(X \sqrt{t})^2}{\Gamma(3) 2!} - \cdots\right)$$

$$= t \iota(X \sqrt{t}),$$

where $\iota(z)$ is defined by

$$\iota(z) = \sum_{n=0}^{\infty} \frac{(-z)^n}{\Gamma(2 + \frac{z}{2}) n!}. \quad (3.63)$$

This series definition of $\iota(z)$ would initially appear to be uninformative. However, we can use equation (3.63) to show that $\iota(z)$ is in the appropriate form to be a Wright function. As described by Gorenflo et al. [68], the Wright function is defined by

$$\phi(\rho, \beta; z) = \sum_{k=0}^{\infty} \frac{z^n}{\Gamma(\rho k + \beta) k!}, \quad (3.64)$$

where $\rho > -1$ is a real number and $\beta \in \mathbb{C}$. It is interesting to note that the Wright function is an entire function of $z$, but we will only concern ourselves with the case where $z$, $\rho$ and $\beta$ are all real.

By inspection of equation (3.63), we find that

$$\iota(z) = \phi(\frac{1}{2}, 2; -z).$$

Gorenflo et al. [68] give several analytical results concerning the Wright function, one of which is particularly useful to us. For $\rho > 1$ and $\beta \in \mathbb{R}$, an asymptotic expansion of $\phi(\rho, \beta; z)$ exists when $z$ is on the negative real semi-axis and $|z|$ is

---

30The Wright function was originally developed in the 1930s in connection with the asymptotic theory of partitions and it has ongoing applications in the area of fractional calculus. A more detailed history of the Wright function and a description of many analytical results (including all of the results that we use here) can be found in Gorenflo et al. [68].
large. This takes the form

\[
\phi(\rho, \beta; -x) \sim x^p \left( \frac{1}{\pi^p} \exp \left( \sigma x^p \cos(\pi p) \right) \cos \left( \pi p \left( \frac{1}{2} - \beta \right) + \sigma x^p \sin(\pi p) \right) \right) \cos \left( \pi p \left( \frac{1}{2} - \beta \right) \right) + O(x^{-p}) \times (c_1 + O(x^{-p})) ,
\]

where \( p = \frac{1}{1 + \rho} \), \( \sigma = (1 + \rho) \rho^{-(1-p)} \) and \( c_1 \) can be determined using another asymptotic result stated in Gorenflo et al. [68] (not given here).

In our case, we have \( \rho = \frac{1}{2} \) and \( \beta = 2 \). Thus, it follows that \( p = \frac{2}{3} \), \( \sigma = \frac{3^\frac{3}{2} \sqrt{2}}{4} \) and \( c_1 = \frac{4}{\sqrt{3} \pi} \). Rearranging, we find that

\[
\iota(z) \sim c z^{-1} \exp \left( -a z^{\frac{3}{2}} \right) \cos \left( b z^{\frac{3}{2}} \right) + O(z^{-\frac{5}{3}} \exp(-a z^{-\frac{5}{3}})),
\]

where \( a = \frac{3^\frac{3}{2} \sqrt{2}}{4} \), \( b = \frac{3^\frac{3}{2} \sqrt{3}}{4} \) and \( c = -\frac{4}{\sqrt{3} \pi} \).

Having obtained an interpretable approximation of \( \iota(z) \) for large \( z \), we now wish to work backwards to determine what this says about the solution of (3.39).

Firstly, we recall that

\[
Y_0(X, t) = t \iota(X \sqrt{t}).
\]

Hence, for sufficiently large values of \( X \sqrt{t} \), we find that

\[
Y_0(X, t) \sim \frac{c \sqrt{t}}{X} e^{-a \frac{3}{2} X^{\frac{3}{2}} t} \cos \left( b \frac{3}{2} X^{\frac{3}{2}} t \right),
\]

where \( a, b \) and \( c \) are defined as above.

Now, \( Y_0(X, t) \) is the leading order term in outer solution of equation (3.39). That is, \( Y_0(X, t) \) is a valid approximation for \( y(x, t) \) when \( X = x \sqrt{s} \) is of order one. This means that

\[
y(x, t) \sim \frac{c \sqrt{t}}{x \sqrt{s}} e^{-a \frac{\sqrt{s}}{\sqrt{x^2}} t} \cos \left( b \frac{\sqrt{s}}{\sqrt{x^2}} x^2 t \right)
\]

is a valid asymptotic approximation for \( y(x, t) \) if we can satisfy the following three conditions:
• $s$ is small,
• $x$ is of the same order as $s^{-\frac{1}{2}}$ or larger, and
• $x \sqrt{s} \sqrt{t}$ is large.

Furthermore, we recall that $y(x, t) = u_1(x, t)$, which is the first nonzero term in the asymptotic expansion of $u(x, t)$, the solution of equation (3.27). Thus, equation (3.67) implies that

$$u(x, t) \sim \frac{\tau c \sqrt{t}}{x \sqrt{s}} e^{-a \sqrt{s} x^2 t} \cos (b \sqrt{s} x^2 t),$$

(3.68)

if all of the necessary conditions are satisfied. This is illustrated in Figure 3.12.

Figure 3.12: Numeric simulations of $u(x, t)$ from the caricature model (i.e. equations (3.25) to (3.27)) compared with the asymptotic approximation given in equation (3.68). Numeric solutions to the caricature model are shown as continuous lines while the asymptotic approximations are shown as dotted lines. The parameter values used are $\tau = 0.5$ and $s = 1$. In order to ensure stability of the numeric simulations, a finer spatial mesh with points regularly spaced at intervals of 0.025 was found to be necessary.

As previously, the initial condition is shown as a dashed line. The profiles corresponding to $t = 0.05$ are shown as very thick lines, the profiles corresponding to $t = 0.1$ are shown as thick lines and profiles representing $t = 0.2$ and $t = 0.4$ are shown in normal thickness. Note that the approximation improves at larger values of $x$. 
These results give us a mathematical explanation of the decaying oscillations that we observed in the numeric simulations of \( y(x, t) \) and \( u(x, t) \) at large values of \( x \). However, we note that our analytic description of the oscillatory behaviour is quite peculiar; the oscillations are not periodic and the decay in amplitude is not exponential. Our result is highly counterintuitive and it is unlikely that we would have been able to obtain such a good description of the oscillations using a different approach (e.g. by proposing an \textit{ansatz} solution).

Now, let us combine the outer and inner solutions that we have developed for \( y_0(x, t) \). Using standard matching procedures, we find that

\[
y_0^{\text{matched}}(x, t) = y_0^{\text{inner}}(x, t) + y_0^{\text{outer}}(x, t) - y_m(t) = \frac{1}{2}(y^*(x + 1, t) - y^*(x - 1, t)) + C_3(x) + t \left( x \sqrt{s t} \right) - t, \tag{3.69}
\]

where \( y^*(x, t), C_3(x) \) and \( \iota(z) \) are defined in equations (3.47), (3.50) and (3.63) respectively.

Hypothetically, we should be able to use equation (3.69) to find the maximum value of \( y_0(x, t) \). If this is significantly less than unity, it would help us to explain why \( u(x, t) \) was consistently much smaller than \( \tau \) in the numeric simulations of our caricature model. Furthermore, it would allow us to develop a better rescaling of \( u(x, t) \) that might be applicable to the full Tranquillo-Murray model.

Based on our numeric simulations, we expect that the maximum value of \( y(x, t) \) will occur when \( x \approx 1 \) and \( t \) is small. However, we cannot use the approximation of \( \iota(z) \) given in equation (3.66) in this case; it is only valid when \( x \sqrt{t} \) is large. Thus, any attempt to find the maximum of \( y_0^{\text{matched}} \) in equation (3.69) would need to use the series definition of \( \iota(z) \). As several terms in the series need to be taken before a decent approximation is obtained, this is difficult and relatively uninformative. Although we note that a bound on \( u \) is theoretically possible, we will not proceed any further with this line of analysis.

Lastly, it is interesting to note that the matched solution given in equation (3.69)
is not uniformly valid for all values of $x$ and $t$. In particular, we encounter problems at long times. Just as $x \sim O(s^{-\frac{1}{2}})$ involved a change of dominant balance that required us to use matched asymptotics, we see a similar effect when $t \sim O(s^{-1})$. However, at long times it is not obvious that the terms associated with the right hand side of equation (3.39) can be neglected. Furthermore, it would be much more difficult to use matched asymptotics in this case because we would need to match the long time solution with both parts of the spatially-matched short time solution.

Fortunately, it is not necessary to go to this depth. We were mainly interested in using asymptotic analysis to explain two unusual features of the Tranquillo-Murray model: the smallness of displacement and the displacement oscillations. Both of these are principally related to the short time behaviour of the Tranquillo-Murray model and hence there is no need for us to extend our analysis to long times.

In conclusion, we have used asymptotic methods to show that a caricature Tranquillo-Murray model exhibits oscillations in displacement in the region far from the wound. This strengthens our hypothesis that the oscillations observed in the Tranquillo-Murray model are analytic rather than numeric. Furthermore, the fact that the oscillations are observed in a viscous-only model supports the theory that the oscillations are connected to dermal viscosity.

In Appendix A, we extend our analysis further, considering a more general model of dermal wound healing. Ultimately, we are able to show that all linear viscoelastic models of dermal wound healing (including the Tranquillo-Murray model) lead to displacement oscillations. Again, we find that these oscillations are related to the viscous behaviour of the dermis. Interestingly, our results also suggest that the time scale of the displacement oscillations is governed by the time scale of viscous relaxation.

Having shown that displacement oscillations are an intrinsic feature of viscoelastic models of dermal wound healing, we would like to be able to give them some
physical explanation. However, it does not seem realistic for the entire skin to wrinkle in response to a healing wound as the models suggest. There is no obvious way to reconcile the mathematical model with physical observations.

In the following section, we resolve this problem by showing that the viscous time scale of the dermis is extremely short compared to the time scale of wound healing. As we will see, Tranquillo and Murray (and later Olsen et al.) hugely overestimated the parameter $\mu$, leading to the apparent discrepancy between model and reality. On the time scale of wound healing, the ECM effectively behaves as a purely elastic material.

3.4 Elasticity vs. viscoelasticity

3.4.1 Experimental measurements of dermal viscoelasticity

Many of the existing mechanical models of dermal wound healing use a viscoelastic constitutive law to describe the mechanics of the ECM. For example, the Tranquillo-Murray model [189] and the Olsen models [136, 137, 140, 141] treat the ECM as a Kelvin-Voigt linear viscoelastic solid. However, the use of viscoelasticity is not universal; as discussed in Section 2.3.2, Cook [33] hypothesised that much of the ‘viscous’ behaviour observed in dermal wound healing is actually related to matrix turnover. If this is the case, a viscoelastic model is inappropriate and it would be better to model the ECM as an elastic solid with an evolving zero stress state. Mathematical techniques for achieving this are described in the following chapters.

At the time that Tranquillo, Murray, Olsen and Cook were working on models of dermal wound healing, very few data on the viscoelastic properties of human skin were available. Over the past decade, this situation has improved significantly as
a more quantitative research on dermal biomechanics has been published [171]. Unfortunately, there is a lack of uniformity in experimental reporting. Given the wide range of experimental approaches used and the variety of definitions of ‘viscoelasticity’, it is difficult to find data that are directly comparable with the mathematical models presented here.

For example, let us contrast two different approaches to measuring dermal viscoelasticity. Firstly, consider the results obtained by Silver et al. [170]. This group carried out in vitro experiments on human skin samples, considering large strains (up to 1) and relatively long times (on the scale of minutes). In their experiments, the dermal samples were gradually strained to predefined levels and the stress response was recorded. For each imposed strain, two stress measurements were made: one as soon as the required strain was reached and one after any viscous relaxation had taken place. The stress after relaxation was referred to as the ‘elastic stress’, while any difference between the earlier stress measurement and the elastic stress was referred to as the ‘viscous stress’.

This approach would appear to be consistent with the Kelvin-Voigt model of viscoelasticity used by Tranquillo and Murray. However, we actually find that the results obtained by Silver et al. are not directly comparable with the results predicted by the Tranquillo-Murray model. One reason for this is that Silver and coworkers do not observe any significant viscous stresses until a strain of 0.4 is reached [170]. This is well outside the range of strains observed in the Tranquillo-Murray model and is much greater than the strains that we would expect to see in human wound healing.

The fact that Silver and coworkers first observe viscous behaviour at such a large value of strain suggests that they are not actually measuring the reversible viscoelastic behaviour that would be described by a Kelvin-Voigt model. Instead, it seems reasonable to assume that Silver et al. exceeded the yield stress of the material that they were studying. The viscous relaxation observed may actually be the plastic flow associated with the internal rearrangement of ECM fibres at
high stress. If this is the case, it would be more appropriate to refer to the difference between the two stress measurements as the residual stress, not the ‘viscous’ stress.

In fact, it is interesting to note that this change would make it easier to justify the mathematical analysis given in Silver et al. [170]. If Silver and coworkers were dealing with a truly viscoelastic material, it would be important for them to pay careful attention to the relationship between stress and the rate of strain. However, Silver et al. do not discuss the time-dependent evolution of stress in much detail at all; most of their results are presented in terms of the relationship between strain and stress, ignoring time-dependent evolution altogether. Although this is acceptable in the analysis of residual stresses, the stress-strain plots given by Silver et al. are of limited use in any formal analysis of viscoelastic behaviour.

Unlike Silver et al., ŠDoubal and coworkers [50, 105] investigated dermal mechanics using in vivo experiments. These involved following the evolution of small stresses and strains over a number of seconds. Importantly, ŠDoubal and coworkers did not attempt to construct a static relationship between stress and strain. Instead, they directly fitted a viscoelastic mathematical model to the time series data that they obtained from their experiments.

Interestingly, this led ŠDoubal and coworkers to the conclusion that there are two different time scales inherent in the mechanical behaviour of the skin. At very short times (≈10 ms), it is necessary to consider elastic, viscoelastic and inertial effects. At longer times, the inertial effects become negligible and it is appropriate to use a classical Kelvin-Voigt model.

---

31In order to confirm this hypothesis, we would need to repeat Silver et al.’s experiments and collect unloading data as well as loading data. As we are only given loading data, we have no way of distinguishing plastic flow from viscous flow.

32See, for example, Krishnan and Rajagopal [104]:
‘In viscoelasticity, for 1D response it only makes sense to plot stress versus time and strain versus time. A stress versus strain plot makes no sense because an infinite number of them can be plotted for the same response. While, at a fixed time, we can find the ratio of the value of stress to that of the strain, the ratio has no physical meaning. Extrapolating from linearized elasticity to call that ratio the “stiffness” does not lead to a meaningful description of the material.’ – Krishnan and Rajagopal (2003) [104] p172
The results obtained by Ďoubal and coworkers are presented in terms of the parameters of a linear Kelvin-Voigt model (with or without inertial effects). Thus, we can use their results directly to determine an appropriate value for the dimensionless parameter \( \mu \) of the Tranquillo-Murray and Olsen models. As described in Kuchařová et al. [105], typical parameter values for their ‘long time’ model (i.e. without inertial effects) are

\[
N = 651.24 \text{ kg s}^{-1}, \quad H = 223.4 \text{ kN m}^{-1};
\]

where \( N \) is the effective coefficient of viscosity and \( H \) is the effective coefficient of elasticity.

Using the same nondimensionalisation as Tranquillo and Murray, we find that the dimensionless parameter \( \mu \) is given by \( \mu = \frac{N}{HT} \), where \( T \) is the time scale used in the model nondimensionalisation. If we follow Olsen et al. [140] and take \( T \) to be one day, we find that

\[
\mu \approx 3.3 \times 10^{-8}.
\]

This is very different from the values of \( \mu \) used in the Tranquillo-Murray and Olsen models. As we have seen, Tranquillo and Murray focused their attention on the case where \( \mu = 1 \). More interestingly, Olsen et al. fitted their model to experimental data and obtained the result that \( \mu = 20 \), nine orders of magnitude larger than the experimental value given above.

There is a simple reason for the huge discrepancy between the experimental value of \( \mu \) and the value that Olsen and coworkers obtained from fitting their model to data. As noted by Cook [33], much of the apparently viscous behaviour of a healing wound can be attributed to the internal remodelling of the ECM. However, the Olsen models do not include any mathematical representation of ECM remodelling. Thus, when Olsen and coworkers attempted to fit their model to experimental data, they obtained an apparent dermal viscosity that was much larger than it should have been. The value \( \mu = 20 \) stated by Olsen et al. reflects
the ‘pseudoviscous’ behaviour associated with remodelling more than the true dermal viscosity.

Although they may yield some similar results, a morphoelastic model of wound healing (e.g. Cook’s model) will generally behave differently from a viscoelastic model. One important point of difference is that the displacement oscillations analysed in Section 3.3 are associated with viscoelasticity, not morphoelasticity. If we use a small value of $\mu$ instead of a large value of $\mu$, we find that any displacement oscillations will settle in the first few moments of wound healing.\(^{33}\)

Furthermore, we find that we can finally give the displacement oscillations a physical interpretation. Given that the viscous relaxation of the dermis happens over the course of a few hundredths of a second, it is reasonable to expect that the displacement oscillations will occur over a similar time scale. If the skin is pinched and released, it is interesting to note that movement can be felt (and sometimes seen), which suggests small vibrations spreading from the site of the pinch. Possibly, these vibrations are equivalent to the oscillations predicted by a viscoelastic model of the dermis.

The experimental value of $\mu$ given above is so small as to be practically zero. If we were to consider a perturbation in small $\mu$, we would find that equation (3.3) predicts that there will be an initial transient (i.e. a boundary layer in short time). However, after a very short period of time we find that the viscous term can be ignored completely.\(^{34}\)

Since we are mainly concerned with the progression of wound healing over days

\(^{33}\)This can be justified using the analysis given in Appendix A.

\(^{34}\)This is consistent with the discussion of biological time scales in Goriely and Ben Amar [70]. They propose that there are four important time scales for growing tissues. In decreasing order of length, these are the time scale of growth (i.e. mechanical restructuring), the time scale of external loading, the time scale of viscous relaxation and the time scale of elastic wave propagation. By neglecting inertial effects, Tranquillo and Murray effectively assume that the time scale of wound healing is much longer than the time scale of elastic wave propagation. However, by including viscous effects, they assume that the time scale of viscosity is comparable to the time scale of wound healing. In actuality, we see that viscous effects are also negligible, but that internal remodelling (the growth described by Goriely and Ben Amar) is not.
and weeks, we are not interested in this initial transient behaviour. Furthermore, there are many ways in which our initial condition does not reflect the real initial state of a wound (e.g. a real wound would not contain a homogeneous provisional matrix). Thus, even if we were to find the initial transient solution, it is unlikely to correspond to the physical behaviour of a new wound.

In conclusion, we are justified in removing the viscous term from our equations altogether. Although Olsen et al. [136, 140] found that a large viscous term was necessary in order to make their model fit experimental results, this is simply a result of the fact that they did not take remodelling into account.35 Like Cook [33], we propose that the best approach to modelling the mechanics of dermal wound healing is to combine a purely elastic model of the dermis with a description of the tissue plasticity.

However, considering a purely elastic model introduces its own problems. As we noted in Section 3.2.3, taking $\mu$ to zero is a singular limit of the Tranquillo-Murray model. In order to deal with an elastic-only model appropriately, we need to think more carefully about our initial conditions and our numeric methods. We briefly discuss these issues in the following section.

---

35 As we will see in Chapter 6, the same theoretical framework that we use to describe tissue remodelling can be used to model some forms of viscoelasticity. Thus, it is not surprising that Olsen et al. found that $\mu$ was large when they fitted their results to data. Tissue remodelling leads to results that are mathematically similar to the results that we obtain from viscoelasticity.
3.4.2 Analysis of an elastic-only model

For now, let us consider the elastic-only model obtained from the Tranquillo-Murray model when \( \mu = 0 \). The equations for this model are given below:

\[
\frac{\partial n}{\partial t} + \frac{\partial}{\partial x} \left( n \frac{\partial u}{\partial t} \right) = \frac{\partial^2 n}{\partial x^2} + r_0 n (1 - n),
\]

(3.70)

\[
\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x} \left( \rho \frac{\partial u}{\partial t} \right) = 0,
\]

(3.71)

\[
\frac{\partial^2 u}{\partial x^2} + \frac{\partial}{\partial x} \left( \frac{\tau_0 \rho n}{1 + \gamma n^2} \right) = s \rho u.
\]

(3.72)

Equations (3.70) and (3.71) are identical to equations (3.1) and (3.2). However, (3.72) is different from (3.3) in that it does not depend on the time derivative of \( u \). In fact, equation (3.72) is an ordinary differential equation in \( x \) that must be satisfied at all times. This has some interesting effects. For example, we find that the initial conditions on \( n, \rho \) and \( u \) must satisfy

\[
\frac{d^2 u(x, 0)}{dx^2} + \frac{d}{dx} \left[ \frac{\tau_0 n(x, 0) \rho(x, 0)}{1 + \gamma (n(x, 0))^2} \right] = s \rho(x, 0) u(x, 0).
\]

(3.73)

In the original viscoelastic model, the following initial conditions were used:

\[
n(x, 0) = \begin{cases} 0 & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases}, \quad \rho(x, 0) = 1, \quad u(x, 0) = 0.
\]

Substituting these conditions into equation (3.73), we find that they are no longer appropriate for an elastic-only model. This is consistent with our earlier comments about an initial transient solution for \( u \) when \( \mu \) is small. We observe

As described above, a physically realistic model of dermal wound healing should also include some representation of tissue plasticity. Without this, an elastic-only model may even be a worse model of wound healing than an equivalent viscoelastic model. At this stage, however, we are concerned with elucidating the mathematical issues that arise in an elastic-only model rather than constructing a physically realistic model.
boundary layer behaviour when \( t \sim O(\mu) \) because the initial condition on \( u \) is inconsistent with the equation obtained in the limit as \( \mu \) tends to zero.

If we wish to investigate the case where \( \mu = 0 \), we need to use equation (3.73) to obtain a valid initial condition for \( u \). Thus, consider the result of substituting the \( n \) and \( \rho \) initial conditions given above into equation (3.73). This yields

\[
\frac{d^2 u(x, 0)}{dx^2} - s u(x, 0) = \frac{-\tau_0}{1 + \gamma} \delta(x - 1),
\]

where \( \delta(x) \) is the Dirac delta function.

Solving this subject to the boundary conditions given in (3.9) and demanding continuity of \( u \), we find that

\[
u(x, 0) = \begin{cases} 
\frac{\tau_0}{(1 + \gamma) \sqrt{s}} e^{-\sqrt{s}} \sinh \left( x \sqrt{s} \right), & 0 \leq x \leq 1; \\
\frac{\tau_0}{(1 + \gamma) \sqrt{s}} \sinh \left( \sqrt{s} \right) e^{-\sqrt{s}x}, & x \geq 1.
\end{cases}
\]  

(3.74)

Interestingly, we note that \( u(x, 0) \) is positive everywhere. This means that the elastic behaviour of the skin causes the wound boundary to retract immediately at \( t = 0 \). For the elastic-only Tranquillo-Murray model, it is even possible to show that \( u \) attains its global maximum when \( t = 0 \) and \( x = 1 \). That is,

\[
u_{\text{max}} = \frac{\tau_0}{(1 + \gamma) \sqrt{s}} e^{-\sqrt{s}} \sinh \left( \sqrt{s} \right) = \frac{\tau_0}{1 + \gamma} \frac{1 - e^{-s}}{2 \sqrt{s}}.
\]  

(3.75)

This result is particularly interesting because we expect the elastic-only model to predict larger displacements than the viscoelastic model. Thus, the expression for \( u_{\text{max}} \) above can be thought of as an upper bound on displacement for the full

\[37\text{If we included a model of tissue plasticity, we may be able to observe some further wound retraction. However, this is not possible in the elastic-only Tranquillo-Murray model. As the cells migrate into the wound space, they reduce the imbalance in stress and pull the displacements back towards zero.}\]
Another interesting observation from equation (3.74) is that \( u \) decays to zero without oscillating as \( x \) tends to infinity. In fact, it is possible to show that displacement oscillations will only occur in the elastic-only model if they are driven by significant oscillations in \( n \) and/or \( \rho \). Since \( n \) tends to a stable and spatially homogeneous steady state and \( \rho \) is effectively constant, we can safely conclude that removing viscosity eliminates the possibility of displacement oscillations.

Thus, a purely elastic model has several advantages over the full Tranquillo-Murray model. At a stroke we have removed the displacement oscillations and obtained a simple expression for the maximum displacement. However, the disadvantage of the purely elastic model is that it is more difficult to implement numerically.

In Section 3.2.3, we commented on the fact that our numeric method is unable to deal with the case where \( \mu = 0 \). To see why this failure occurs, we note that standard parabolic-elliptic solvers are designed to solve systems of equations of the form

\[
\frac{\partial}{\partial t} (\text{species}) = -\nabla \cdot (\text{flux}) + \text{source}.
\]

If we introduce an additional species, \( v = \frac{\partial u}{\partial t} \), we can easily rearrange the equations of the elastic-only model into this standard form. This yields the following system:

\[
\begin{align*}
\frac{\partial n}{\partial t} &= \frac{\partial}{\partial x} \left( \frac{\partial n}{\partial x} - v n \right) + r_0 n (1 - n), \\
\frac{\partial \rho}{\partial t} &= \frac{\partial}{\partial x} (-v \rho), \\
\frac{\partial u}{\partial t} &= v,
\end{align*}
\]

\(^{38}\)It should be noted that we have not rigorously demonstrated \( u_{\max} \) to be an upper bound. Further work is needed to determine whether this is the case.
We note that equations (3.76) to (3.78) involve time derivatives but (3.79) does not. Thus, a conventional numeric method (e.g. using the method of lines) would use equations (3.76) to (3.78) to step forward in time for $n$, $\rho$ and $u$. Based on these results, the method would then attempt to update $v$ using equation (3.79). However, equation (3.79) does not depend on $v$ in any way. Hence, a naïve numeric method will always fail.

One way of resolving this problem would be to use the following procedure at each time step:

- Use equations (3.76) and (3.77) to find $n$ and $\rho$ at the new time $(t + \Delta t)$.
- Solve for $u(x, t + \Delta t)$ by substituting the updated values of $n$ and $\rho$ into equation (3.79).
- Approximate $v(x, t + \Delta t)$ using a numeric derivative of $u$; for example, $v(x, t + \Delta t) = (u(x, t + \Delta t) - u(x, t))/\Delta t$.
- Use an appropriate method (e.g. an iterative scheme) to improve the accuracy of these solutions before moving to the next time step.

This is effectively the procedure used by Cook [33] to solve his full model of dermal wound healing.

Unfortunately, there are two difficulties associated with this procedure. Firstly, and most importantly, numeric differentiation is highly unstable and it is very unlikely that we would be able to obtain accurate results for velocity. Although this is not particularly important in the model described in this section, we will see that the accurate estimation of velocity is essential to solving the morphoelastic evolution equation for strain.
A second difficulty is that it is not easy to implement the procedure described above using standard numeric packages. Given the complexity of modelling wound healing, it is desirable to take advantage of the robust, efficient and accurate algorithms that are available for solving partial differential equations. However, if we follow the procedure above, it would probably be necessary for us to write our own specialist code.

Fortunately, this procedure is not the only possible way of obtaining a solution to the elastic-only problem. If we differentiate equation (3.79) with respect to time,

$$0 = \frac{\partial}{\partial x} \left( \frac{\partial v}{\partial x} + \left( \frac{\tau_0 \rho n}{1 + \gamma n^2} \right) \frac{\partial n}{\partial t} + \left( \frac{\tau_0 \rho n}{1 + \gamma n^2} \right) \frac{\partial \rho}{\partial t} \right) - s \frac{\partial \rho}{\partial t} u - s \rho v. \quad (3.80)$$

We can now substitute expressions for $\frac{\partial n}{\partial t}$ and $\frac{\partial \rho}{\partial t}$ from equations (3.76) and (3.77) into equation (3.80). This leads to an ordinary differential equation in space that can be solved to obtain $v$ at each time step. The full equation obtained in this way is too complicated to be readily informative. For simplicity, we will therefore consider the situation where $\gamma = r_0 = 0$. This yields

$$0 = \frac{\partial}{\partial x} \left( \frac{\partial v}{\partial x} + \tau_0 \rho \left( \frac{\partial^2 n}{\partial x^2} - \frac{\partial}{\partial x} (n v) \right) - \tau_0 n \frac{\partial}{\partial x} \left( \rho v \right) \right) - s \frac{\partial}{\partial x} \left( \rho v \right) - s \rho v. \quad (3.81)$$

Using this approach, we were able to obtain numeric results for the elastic-only Tranquillo-Murray model. A plot of the evolving displacement profile at early times is shown in Figure 3.13.

Although equation (3.81) allows us to use a standard numeric method, there are a number of potential problems with this approach. Most importantly, we cannot guarantee that replacing the force balance equation with its time derivative will lead to consistent results. When we replaced equation (3.79) with equation

---

39In order to remain consistent with the earlier system of equations, we also need to impose the initial condition on $u$ given in equation (3.74). Otherwise, differentiating with respect to time would cause us to lose some information.
Figure 3.13: Numeric simulations of $u(x, t)$ from the elastic-only model. Parameter values are $\tau = 0.2$, $s = 1$, $\gamma = r_0 = 0$. Times shown are $t = 0$ to $t = 0.1$ in steps of 0.01. The initial condition is shown as a dashed line and the profiles at $t = 0.01$ and $t = 0.02$ are shown as very thick and thick lines respectively. Note that there is no evidence of oscillations at large values of $x$.

(3.81), our numeric results were not completely stable. As well as using (3.81) to define $v$, we found that it was necessary to use (3.79) to define $u$ (instead of (3.78)). The best results, such as those illustrated in Figure 3.13, were obtained by algebraically rearranging the equations obtained in order to make them as similar to reaction-diffusion equations as possible.

In conclusion, we note that removing viscosity has significant consequences for the Tranquillo-Murray model. From an analytical perspective, we find that we lose the ability to specify an initial condition for $u$. However, we also find that we completely eliminate the displacement oscillations observed in viscoelastic models.

Unfortunately, the elastic-only model is more difficult to implement numerically than the viscoelastic model. We have outlined two possible approaches for dealing with the problems that arise in elastic-only models. At this stage, however, it is not possible to determine which approach is best. There may even be a different
approach that avoids the disadvantages of the two approaches that we describe here. Further work is necessary in order to decide how best to proceed with the numeric simulation of elastic-only models.
3.5 Towards a new model of dermal wound healing

In this chapter, we have thoroughly analysed the dermal wound healing model proposed by Tranquillo and Murray. This work has been motivated by our desire to construct a new mechanical model of dermal wound healing that builds on the insights of earlier researchers. To conclude this chapter, we summarise some of our most important findings, focusing on results that will help to shape the wound healing model that we propose in Chapter 6.

A morphoelastic model is more realistic than a viscoelastic model

This chapter has been dominated by our analysis of the constitutive law that Tranquillo and Murray proposed for the ECM. We found that their assumption of linear viscoelasticity leads to unphysical results. In a viscoelastic model of dermal wound healing, the viscous stress, the cellular traction stress and the subdermal tethering force will always interact to cause apparent oscillations in displacement. However, these oscillations would not appear to be realistic on the time scale of wound healing. Thus, we seek to construct an improved model that does not have this problem.

At first glance, it is surprising that a viscoelastic constitutive law should lead to unrealistic oscillations; it is well known that human skin can behave as a viscoelastic material [171]. However, experimental results indicate that the time scale of viscous relaxation in the skin is extremely short compared to the time scale of wound healing. This leads us to conclude that the dimensionless parameter of viscosity in a wound healing model is so small as to be practically zero. Thus, it is more realistic to use a purely elastic model than a viscoelastic model.

Interestingly, earlier modellers were led to the conclusion that the viscosity of the skin is large, not small [140]. Not only is this inconsistent with experimental
Analysis of the Tranquillo-Murray wound healing model

studies, but it also tends to magnify the observed displacement oscillations. We suspect that the viscosity of the skin was overestimated by previous researchers because the active remodelling of the dermal ECM by fibroblasts yields results that are similar in appearance to viscous flow. Thus, a purely elastic model will probably not be sufficient for modelling wound healing. We also need to include some representation of the evolving plastic behaviour of the material. This requires a theory of morphoelasticity, which we will consider in detail over the course of the next two chapters.

Advection is unnecessary

A second important point is that removing advection from the Tranquillo-Murray model did not have any significant effect on the output. Because the displacements and velocities predicted by the model were universally small, we found that the advective fluxes were practically zero. Thus, it is unnecessary for us to include passive advection in the model of dermal wound healing that we develop.

Interestingly, we also note that the linear constitutive laws used by Tranquillo and Murray are only valid if it is assumed that the displacement gradient is always small. If the velocities were large, this assumption would probably be violated. Thus, in any case where advection is important, we would need to alter the ECM constitutive law to take into account the fact that the displacement gradients may become large.

Cells respond to their mechanical environment

If advection is removed from the Tranquillo-Murray model, we find that the force balance equation decouples from the rest of the system. Thus, the mechanical model becomes an ‘optional add-on’ to a conventional system of reaction diffusion equations. However, experimental results indicate that the mechanical environ-
ment of a wound affects the quality of healing (see, for example, Aarabi et al. [1]). This is because cells change their behaviour (and even their phenotype) according to the stress and/or strain of the ECM.

This is an important part of wound healing that is yet to be investigated using mathematical models. The Tranquillo-Murray model [189] and the Olsen models [136, 137, 138, 139, 140, 141] can all be decoupled by removing advection. Moreover, advection is observed to be very small within the range of parameter values explored. Hence, the mechanical models proposed by these earlier researchers are effectively irrelevant to the rest of the system.

Interestingly, even the Cook model [33] does not take into account the response of cells to stress or strain. Although he developed a theory to describe cellular motion in an oriented environment, Cook did not consider the possibility that the stress exerted by cells may be dependent on ECM mechanics. We will consider this further in Chapter 6.

**Conclusion**

The Tranquillo-Murray model laid the foundation for mechanochemical modelling of dermal wound healing. Although their model suffers from a number of flaws, Tranquillo and Murray’s work was truly groundbreaking. When we come to describe our own model of wound healing in Chapter 6, we will essentially build on the Tranquillo-Murray model, although we will also incorporate insights from Olsen *et al.* and Cook. Fundamentally, all of our work is underpinned by the careful and critical analysis of the Tranquillo-Murray model that we have described here.
Chapter 4

Zero stress state theory and the definition of strain

4.1 Introduction

As demonstrated in the previous chapter, it is not appropriate to model the mechanics of the dermis using a classical viscoelastic constitutive law. Instead, it is better to use an elastic model that has been modified to take into account the evolving mechanical structure of the dermal ECM (i.e. the plasticity of granulation tissue). In his doctoral work, Cook [33] developed an innovative mathematical approach to deal with the problem of wound plasticity. This involved modelling residual strains using a tensorial representation of the zero stress state of the ECM.

The techniques developed by Cook for modelling wound plasticity are very similar to those developed by Rodriguez et al. [159] for modelling the growth of soft tissues.\(^1\) Both approaches depend on the idea that the deformation gradient

\(^1\)It is interesting to note that these two developments occurred independently and at roughly the same time. Rodriguez et al. [159] was published before Cook submitted his thesis, but it
tensor used in classical mechanics can be expressed as the product of two tensors. One of these represents the elastic deformation while the other represents the plastic deformation associated with wound healing or growth. As we will see, this multiplicative decomposition is a useful and informative way of representing the changing zero stress state of biological materials. Furthermore, the multiplicative decomposition of the deformation gradient gives us a theoretical framework for modelling residual strains.

Since the time of Cook and Rodriguez et al., considerable theoretical progress has been made in this area (see, for example, Lubarda and Hoger [113]). Also, a number of new biomechanical applications of zero stress state theory have been developed (see, for example, the recent review by Goriely et al. [71]).\(^2\) Despite this progress, however, there is a need for simple methods that can be used to incorporate an evolving zero stress state into a mathematical model.

As it stands, the complexity of zero stress state theory is the largest obstacle to its general use. Although the plastic behaviour of granulation tissue is very important in dermal wound healing, no mathematical models since Cook [33] have included an evolving zero stress state. Furthermore, many of the theoretical papers describing the zero stress state are very dense and very general; there is still a lot of ambiguity surrounding the ‘correct’ way to model the evolution of the zero stress state in response to biological processes such as growth and tissue remodelling.

We seek to remedy this perceived deficiency by developing a novel theory of morphoelasticity to describe the evolving zero stress state of a tissue that undergoes changes to its internal structure. As we will see, the theory that we develop is considerably simpler and more applicable than the other theories in the existing literature. Before we consider the problem of an evolving zero stress state, it behooves us to clearly define the zero stress state and describe how it relates to

---

\(^2\)We described many of these developments (both theoretical and applied) in Section 2.4.
conventional mechanical concepts like strain. In this chapter, we focus entirely on ‘static’ zero stress state theory; the evolution of the zero stress state is dealt with in Chapter 5.

Firstly, in Section 4.2, we review the classical theory of deformation and strain; this lays the foundation for the extension of classical mechanics to take into account the zero stress state. In this section, we define important classical measures of deformation (e.g. the deformation gradient tensor and the strain tensor) and discuss some well-known results that are important to our development of zero stress state theory, such as the polar decomposition theorem.

Next, in Section 4.3, we explore the concept of a zero stress state. Much of the material in this section is analogous to earlier treatments of the zero stress state, particularly those described by Hoger and coworkers (see, for example, Rodriguez et al. [159], Lubarda and Hoger [113] or Hoger et al. [94]). By introducing the concept of a stress-relieving deformation, our development follows classical deformation theory as closely as possible and shows clearly why the multiplicative decomposition of the deformation gradient tensor is appropriate. Also, we are able to construct necessary conditions for the existence and uniqueness of the zero stress state (see Appendix B for details of this proof).

One interesting feature of our analysis is that the definition of strain that we obtain is different from that originally used by Rodriguez et al. [159]. In fact, we are able to show that the strain tensor proposed by Rodriguez et al. may not be well-defined in some situations. Furthermore, we show that it is very difficult to construct a stress tensor that can be associated with Rodriguez et al.’s strain.

Finally, in Section 4.4, we use our definition of finite strain to construct a theory of infinitesimal strain that is appropriate in cases where the deformed body is close to its zero stress state. That is, we consider the case where the body may have deformed significantly from its initial state, but the strain (measured as the difference between the current state and the zero stress state) is small.
4.2 The classical theory of elastic deformation

Following Spencer’s *Continuum Mechanics* [176] and elsewhere, standard notational practices such as the use of bold type to represent vectors and tensors and Einstein summation notation are used throughout this chapter and subsequent chapters. However, formal covariant/contravariant notation is not used (*e.g.* metric tensors are omitted from the definition of the Cauchy-Green tensors).

In addition to Spencer’s notation, we use the expression

\[ A = \frac{\partial b}{\partial c} \]

to indicate that the tensor \( A \) has components given by

\[ A_{ij} = \frac{\partial b_i}{\partial c_j}, \]

where \( b \) and \( c \) are vectors. Any other nonstandard notation will be defined as it appears.

4.2.1 Motions, deformations and coordinate systems

Consider a continuous body \( \mathcal{B} \) that initially occupies the region of space given by the points in \( \mathcal{R}_0 \subseteq \mathbb{R}^3 \). At a later time \( t \), the body \( \mathcal{B} \) has moved and deformed so that it now occupies the space \( \mathcal{R}_t \subseteq \mathbb{R}^3 \). We can describe the motion and deformation of \( \mathcal{B} \) from time 0 to a general time \( t \in \mathbb{R}^+ \) using a mapping of the form

\[ f : (\mathcal{R}_0, \mathbb{R}^+) \to \mathcal{R}_t. \]

That is, given a particle in \( \mathcal{B} \) that is initially at the point with position vector \( X \), we can find the position vector \( x \) that corresponds to the same particle at time \( t \) by applying \( f \):

\[ x = f(X, t). \quad (4.1) \]
At any given time we expect the deformation of $\mathcal{B}$ to be continuous; that is, connected regions of the initial configuration map to connected regions of the current configuration and no two particles in the initial configuration have the same position in the current configuration. If this is the case, $f$ will be a continuous and bijective function of $x$ for any given value of $t$. Thus, it is possible to construct an inverse function, $f^{-1}$, with the property that

$$X = f^{-1}(x, t), \quad (4.2)$$

where $X$ and $x$ are defined as before. Moreover, if $f$ represents a real deformation in response to applied forces, we expect $f$ and $f^{-1}$ to be smooth.

This one-to-one correspondence between $X$ and $x$ means that we have two different ways of describing any motion or deformation of $\mathcal{B}$: we can use the original position $X$ as an independent variable and define the changes to $\mathcal{B}$ using $f$, or we can use the current position $x$ as an independent variable and define the changes to $\mathcal{B}$ using $f^{-1}$. The former option is referred to as the Lagrangian or material coordinate system, while the latter is referred to as the Eulerian or spatial coordinate system.

Theoretically, there is a wide range of other coordinate systems that could also be used. As described by Malvern [118], it is sometimes useful to construct an alternative coordinate system based on some other configuration of the body $\mathcal{B}$ (i.e. not the initial configuration or the current configuration). All that is required is that each particle in $\mathcal{B}$ is labelled uniquely in the given coordinate system and contiguous regions of $\mathcal{B}$ are represented by contiguous sets of points. Hence, $\chi$ defines an appropriate coordinate system if and only if we can construct bijective, continuous (and preferably smooth) functions $a$ and $b$ such that

$$X = a(\chi, t), \quad (4.3)$$
$$x = b(\chi, t). \quad (4.4)$$
Given the earlier definitions of $f$ and $f^{-1}$, we find that $f = b \circ a^{-1}$ and $f^{-1} = a \circ b^{-1}$. Hence, if we are given $a$ and $b$, it is always possible to construct the functions $f$ and $f^{-1}$ and thus convert back to a Lagrangian or Eulerian system.

One physically meaningful way of defining the $\chi$ coordinate system would be to use a zero stress state of the body. That is, we could define $\chi$ so that the body would be free of elastic stress if $x = b(\chi, t) = \chi$ throughout. In such a coordinate system, the function $a$ can be thought of as representing plastic changes to $\mathfrak{B}$ while $b$ represents the elastic deformation. Unfortunately, it is not generally possible to construct an internally consistent zero stress coordinate system. Although theoretically interesting, the concept of a zero stress coordinate system is of limited practical use.

The function $f$ in Lagrangian coordinates, the function $f^{-1}$ in Eulerian coordinates and the pair of functions $a$ and $b$ in general coordinates can all be thought of as complete mathematical descriptions of the physical changes that occur to $\mathfrak{B}$ over time. That is, each of these formulations can be used to describe the rigid motions (i.e. translations and rotations) of $\mathfrak{B}$ as well as the deformation that it experiences.

In solid mechanics we are principally concerned with material properties that are independent of rigid motions. Thus, we do not need all of the information that can be obtained from the functions $f$, $f^{-1}$, $a$ and $b$. Instead, it is more practical for us to use a simpler measure of deformation that is unaffected by translations and rotations. In classical mechanics, the most commonly used measure of deformation is the strain tensor, which we describe in Section 4.2.5. Before we define strain, however, we need the deformation gradient tensor.

---

3Specifically, problems arise if residual stresses are present, as described by Hoger [91, 92] and by Olsson [143]. Alternative approaches for representing the zero stress state exist and we describe them in Section 4.3.2.
4.2.2 The deformation gradient tensor

Let $\mathcal{P}$ be a particle in $\mathcal{B}$ and let $\mathcal{Q}$ be another particle in the neighbourhood of $\mathcal{P}$. A rigid motion of $\mathcal{B}$ will leave the distance between these two particles unchanged but a deformation might cause $\mathcal{P}$ and $\mathcal{Q}$ to become closer to each other or further apart. One way of describing the distance between $\mathcal{P}$ and $\mathcal{Q}$ is the deformation gradient tensor, which we describe in this section.

As we will see, the deformation gradient tensor allows us to calculate the vector between $\mathcal{P}$ and $\mathcal{Q}$ in the current configuration using the vector between $\mathcal{P}$ and $\mathcal{Q}$ in the initial configuration. However, we ultimately wish to develop a measure of strain that indicates how the scalar distance between $\mathcal{P}$ and $\mathcal{Q}$ changes over time. As such, we find that the deformation gradient tensor contains too much irrelevant information\footnote{Specifically, the deformation gradient tensor contains irrelevant information about the relative orientation of $\mathcal{P}$ and $\mathcal{Q}$.} to be a useful measure of deformation.

![Diagram](image.png)

Figure 4.1: Over time, the particles $\mathcal{P}$ and $\mathcal{Q}$ will move through space and the vector that connects them will change. The initial vector connecting $\mathcal{P}$ and $\mathcal{Q}$ is labelled $\delta \mathbf{X}$ and the vector connecting them at an arbitrary later time is labelled $\delta \mathbf{x}$. In a rigid motion, the distance between $\mathcal{P}$ and $\mathcal{Q}$ remains constant, but the direction of the vector can change.

Now, as shown in Figure 4.1, let $\mathbf{X}$ and $\mathbf{X} + \delta \mathbf{X}$ be the position vectors of the particles $\mathcal{P}$ and $\mathcal{Q}$ at time 0. Similarly let $\mathbf{x}$ and $\mathbf{x} + \delta \mathbf{x}$ be the position vectors...
of $P$ and $Q$ at time $t$. With $f$ and $f^{-1}$ defined as in equations (4.1) and (4.2), we write

$$
\delta x = f(X + \delta X, t) - f(X, t), \quad \delta X = f^{-1}(x + \delta x, t) - f^{-1}(x, t). \quad (4.5)
$$

Now, let $\delta S = ||\delta X||$ be the initial distance between $P$ and $Q$ and let $\delta s = ||\delta x||$ be the distance between the same particles at time $t$. Furthermore, let the vectors $\hat{m}$ and $\hat{n}$ be unit vectors in the direction of $\delta X$ and $\delta x$ respectively. Hence, in Lagrangian coordinates we find that

$$
\hat{n} \delta s = f(X + \hat{m} \delta S, t) - f(X, t).
$$

Assuming that $f$ is differentiable (as we expect), it follows from Taylor’s theorem that

$$
\hat{n} \delta s = f(X, t) + \frac{\partial f}{\partial X} \hat{m} \delta S + O(\delta S^2) - f(X, t)
= \frac{\partial f}{\partial X} \hat{m} \delta S + O(\delta S^2).
$$

In the limit as $\delta S$ approaches 0, we replace $\delta S$ by $dS$ and $\delta s$ by $ds$ to find that

$$
\frac{ds}{dS} \hat{n} = F \hat{m}, \quad (4.6)
$$

where $F = \frac{\partial f}{\partial X} = \frac{\partial x}{\partial X}$ is called the deformation gradient tensor. Using Einstein summation notation, equation (4.6) can be written as

$$
\frac{ds}{dS} \hat{n}_i = F_{iJ} \hat{m}_J = \frac{\partial x_i}{\partial X_J} \hat{m}_J. \quad (4.7)
$$

Rearranging equation (4.6), we find that

$$
dx = F \, dX. \quad (4.8)$$
Thus, the deformation gradient tensor $F$ is effectively a way of describing how infinitesimal line segments are changed by a deformation. The deformation gradient will generally vary throughout $\mathcal{B}$, but for continuous deformations it will always be nonsingular. This means that we can also construct $F^{-1} = \frac{\partial X}{\partial x}$ such that

$$dX = F^{-1} dx$$  \hspace{1cm} (4.9)

at all points in $\mathcal{B}$.\footnote{Equation (4.9) can also be obtained by starting from the Eulerian definition of $\delta X$ in (4.5b) and using a Taylor expansion in $\delta s$.}

The deformation gradient and its inverse have a number of applications; for example, they can be used to convert line integrals from Lagrangian coordinates to Eulerian coordinates and vice versa. However, $F$ is not an appropriate measure of the true deformation because it is affected by rigid rotations, as we shall see in the following section.

### 4.2.3 Rotations and orthogonal tensors

Throughout this chapter and the following one, we will make use of the well-known fact that rotations can be represented using proper orthogonal tensors. In this section, we define orthogonal tensors and describe some important and useful results.

A second-order tensor $Q$ is called orthogonal if its transpose is equal to its inverse (i.e. $Q^T = Q^{-1}$). The determinant of a real-valued orthogonal tensor must either be 1 or $-1$. If $\det Q = 1$, the tensor is called proper orthogonal; if $\det Q = -1$, the tensor is called improper orthogonal.

The set of orthogonal tensors form a group under multiplication and the set of proper orthogonal tensors is a subgroup of the set of orthogonal tensors. Any rotation can be represented using a proper orthogonal tensor and any proper
orthogonal tensor represents a rotation. That is, rotating a vector by a given angle around a given axis is equivalent to multiplying that vector by a proper orthogonal tensor. It is relatively straightforward to convert between the tensor and axis-angle representations of a rotation.

In the case where a rigid motion is applied to a body, it can be shown that the deformation gradient tensor, $F$, is proper orthogonal. To see this, we note that any rigid motion of a body can be represented as a rotation followed by a translation of the origin. That is, if $x = f(X)$ defines a rigid motion at a particular time, then

$$f(X) = QX + c,$$

where $Q$ is proper orthogonal (describing a rotation), and $c$ is a vector that indicates the translation of the origin. By differentiating this expression with respect to $X$, it can be seen that the deformation gradient tensor is given by

$$F = \frac{\partial f}{\partial X} = Q.$$

Now, let us consider what happens to $F$ when a rigid motion is combined with another deformation. Firstly, let us define a general deformation by $x = f(X)$ throughout the body $\mathcal{B}$. This will have a deformation gradient tensor, $F$, with components

$$F_{iJ} = \frac{\partial f_i}{\partial X_J}.$$

Now consider the deformation that is obtained when $f$ is followed by a rotation of $\mathcal{B}$. That is, we define $f^*$ so that $f^*(X) = Qf(X)$ and $Q$ is a proper orthogonal tensor that is independent of $X$. The new deformation gradient, $F^*$, will have

---

6Similarly, a rigid motion that is not a rotation and does not involve translation of the origin (i.e. a reflection or a rotoinversion) can be represented using an improper orthogonal tensor and vice versa. However, such operations are not generally important in continuum mechanics.
components
\[
F_{iJ}^* = \frac{\partial f_i^*}{\partial X_j}
\]
= \frac{\partial (Q_{ik} f_k)}{\partial X_j}
= Q_{ik} \frac{\partial f_k}{\partial X_j}.
\]

Hence,
\[
F^* = Q F. \tag{4.10}
\]

That is, the overall deformation gradient for a deformation followed by a rotation can be obtained by premultiplying the deformation gradient for the original deformation by the proper orthogonal tensor associated with the rotation.

Similarly, consider the deformation that is obtained when a body is rotated before the deformation \( f \) is applied. That is, let \( f^{**}(X) = f(Q X) \). Now,
\[
F_{iJ}^{**} = \frac{\partial f_i^{**}}{\partial X_J}
\]
= \frac{\partial f_i(QX)}{\partial X_J}.
\]

Now, let us define the vector \( \bar{X} \) so that \( \bar{X} = QX \) (i.e. \( \bar{X}_A = Q_{AB} X_B \)). Using the chain rule, we find that
\[
F_{iJ}^{**} = \frac{\partial f_i(\bar{X})}{\partial \bar{X}_K} \frac{\partial \bar{X}_K}{\partial X_J}
\]
= \( F_{iK} \frac{\partial (Q_{KL} X_L)}{\partial X_J} \)
= \( F_{iK} Q_{KL} \delta_{LJ} \),
where $\delta$ is the Kronecker delta. Thus,

$$F_{iJ}^{**} = F_{iK} Q_{KJ},$$

or alternatively,

$$F^{**} = FQ. \tag{4.11}$$

One well-known theorem concerning orthogonal tensors is the polar decomposition theorem (see, for example, Spencer [176]). This theorem states that every square matrix $A$ (and similarly every second-order tensor) can be expressed in the form

$$A = QU = V Q,$$

where $Q$ is orthogonal and $U$ and $V$ are both symmetric positive semi-definite. In the case where $A$ is nonsingular, the polar decomposition is unique. Moreover, if $A$ has a positive determinant, then $Q$ is proper orthogonal and $U$ and $V$ are both symmetric positive definite. For further details of this theorem see, for example, Anton and Busby [6] or Spencer [176].

Importantly, the determinant of the deformation gradient tensor is always positive.\footnote{If the deformation of $\mathfrak{B}$ is continuous, it follows that $\det F$ will be a continuous function of $t$ and $x$. Furthermore, it can be shown that $\det F = 0$ corresponds to an impossible situation where there are superimposed particles. Since $\det F$ is initially 1, it follows that $\det F$ must always be positive.} Thus, the polar decomposition theorem yields

$$F = RU = V R, \tag{4.12}$$

where $R$ is proper orthogonal and $U$ and $V$ are symmetric positive definite. The tensor $R$ is called the rotation tensor of the deformation, while $U$ and $V$ are referred to as the right stretch tensor and left stretch tensor respectively. Since the stretch tensors are unaffected by rigid rotations,\footnote{Strictly speaking, this is not true. The right stretch tensor is affected by a rotation of the initial orientation of the body (i.e. a rotation in Lagrangian space), while the left stretch tensor is affected by a rotation of the final orientation of the body (i.e. a rotation in Eulerian space). Thus, we tend to use the right stretch tensor in conjunction with Lagrangian coordinates and} it is possible to use them...
as measures of the pure deformation. However, they are not used as widely as the closely-related Cauchy-Green tensors described in the next section. For more information about the use of stretch tensors in solid mechanics, see Spencer [176].

### 4.2.4 The Cauchy-Green tensors

Consider what happens if we take equation (4.6) and take the scalar product of each side with itself. That is,

$$(\frac{ds}{dS} \hat{n}) \cdot (\frac{ds}{dS} \hat{n}) = (F \hat{m}) \cdot (F \hat{m}).$$

This yields the result that

$$(\frac{ds}{dS})^2 \hat{n}^T \hat{n} = \hat{m}^T F^T F \hat{m}.$$

Since $\hat{n}$ is a unit vector, we find that $\hat{n}^T \hat{n} = 1$. Thus,

$$(\frac{ds}{dS})^2 = \hat{m}^T F^T F \hat{m}.$$

Using the fact that $\hat{m} dS = dX$, we can multiply through by $dS^2$ to find that

$$ds^2 = dX^T F^T F dX = dX^T C dX, \quad (4.13)$$

where $C = F^T F$ is called the right Cauchy-Green tensor. Since $C$ is the product of a tensor with its own transpose, we note that $C$ will be symmetric.

Now, consider the case where we know the infinitesimal vector, $dX$, that connects two neighbouring particles in the initial configuration. Equation (4.13) means that $C$ can be used to calculate the distance, $ds$, between the same particles at time $t$. Thus, the right Cauchy-Green tensor can be interpreted as a tensorial the left stretch tensor in conjunction with Eulerian coordinates. For a fixed coordinate system, this ensures that the stretch tensor being used is invariant to rigid motions.
description of the relationship between distance in the current configuration and displacement in the initial configuration.

Similarly, we can use the fact that

\[
\frac{dS}{ds} \dot{\mathbf{m}} = F^{-1} \dot{\mathbf{n}},
\]

to derive the Eulerian equivalent of equation (4.13). This takes the form

\[
ds^2 = d\mathbf{x}^T F^{-T} F^{-1} d\mathbf{x} = d\mathbf{x}^T B^{-1} d\mathbf{x}, \quad (4.14)
\]

where \( F^{-T} \) denotes the inverse of the transpose of \( F \).\(^9\) The tensor \( B = FF^T \) is called the left Cauchy-Green tensor and, like the right Cauchy-Green tensor, it is symmetric. Given an infinitesimal vector \( d\mathbf{x} \) between neighbouring particles in the current configuration, \( B^{-1} \) can be used to calculate \( dS \), the original distance between those particles.

In a rigid motion, \( F \) is proper orthogonal (i.e. \( F^T = F^{-1} \)). In this case, it follows that \( B \) and \( C \) will both be equal to the identity tensor. This is consistent with the physical interpretation of the Cauchy-Green tensors, because the distances between any two particles are unchanged by a rigid rotation.

If a deformation is followed by a rotation, we recall that

\[
F^* = Q F,
\]

where \( F \) represents the original deformation and \( Q \) is a proper orthogonal tensor

\(^9\)We will use this notation for the inverse-transpose throughout the remainder of the thesis.
that represents the rotation. In this case,

\[ C^* = F^{*T} F^* \]
\[ = F^T Q^T Q F \]
\[ = F^T F \]
\[ = C. \]

Thus, the right Cauchy-Green tensor is unaffected by a rotation that is applied after a deformation. This makes sense because the right Cauchy-Green tensor relates vectors in the initial configuration to distances in the current configuration. Rotating \( B \) after some initial deformation will leave the distances between particles in the current configuration unchanged.

In contrast, \( C \) is affected if the body is rotated before it is deformed. If we define \( F^{**} = F Q \), we ultimately find that

\[ C^{**} = Q^T C Q. \quad (4.15) \]

This can be explained by noting that rotating a body before deforming it is equivalent to rotating the Lagrangian coordinate system. Thus, we would expect equation (4.15) to be consistent with the results that we would obtain from rotating the coordinate system of a one-point tensor.\(^{10}\) This is indeed what we observe.

In contrast, it can be shown that the left Cauchy-Green tensor, \( B \), is unaffected by performing a rotation before the deformation. However, it is affected by

---

\(^{10}\)A one-point tensor is a tensor where all of the subscripts refer to the same coordinate system and need to be rotated together. For example, the value of \( C_{12} \) can be thought of as measuring the interaction of \( dX_1 \) and \( dX_2 \) (both in the Lagrangian coordinate system) following the deformation. In contrast, a two-point tensor has subscripts that refer to different coordinate systems. A classic example of a two-point tensor is the deformation gradient. By its definition, \( F_{iJ} \) describes how \( dx_i \) relates to \( dX_J \). When rotating the coordinate system of a one-point tensor, it is necessary to rotate both subscripted dimensions together. In contrast, it is possible to rotate the two dimensions of a two-point tensor independently. Three-point, four-point etc. tensors of higher-order are defined analogously.
performing a rotation afterwards. This is because the left Cauchy-Green tensor is a one-point tensor defined with respect to the Eulerian coordinate system.

It is interesting to note that the left and right Cauchy-Green tensors are closely related to the left and right stretch tensors that we defined at the end of Section 4.2.3. We recall from equation (4.12) that \( F = RU \), where \( R \) is the rotation tensor of \( F \) and \( U \) is the right stretch tensor. Using this equation, we find that

\[
C = F^T F = U^T R^T RU = U^2,
\]

since \( R \) is orthogonal and \( U \) is symmetric. Similarly, it can be shown that \( B = V^2 \).

Since the stretch tensors are symmetric positive definite, we find that the Cauchy-Green tensors are also symmetric positive definite. Furthermore, any symmetric positive definite matrix has a unique symmetric positive definite square root (see, for example, [165]). Thus, it is theoretically possible to obtain the stretch tensors from the Cauchy-Green tensors.\(^{11}\)

### 4.2.5 The classical definition of strain

In the case of a rigid motion, the Cauchy-Green and stretch tensors all reduce to the unit tensor, \( I \). However, it is more useful for us to construct a measure of deformation with the property that a rigid motion leads to a deformation of zero. This leads us to the strain tensors.

As described by Malvern [118], the Lagrangian strain tensor \( E^L \) is defined so that

\[
(dS)^2 - (ds)^2 = 2 dX^T E^L dX,
\]  

\(^{11}\)Since there is a bijective relationship between the Cauchy-Green tensors and the stretch tensors, they can be considered to be equivalent measures of deformation. However, in certain situations it is preferable to use one over the other.
and the Eulerian strain tensor $E^E$ is defined so that

$$(ds)^2 - (dS)^2 = 2 \frac{d}{dx^T} E^E d\bf{x}.$$  \hfill (4.17)

As we will see, these equations can be combined with the definition of the Cauchy-Green tensors to construct formal definitions for the strain tensors.

However, let us first reflect on how equations (4.16) and (4.17) can be interpreted physically. Given two neighbouring particles $P$ and $Q$ such that $d\bf{X}$ is the vector separating them in Lagrangian coordinates, we find that the Lagrangian strain tensor allows us to calculate how much that line segment is stretched in the current configuration relative to the initial configuration (as measured by $(ds)^2 - (dS)^2$). Thus, the strain tensor allows the degree of stretching to be evaluated for any given direction of the vector $d\bf{X}$. Importantly, this means that the strain tensor will be equal to the zero tensor at a point if and only if the body is completely unstretched in the neighbourhood of that point (i.e. $E^L = 0$ implies that $ds = dS$ throughout the neighbourhood of $P$ and vice versa).

At first glance, the factor of two appearing on the right hand sides of equations (4.16) and (4.17) seems to be unnecessary. However, this can be justified by considering the case where $ds$ and $dS$ are almost equal (i.e. $\frac{ds}{dS} = 1 + \epsilon$). With this assumption, we find that equation (4.16) becomes

$$2 d\bf{X}^T E^L d\bf{X} = (1 + \epsilon)^2 \cdot (dS)^2 - (dS)^2$$

$$= dS^2 + 2 \epsilon dS^2 + \epsilon^2 dS^2 - dS^2$$

$$= 2 \epsilon dS^2 + \epsilon^2 dS^2.$$

Noting that $\epsilon = \frac{dx-dS}{dS}$ and neglecting terms of order $\epsilon^2$, we have

$$(ds - dS) dS \approx d\bf{X}^T E^L d\bf{X}.$$  

Letting $\hat{m}$ represent a unit vector in the direction of $d\bf{X}$ as before, we now divide
through by \(dS^2\) to find that

\[
\frac{d(s - S)}{dS} \approx \hat{m}^T E^L \hat{m} \tag{4.18}
\]

Similarly, if we begin with equation (4.17), we find that

\[
\frac{d(s - S)}{ds} \approx \hat{n}^T E^E \hat{n} \tag{4.19}
\]

where \(\hat{n}\) is a unit vector in the direction of \(d\mathbf{x}\).

Thus, the Lagrangian (Eulerian) strain allows the change in length of an infinitesimal line segment per unit of original (current) length to be approximated for a given point and direction in the initial (current) configuration. These interpretations of strain are physically useful and some texts even use equations (4.18) and (4.19) as their main definitions for strain.

It should be noted that this analysis depends on \(ds - dS\) being small relative to \(dS\) (\(ds\) for Eulerian strain). However, this may not be the case if the deformations and stretches are large. In order to take large changes like this into account, an alternative definition of strain called ‘true strain’ or ‘logarithmic strain’ is sometimes used. If logarithmic strain is used, we find that (4.18) and (4.19) are equalities for all relative sizes of \(ds\) and \(dS\). However, logarithmic strain is significantly more difficult to apply computationally than conventional strain.

Throughout this work, we will use conventional strain rather than logarithmic strain. One reason for this is that the zero stress state theory associated with conventional strain is better developed than the equivalent theory for logarithmic strain. However, recent theoretical developments indicate that it may be possible to use logarithmic strain to describe the evolution of the zero stress state (see, for example, Xiao et al. [196]). Furthermore, advances in algorithms and computing power are making logarithmic strain more widely used. Despite this, we remain focused on conventional strain.
Now, the conventional Lagrangian and Eulerian strain tensors may be defined in terms of the Cauchy-Green tensors. Combining equations (4.13) and (4.16), we find that

\[ dX^T C dX - dS^2 = 2dX^T E^L dX, \]

or

\[ dX^T (C - I) dX = 2dX^T E^L dX. \]

As this equation must hold for all possible directions \( dX \), we conclude that

\[ E^L = \frac{1}{2}(C - I), \quad (4.20) \]

where \( C \) is the right Cauchy-Green tensor. Similarly, combining (4.14) and (4.17) yields

\[ E^E = \frac{1}{2}(I - B^{-1}), \quad (4.21) \]

where \( B \) is the left Cauchy-Green tensor.

The tensors \( E^E \) and \( E^L \) are relatively unwieldy and difficult to calculate. However, it is possible to simplify equations (4.20) and (4.21) in the case where all displacements are small (i.e. the tensors \( F \) and \( F^{-1} \) are both close to the identity).

In order to simplify strain in this case we can use the displacement gradient tensors, \( H^L \) and \( H^E \). These are defined by

\[ H^L = \frac{\partial(x - X)}{\partial X} = F - I, \quad (4.22) \]

\[ H^E = \frac{\partial(x - X)}{\partial x} = I - F^{-1}. \quad (4.23) \]

In the case of small deformation, \( H^L \) and \( H^E \) will both be close to the zero tensor. We can now substitute \( F = I + H^L \) and \( F^{-1} = I - H^E \) into the definitions of
C and $B^{-1}$ given in equations (4.13) and (4.14). This yields

$$C = (I + H^L)^T (I + H^L) = I + H^L + (H^L)^T H^L,$$

$$B^{-1} = (I - H^E)^T (I - H^E) = I - H^E - (H^E)^T H^E.$$  

Now, we note that the $(H^L)^T H^L$ and $(H^E)^T H^E$ terms will be much smaller than $H^L$ and $H^E$ respectively. Thus, to leading order they can be neglected. We can now substitute the expressions above (without the quadratic terms) into equations (4.20) and (4.21) to find that

$$E^L \simeq e^L = \frac{1}{2} \left( H^L + (H^L)^T \right), \quad (4.24)$$

and

$$E^E \simeq e^E = \frac{1}{2} \left( H^E + (H^E)^T \right). \quad (4.25)$$

Equivalently, we can express equations (4.24) and (4.25) in terms of $F$ as follows:

$$e^L = \frac{1}{2} (F + F^T) - I,$$

and

$$e^E = I - \frac{1}{2} (F^{-1} + F^{-T}).$$

The tensors $e^L$ and $e^E$ are called the infinitesimal strain tensors. Since they are much easier to calculate and manipulate than the finite strain tensors ($E^L$ and $E^E$), the infinitesimal strain tensors are used extensively in engineering and practical mechanics. However, they can only be used as valid representations of strain in the case where all deformations are small.

Moreover, the infinitesimal strain tensors are sensitive to rotations. In this sense, they are not true measures of deformation because they are not invariant to all
rigid motions.

4.3 Stress, strain and the zero stress state

4.3.1 Stress and plasticity

The internal forces acting on a continuum are represented using the stress tensor, \( \sigma \). The concept of stress is fundamental to continuum mechanics, and detailed descriptions may be found in Spencer [176] or Malvern [118]. Stress has units of force per area and is effectively a measure of the internal forces exerted over an arbitrary internal area element. Generally, the stress tensor \( \sigma \) is defined so that if \( \delta A \) is an infinitesimal area element with normal vector \( \hat{n} \) and \( F \) is the total force acting over the surface \( \delta A \),

\[
F = \sigma \cdot \hat{n} \delta A.
\]

Note that, as with strain, there are multiple definitions of stress depending on the coordinate system used. The most physically meaningful of these is the Cauchy stress, which is defined entirely with reference to the current (Eulerian) configuration. In Lagrangian coordinates, stress is defined using the first and second Piola-Kirchoff stress tensors, which do not lend themselves to physical interpretation as easily as the Cauchy stress tensor. Spencer [176] and Malvern [118] show that the Cauchy stress tensor and the second Piola-Kirchoff tensor are symmetric (due to the balance of angular momentum), but this is not necessarily true of the first Piola-Kirchoff stress tensor.

In an elastic solid, the presence of a stress will generally lead to a strain. That is, a solid will become stretched or compressed in response to the internal forces that it experiences. The relationship between stress and strain depends strongly on the material being considered and is defined using an equation called the constitutive
law for the material. This relationship is normally bijective and it is generally convenient to express the stress as a function of strain.

The simplest example of a constitutive law for a solid is the isotropic Hooke’s law. This is given by

$$\sigma_{ij} = \lambda \delta_{ij} E^E_{kk} + 2 \mu E^E_{ij},$$

(4.26)

where $\sigma$ is the Cauchy stress tensor, $E^E$ is the Eulerian strain and $\lambda$ and $\mu$ are the elastic (or Lamé) coefficients, which are characteristic of the solid being considered [176]. The isotropic Hooke’s law is appropriate for describing simple, isotropic, perfectly elastic solids such as steel under small deflections, but more complex materials require more complex constitutive equations.

Many materials undergo plastic deformations when they are stressed beyond a certain limit (known as the yield stress). Once the yield stress is reached, the particles of the solid are no longer bound together in a fixed lattice. Instead, the particles slip relative to each other, allowing the strain to change significantly without the corresponding change in stress that would be predicted by Hooke’s law. If the load which caused the plastic deformation is then removed, it is normally found that the relationship between stress and strain is different to the original relationship. This change to the constitutive law reflects the hardening and irreversible deformation of the material that took place during the plastic deformation.

In classical plasticity theory, the most common way to deal with these plastic deformations is to reformulate the stress-strain relationship in terms of infinitesimal changes. That is, instead of stating the constitutive law as an explicit relationship between $\sigma$ and $E$, we deal with infinitesimal changes in stress and strain, represented as $d\sigma$ and $dE$.

The first step in this approach is to separate $dE$ into elastic and plastic components so that

$$dE = dE^e + dE^p.$$
Now, we write separate constitutive laws for these two components of strain. The relationship between $dE^e$ and $d\sigma$ is generally taken to be linear, often using the isotropic Hooke’s law given in (4.26). In contrast, $dE^p$ is defined so that it is only nonzero during plastic flow and its constitutive law is normally derived using thermodynamic arguments about the irreversible work done. For more details of the classical treatment of plasticity, see Hill [89].

This plasticity theory is very useful for describing polycrystalline metallic materials. However, such materials generally only undergo plastic changes to their fundamental structure in response to physical stress or a change in temperature. In contrast, biological materials display much more complicated ‘plastic’ behaviour. Some biological plasticity is like metal plasticity in that it occurs in response to mechanical stress. More often, however, the fundamental mechanical structure of a tissue is directly modified by the action of cells. In this case, the thermodynamic arguments used to specify $dE^p$ for a metal are no longer useful and we need to develop some other constitutive law for tissue plasticity.

One alternative to the infinitesimal approach of classical plasticity is to focus on the changes to the zero stress state of the body being considered. That is, we separate the current deformation into an elastic part and a plastic part (not necessarily by writing the total strain as a sum of elastic and plastic strains). Then, we use evolution equations to describe how the elastic and plastic deformations change over time.

In Section 4.2.1, we mentioned the possibility of defining a zero stress coordinate system in which the current position and the original position of a particle are both considered to be functions of its position in a hypothetical zero stress form of the body. This would give us independent functions that describe the elastic and plastic components of deformation. However, although this approach allows us to describe some plastic deformations, it is unable to deal with residual stresses.

As we described in Section 2.4, Lee [107] proposed a mathematical model of plasticity that uses a multiplicative decomposition of the deformation gradient.
Later, this model was introduced to biomechanical modelling by Rodriguez et al. [159]. In the following section, we investigate the approach to biological plasticity taken by Rodriguez et al. [159]. We describe an intuitive but rigorous way of approaching the multiplicative decomposition of the zero stress state that leads us naturally to the definitions of strain that we develop in Section 4.3.3.

4.3.2 Residual stresses and the zero stress state

As well as showing plastic behaviour, many biological materials experience residual stresses. Residual stress occurs when a body is bound to itself in such a way that it cannot deform to relieve its internal stresses. For example, consider a flat sheet of a flexible, elastic metal that has been bent into a cylinder and then welded to itself. Because of the weld, the metal sheet cannot return to being flat; thus, even when the force that was used to bend the sheet into a cylinder is released, the sheet remains bent and under stress. If we cut the cylinder, it will spring open (and, in the absence of plastic effects, it will return to being a flat sheet), thus relieving the residual stresses. Similarly, the residual stress in the cylinder can be relieved by annealing (i.e. heating the cylinder and then allowing it to cool slowly). In this case, the particles within the cylinder rearrange themselves so that they are no longer under stress in their current configuration.

The classic example of residual stress in biology was first described by Chuong and Fung [28] when they showed that arterial cross-sections open significantly when a longitudinal cut is made. That is, cylindrical arteries experience residual stresses much like the metal cylinder described in the previous paragraph. When the cut is made, the artery deforms out of its cylindrical shape (although not to a completely flat sheet) and the residual stresses are relieved [62].

\[\text{Note that it is not necessarily true that all residual stresses will be relieved with a single cut. Theoretically, there is no limit to the number of cuts that would be necessary to relieve the residual stresses of a body. However, Fung and coworkers [63, 86] found that one cut is sufficient to relieve almost all of the residual stress within an artery, although further cuts are necessary if the artery wall is thick.}\]
and Taber [180].

When determining how a body responds to an applied force, it is necessary to take residual stresses into consideration. In fact, the different mechanical properties of residually stressed materials has led to their extensive use in industrial applications. For example, prestressed concrete (a residually stressed material) is commonly used in large buildings because it can tolerate greater tensile stresses than reinforced concrete. In the rest of this section, we will introduce and describe some of the mathematical techniques that have been used to model and describe residual stress (and plasticity) in biomechanical applications.

In order to account for residual stresses in growing tissues, Rodriguez et al. [159] proposed a model in which the deformation gradient $F$ (written as $F_{eg}$ in Rodriguez’s notation) of a growing tissue is expressed as the product of a growth component, $F_g$, and an elastic component, $F_e$, in the form

$$F = F_e F_g.$$  

As we will see, $F_g$ effectively maps infinitesimal vectors in the initial configuration of the body to infinitesimal vectors in a local zero stress configuration of the body while $F_e$ maps infinitesimal vectors in the local zero stress configuration to infinitesimal vectors in the current configuration. That is, $F_g$ describes the deformation from ‘how the body was’ to ‘how the body would like to be’, while $F_e$ describes the deformation from ‘how the body would like to be’ to ‘how the body is’. A graphical representation of these deformations (using our notation of $Z \equiv F_g$ and $Y \equiv F_e^{-1}$) is shown in Figure 4.2. We will now formalise these concepts mathematically and show how they can be used in practical contexts.

The elastic part of the deformation, $F_e$, is used for the calculation of strain and, from this strain, elastic stress. Hence, if $F_e = Q$ at a point, where $Q$ is a proper

---

13Later in this section, we will introduce our own notation for these two tensors. The tensor $F_g$ is equivalent to $Z$, the Lagrangian zero stress deformation gradient, while $F_e$ is equivalent to $Y^{-1}$, the inverse of the Eulerian zero stress deformation gradient. Lee [107], Cook [33] and Goriely et al. [70, 71] each have their own system of notation.
Zero stress state theory and the definition of strain

Figure 4.2: Lee [107] and Rodriguez et al. [159] both proposed that residual stresses could be accounted for by decomposing the deformation from the initial configuration to the current configuration as the product of an elastic deformation and a growth deformation. This is illustrated above.

The initial configuration is shown as a regular grid. In order to be at a state of zero stress, some of the elements of this grid ‘want’ to be smaller than they are in their initial configuration. In our example, it is impossible for the entire grid to be simultaneously stress-free without tearing. In this case, the zero stress state is said to be incompatible.

For a given zero stress state, it is generally possible to calculate the current configuration if the bijective relationship between stress and strain is known. This is a direct result of the fact that linear momentum must be conserved. If inertial effects are ignored, this means that the divergence of the stress tensor at any point must be equal to the body force. Thus, it is possible to use the boundary conditions and body forces to calculate the distribution of stress throughout the body; using this, we can determine the strain (and hence the current deformation).

In our notation, $F$ represents the overall deformation gradient from the initial configuration to the current configuration, while $Z$ (the Lagrangian zero stress deformation gradient) represents the deformation from the initial state to the zero stress state and $Y$ (the Eulerian zero stress deformation gradient) represents the deformation from the current state to the zero stress state.
orthogonal tensor, there is taken to be no stress at that point. Residual stresses arise when, for a given $F_g$ (i.e. for a given description of ‘how the body would like to be’), it is not possible for $F_e$ to be proper orthogonal at every point in the body while retaining the property that $F$ is the gradient of some universally defined function (the deformation map). In this case, it is not possible for the body to experience zero stress everywhere unless it has been torn.

It is instructive to compare the approach of Rodriguez et al. with the zero stress coordinate system we described in Section 4.2.1. Within the zero stress coordinate system, we deal with plastic deformations by considering the mapping from the initial state to the zero stress state separately to the mapping from the zero stress state to the current state. That is,

$$\chi = a^{-1}(X, t),$$
$$x = b(\chi, t);$$

where $\chi$ is a position vector in zero stress state coordinates.

The approach developed by Rodriguez et al. effectively involves the deformation gradients associated with these mappings. That is, for given functions $a^{-1}$ and $b$ defining the zero stress state, we can construct tensors

$$F_g = \frac{\partial a^{-1}}{\partial X} = \frac{\partial \chi}{\partial X},$$

and

$$F_e = \frac{\partial b}{\partial \chi} = \frac{\partial x}{\partial \chi}.$$  

Now, we note that

$$F = \frac{\partial x}{\partial X} = \frac{\partial x}{\partial \chi} \frac{\partial \chi}{\partial X} = F_e F_g.$$  

Also, we find that $F_e = Q$ corresponds to the case where $x = b(\chi, t)$ is a rigid motion (i.e. there is no stress induced in going from the zero stress state to the
current state).

It follows that the $F_g$ and $F_e$ defined by the deformation gradients obtained from a zero stress coordinate system are identical to the $F_g$ and $F_e$ defined by Rodriguez et al. This means that given an appropriately defined zero stress coordinate system, it is always possible to construct a multiplicative decomposition of the deformation gradient in the form described above. However, it should be noted that the converse is not true; $a^{-1}$ and $b$ cannot necessarily be constructed from $F_g$ and $F_e$.

To show that this is the case, we note that the ‘$i$’th row of $F_g$ contains the components of $\text{grad } a_i^{-1}$, where $a_i^{-1}$ is the ‘$i$’th component of $a^{-1}$. Thus, going backwards from $F_g$ to $a^{-1}$ is equivalent to solving for three different scalar fields, $a_i^{-1}$, given $\text{grad } a_i^{-1}$ in each case. This will only be possible if each row of $F_g$ yields a conservative (i.e. irrotational) vector field. Hence, it is not always possible to obtain $a^{-1}$ from $F_g$. By the same argument, it can also be shown that it is not generally possible to find $b$ given $F_e$.

This means that the technique used by Rodriguez et al. is more general than the technique of zero stress state coordinates which we outlined in Section 4.2.1. Importantly, defining the zero stress state using deformation gradient tensors instead of a coordinate system allows us to consider situations in which residual stresses arise (i.e. situations where $F_g$ is not a real deformation of the body and so $F_e = Q$ is never possible).

In fact, this is one of the main applications of the multiplicative decomposition of the deformation gradient. If a given choice of $F_g$ cannot be used to define a universal zero stress coordinate system, the zero stress state is called incompatible. For a clear and detailed description of the different types of incompatibilities and how they can arise, see Skalak et al. [175].

Given a body with an incompatible zero stress state, the residual stress distribution is defined to be the stress distribution that arises when the body is unloaded.
(i.e. when there are no body forces and all boundaries are stress-free). In order to determine the residual stress distribution of a body, it is necessary to know the constitutive relationship between stress and strain as well as $F_g$. This is because the residual stress distribution must satisfy the conservation of linear momentum; that is, the internal forces experienced by the body must be in balance.

Finding the residual stress distribution when the zero stress state is incompatible is a complex and interesting mathematical problem. A detailed description of this work is beyond the scope of this thesis, but see especially the recent work of Klarbring and Olsson [101, 143] and Ben Amar and Goriely [18] for some promising computational results.\footnote{Both of these research teams concentrate on the biologically relevant problem of modelling residual stresses in hyperelastic incompressible materials. In contrast, we focus on problems where it is appropriate to use more general theories of linear elasticity. It would be interesting to extend our work to incorporate the more realistic constitutive laws considered by these researchers.}

The multiplicative decomposition of the deformation gradient has also been used in models of growth and other biomechanical processes where plastic changes are important [113, 159]. In these cases, the zero stress state of a tissue may change, prompting changes to its shape, size and/or residual stress distribution. Importantly, if we use an evolving zero stress state to model plastic change instead of the relationship between differentials described in Section 4.3.1, we retain the property that zero elastic strain (i.e. $F_e = Q$) corresponds to zero stress. Also, it is simpler and more intuitive to construct a constitutive law for the evolution of the zero stress state instead of employing the arguments of classical plasticity theory.\footnote{It is also likely that tissue growth and similar problems are very difficult or even impossible to express using classical plasticity theory. Classically, plastic changes occur when some measure of the stress tensor crosses a yield threshold. In contrast, biological plasticity occurs continuously; there is no stress condition required for growth to occur. More importantly, the classical theories developed to describe what happens when the yield stress is exceeded generally depend on the principal that the volume change (or, at the very least, the mass change) associated with plastic change is zero [89].

However, most models of biological tissues are not mass-conserving; cell-division, cell death, fibre synthesis and fibre degradation are all treated as increases or decreases to the total mass of cells or ECM within the system. Although it may still be possible to use classical plasticity in this case, it would probably be necessary to restore the conservation of mass by considering a...}
In recent years, Hoger and coworkers have made significant progress in developing and justifying a sound foundation for zero stress state theory, specifically as it applies to growth [26, 94, 113]. Their presentation of the zero stress state arises from considering how the stress response function of a body can be related to its strain energy. In turn, the strain energy is measured with reference to the zero stress state.

Here, we present our own theoretical approach (analogous to, but distinct from, Hoger’s work) by considering hypothetical deformations away from the current configuration, similar to Chen and Hoger’s pure deformations [26]. In Section 4.3.3 we then modify the classical definition of strain \((i.e.\ equations\ (4.16)\ and\ (4.17))\) in order to obtain alternative strain formulations that incorporate the evolving zero stress state.

Although the multiplicative decomposition is a physically intuitive way of dealing with the zero stress state, it should be noted that some theoretical difficulties that need to be addressed remain. For example, Xiao et al. [196] comment on the fact that the nonuniqueness of the multiplicative decomposition makes it difficult to clearly define objective time derivatives. We will encounter this in Section 5.3.3, where we find that our representation of the changing zero stress state is multiply defined. Despite the significant contribution of Hoger and other researchers, there are still many unanswered theoretical questions about the representation of the zero stress state that deserve further investigation.

Now, let \(X, x\) and \(f(X, t)\) be defined as before, so that

\[
x = f(X, t)
\]

defines the deformation of a body \(B\) from its original state to its configuration at time \(t\). Also, let \(X_0\) be the position vector in Lagrangian coordinates of some particle \(P\) in \(B\). We now consider a hypothetical deformation \(g(X, t; X_0)\), called a stress-relieving deformation, defined so that if, at time \(t\), the body immediately\(^{16}\) multiphase model. This would add further layers of complexity to the model being constructed.\(^{16}\)We require an immediate change in order to eliminate the possibility of plastic flow occurring
changed from $x = f(X, t)$ to $x = g(X, t; X_0)$, the elastic stress tensor evaluated at $X_0$ would be zero. That is, $g$ is some deformation of $\mathcal{B}$ which locally relieves the stress at $X = X_0$.

For example, let us consider how $g$ would be defined in the simplest possible case: a perfectly elastic body with no residual stress. In this situation, we can define the stress-relieving deformation, $g$, generally by noting that all the points in $\mathcal{B}$ will be at zero stress when there is no deformation from the body’s initial state to its current state. Thus,

$$g(X, t; X_0) = X$$

would be an appropriate stress-relieving deformation for all possible values of $t$ and $X_0$.

However, this choice of $g$ is not unique, as we note that any rigid motion from the initial condition would also lead to an unstressed body. That is,

$$g(X, t; X_0) = QX + c$$

is also a valid stress-relieving deformation for all values of $t$ and $X_0$ in a perfectly elastic body, where $Q$ is a proper orthogonal tensor and $c$ is a vector.

In fact, for a general body we find that it is possible to find infinitely many stress-relieving deformations for any given point (as long as it is possible to find at least one stress-relieving deformation at that point). This is because following a deformation with a rigid motion will not change the stress distribution within the body.\(^\text{17}\) Thus, if $g$ is a stress-relieving deformation at $X_0$, then so is $\hat{g}$ where

$$\hat{g}(X, t; X_0) = Qg(X, t; X_0) + c.$$

---

\(^{17}\)There are potentially other ways of changing $g$ that leave the stress-relieving nature of the deformation unchanged. In general, if $g$ is followed by a further deformation that does not affect the neighbourhood of $X_0$, the result will be a valid stress-relieving deformation.
For a perfectly elastic body, it is always possible to find a stress-relieving deformation that is independent of \( \mathbf{X}_0 \) and \( t \). However, if we introduce plastic effects, the stress-relieving deformation will need to change over time to reflect this. Similarly, if there are residual stresses, it will not be possible to find a single deformation of the body that relieves the stress at all points and hence \( g \) will depend on \( \mathbf{X}_0 \).\(^{18}\)

In the most general case, we must consider \( g \) to be a function that is not uniquely defined and which varies with \( \mathbf{X} \), \( t \) and \( \mathbf{X}_0 \). Interestingly, we note that the stress-relieving deformation must vary smoothly with \( \mathbf{X} \) for any given choice of \( \mathbf{X}_0 \) in order to correspond to a real, continuous deformation. However, there is no restriction on \( g \) requiring it to vary smoothly with \( \mathbf{X}_0 \) or even with \( t \).

Now, let \( \mathbf{Z} \) be a tensor, which we will call the Lagrangian zero stress deformation gradient tensor,\(^{19}\) that is defined so that

\[
\mathbf{Z}(\mathbf{X}_0, t) = \frac{\partial g(\mathbf{X}, t; \mathbf{X}_0)}{\partial \mathbf{X}} \bigg|_{\mathbf{X}_0}.
\]

That is, \( \mathbf{Z} \) is the deformation gradient tensor of the stress-relieving deformation evaluated at the point at which it relieves stress. This means that if \( \mathbf{F} = \mathbf{Z} \) at any point in \( \mathcal{B} \), the stress at that point is zero.

We recall that \( g \) is not uniquely defined. As a result, it is possible to show that \( \mathbf{Z} \) is not uniquely defined. To see this, let \( g \) be some stress-relieving deformation and let \( \mathbf{Z} \) be the associated zero stress deformation gradient. Using the definition of a stress-relieving deformation, we recall that \( \mathbf{\hat{g}} = \mathbf{Q} g + \mathbf{c} \) will also be a stress relieving deformation if \( \mathbf{Q} \) is proper orthogonal. We can use this \( \mathbf{\hat{g}} \) to generate a new zero stress deformation gradient tensor. Generally, if \( \mathbf{Z} \) is a valid zero stress deformation gradient tensor, it is defined by

\[
\mathbf{Z}(\mathbf{X}_0, t) = \frac{\partial \mathbf{g}(\mathbf{X}, t; \mathbf{X}_0)}{\partial \mathbf{X}} \bigg|_{\mathbf{X}_0}.
\]

\(^{18}\)Interestingly, we note that it is only possible to define a zero stress coordinate system when \( g \) is independent of \( \mathbf{X}_0 \). This explains why it is impossible to represent residual strains using a zero stress state system.

\(^{19}\)In unambiguous cases, we will refer to \( \mathbf{Z} \) as the zero stress deformation gradient tensor.
deformation gradient, then so is

\[ \tilde{Z} = QZ, \]

where \( Q \) is proper orthogonal.

We generally expect that \( Z \) is unique up to premultiplication by \( Q \); however, it is theoretically possible for two or more rotational ‘families’ of zero stress deformation gradient tensors to exist. In Appendix B, we show that this will not occur in most physically realistic situations.

In the case where only one such family of zero stress deformation gradients exists, we can easily construct a unique measure of the zero stress state. In particular, consider the tensor defined by

\[ \Theta = Z^T Z. \]

We note that if \( Z \) is replaced by \( \tilde{Z} \) as defined above, the tensor \( \Theta \) remains unchanged. Thus, \( \Theta \) is a unique representation of the zero stress state. We will refer to \( \Theta \) as the zero stress state Cauchy-Green tensor.

The tensor \( \Theta \) is effectively the right Cauchy-Green tensor associated with a deformation gradient of \( Z \). As a result, we find that \( \Theta \) is symmetric positive definite. Similarly, we could define the right stretch tensor \( U_Z \) associated with \( Z \) and use it as unique representation of the zero stress state. This will also be symmetric positive definite.

In Section 4.4, however, we show that \( U_Z \) is not necessarily a useful way of representing \( Z \). Thus, we will generally prefer to use the zero stress deformation gradient as a representation of the zero stress state, despite the fact that it is not unique to rotations.

It should be pointed out that all of the analysis in this section has depended on the assumption that for any given \( t \) and \( X_0 \), stress-relieving deformations will
exist. In Appendix B, we discuss hypothetical constitutive laws where the zero stress state (as represented by $\Theta$) does not exist or is not unique. However, by developing sufficient conditions for existence and uniqueness we are also able to show that these constitutive laws are mathematical pathologies that are unlikely to arise in real materials. For the rest of our analysis we will assume the existence and uniqueness of the zero stress state Cauchy-Green tensor, $\Theta$. That is, we will assume that the zero stress deformation gradient is unique up to premultiplication by a proper orthogonal tensor.

4.3.3 The improved definition of strain

Consider a particle $P$ in the body $\mathcal{B}$ with initial position vector $X_0$ and another particle $Q$ in the neighbourhood of $P$ with initial position vector $X_0 + dX$. Now, let $\mathcal{B}$ be deformed so that at time $t$, the particle $P$ is at zero stress. That is, the deformation at time $t$ is a stress-relieving deformation around $X = X_0$. As $\Theta$ is the right Cauchy-Green tensor associated with all stress-relieving deformations, it follows from the definition of the right Cauchy-Green tensor in equation (4.13) that
\[ dL^2 = dX^T \Theta dX, \] (4.27)
where $dL$ is the distance between $P$ and $Q$ after the stress-relieving deformation.

Now, $dL$ can be thought of as indicating the ‘desired’ distance between the particles $P$ and $Q$. That is, it is necessary for the distance between $P$ and $Q$ to be $dL$ in order for there to be zero stress between $P$ and $Q$. For a general deformation, we expect the stress at $P$ in the direction of $dX$ to depend in some way on the difference between $ds$, the current distance between $P$ and $Q$, and $dL$, the distance between the same two points after a stress-relieving deformation.\(^{20}\)

We recall that strain is conventionally defined according to the difference beween

\(^{20}\)As before, we are assuming a bijective relationship between stress and deformation. In actuality, the deformation depends on the stress.
square distances in the current and initial configurations. That is,

\[(ds)^2 - (dS)^2 = 2dX^T E_{\text{classical}}^L dX,\]
\[(ds)^2 - (dS)^2 = 2dx^T E_{\text{classical}}^E dx;\]

as given in equations (4.16) and (4.17).

In order to take into account the changing zero stress state, we replace \(dS\) with \(dL\); that is, we define the ‘effective strain’\(^{21}\) using the difference between the current and stress-free states rather than the difference between the current and initial states. Thus,

\[(ds)^2 - (dL)^2 = 2dX^T E^L dX,\] \hspace{1cm} (4.28)
\[(ds)^2 - (dL)^2 = 2dx^T E^E dx.\] \hspace{1cm} (4.29)

We recall that we were able to express the classical definition of strain in terms of the Cauchy-Green tensors. Thus, it seems natural to seek to define our new definition of strain in terms of the zero stress state Cauchy-Green tensor. According to equation (4.27),

\[dL^2 = dX^T \Theta dX.\]

Furthermore, we note from equation (4.9) that

\[dX = F^{-1} dx,\]

and from equation (4.13) that

\[ds^2 = dX^T C dX,\]

where \(F\) is the deformation gradient tensor and \(C\) is the right Cauchy-Green

\(^{21}\)In general, we will use \(E\) to represent the effective strain and \(E_{\text{classical}}\) to represent the classical strain. Throughout the rest of this thesis, the term ‘strain’ refers to the effective strain rather than the classical strain.
tensor.

Applying these results, we find that

\[ (ds)^2 - (dL)^2 = dX^T (C - \Theta) dX, \]

and

\[ (ds)^2 - (dL)^2 = dx^T (I - F^{-T} \Theta F^{-1}) dx. \]

Substituting into equations (4.28) and (4.29), we obtain the following expressions for the Lagrangian and Eulerian strain tensors:

\[
E^L = \frac{1}{2} (C - \Theta) = \frac{1}{2} (F^T F - Z^T Z), \tag{4.30}
\]

\[
E^E = \frac{1}{2} (I - F^{-T} \Theta F^{-1}) = \frac{1}{2} (I - F^{-T} Z^T Z F^{-1}). \tag{4.31}
\]

Equation (4.31) can be simplified by introducing \( Y \), the Eulerian zero stress deformation gradient. We define \( Y \) so that

\[ Y = Z F^{-1}. \tag{4.32} \]

This means that \( Y \) can be used to relate vectors in the Eulerian configuration with vectors in a locally-constructed zero stress coordinate system.\(^{23}\) That is, if \( dx \) is a vector in the current configuration connecting two neighbouring particles, the vector \( dz = Y dx \) has a length that represents the distance between those particles at zero stress.

\(^{22}\)Rodriguez et al. [159] use a different definition of Lagrangian strain. As we discuss in Appendix C, their version of strain is difficult to interpret as well as being potentially ill-defined.

\(^{23}\)We recall that we cannot generally construct a globally valid zero stress coordinate system. However, it is always possible to construct a zero stress coordinate system (indeed a family of zero stress coordinate systems) in the neighbourhood of a given particle.
Given $Y$ defined above, we find that equation (4.31) yields

$$E^E = \frac{1}{2}(I - YT)Y. \quad (4.33)$$

In Eulerian problems, $Y$ has a number of advantages over $Z$. For example, equation (4.33) is directly analogous to the classical definition of Eulerian strain with $F^{-1}$ replaced by $Y$. Thus, we can use the method described in Section 4.2.5 to construct an infinitesimal strain measure based on $Y$:

$$e^E = I - \frac{1}{2}(Y + YT). \quad (4.34)$$

A second advantage of $Y$ is that it allows us to completely avoid the initial configuration in our definitions of stress and strain. This is particularly useful in biological problems; the constituent fibres of most tissues are continually being synthesised and degraded and thus it is difficult to define the ‘original position’ of the present fibres. Hence, referring all mechanical changes to the initial configuration can lead to conceptual problems.

However, this last advantage is not without its problems. Having circumvented all reference to the initial configuration, it is no longer possible to see how the current deformation separates into an ‘elastic’ part and a ‘growth/plastic’ part. Similarly, it becomes more difficult to use the model to predict particle streamlines. This is a significant problem because a number of interesting experiments have been performed in which particle streamlines are tracked.\(^\text{24}\)

\(^{24}\)For example, Roy et al. \cite{163} implanted a collagen lattice with small fluorescent plastic balls and tracked the position of these balls as the lattice was contracted by fibroblasts. Dorie et al. \cite{48, 49} performed similar experiments on tumour spheroids.
4.4 Infinitesimal strain approximations

Equations (4.30) and (4.31) are exact representations of the strain, analogous to the classical definitions given in equations (4.16) and (4.17). However, many common materials respond to the stresses that they normally experience with very small deformations. In classical elasticity, the assumption of small deformations (\(i.e. \ F \simeq I\)) leads to the infinitesimal strain approximations given in equations (4.24) and (4.25), which are used in the theory of linear elasticity.

In this section, we present a novel simplification of equations (4.30) and (4.31) analogous to classical infinitesimal strain. Our approximation is applicable in the case that the deformation away from the zero stress state is small (\(i.e. \ F \simeq Z\), or equivalently, \(Y \simeq I\)). This means that our infinitesimal strain approximation is able to deal with finite deformations in the case that most of the observed deformation can be attributed to the effects of plastic change. That is, the formulation for strain that we develop is most appropriate for materials which yield at small strains and do not exhibit large elastic deformations.

It should be noted that there are two other infinitesimal strain approximations incorporating the zero stress state in the present literature. Cook [33] developed an infinitesimal strain approximation for the case where the elastic and plastic components of deformation are both very small (\(i.e. \ F \simeq I\) and \(Z \simeq I\)). We will describe Cook’s infinitesimal strain later in this section and compare it to our own formulation.

The second infinitesimal strain approximation is that developed by Hoger [92]. Hoger’s infinitesimal strain is applicable for any case where the observed deformation (as represented by the right stretch tensor, \(U\)) is small. That is, Hoger’s infinitesimal strain is valid for arbitrary changes to the zero stress state and arbitrary physical rotations of the body being modelled. This contrasts with Cook’s work where both the deformation and the changes to the zero stress state are required to be small. Hoger’s infinitesimal strain has greater flexibility because
of the representation of the zero stress state that she uses. Hoger models the zero stress state by tracking the residual stress of an undeformed body \((i.e. \text{ the stress, } \sigma, \text{ associated with } F = I)\), whereas Cook, like us, considers the deformation associated with zero stress.

Thus, Hoger’s infinitesimal strain is able to deal with the case where the difference between the current state and the zero stress state is arbitrarily large, as long as the deformation from the initial state to the current state is relatively small. This is particularly useful when modelling tissues like arteries, where size and shape do not change substantially over time but where considerable changes to the zero stress state may lead to significant residual stresses.

In contrast, our infinitesimal strain is able to deal with arbitrarily large deformations, as long as the deformation from the zero stress state to the current state is relatively small. This will be more useful when modelling growing tissues such as developing embryos or solid tumours, where there are significant changes in size but the strains associated with residual stress are small. In wound healing, we expect both the residual strain and the total deformation to be small. Thus, it is reasonable to expect that either Hoger’s approach or our approach would be valid.

As our infinitesimal strain approximation depends on different assumptions to Hoger’s formulation, a direct comparison between the two is not meaningful or useful. Instead, the two different approaches are complementary; the choice of infinitesimal strain formulation depends on the nature of the modelling problem being considered.

### 4.4.1 The principal zero stress deformation gradient

Throughout this section, we are concerned with the case where the observed deformation is close to being a stress-relieving deformation. Expressed mathe-
matically, this corresponds to a situation where the deformation gradient is close to the Lagrangian zero stress deformation gradient (i.e. $F \simeq Z$) or the Eulerian zero stress deformation gradient is close to the identity (i.e. $Y \simeq I$). However, $Z$ and $Y$ are not uniquely defined. Hence, it is necessary for us to think carefully about whether there is an appropriate choice of $Z$ and $Y$ that minimises the differences $F - Z$ and $I - Y$.

Thus, consider a situation where the deformation gradient, $F$, and the Lagrangian zero stress deformation gradient, $Z_0$, are known. Furthermore, let us assume that the elements of $F$ and $Z_0$ are significantly different but the strain tensors $E^E = \frac{1}{2} (I - F^{-T} Z_0^T Z_0 F^{-1})$, and $E^L = \frac{1}{2} (F^T F - Z_0^T Z_0)$, are both small.

Although $F$ and $Z_0$ are different, the fact that $E^L$ is small implies that the Cauchy-Green tensors associated with these two deformations are similar. This means that the stretch tensors associated with $F$ and $Z_0$ are similar and hence the difference between $F$ and $Z_0$ can largely be ascribed to a rigid rotation.

If $Z_0$ is a zero stress deformation gradient and $Q$ is proper orthogonal, we recall that $Z = QZ_0$ will also be a zero stress deformation gradient. By an appropriate choice of $Q$ we should be able to remove the rigid rotation that separates $F$ and $Z_0$, yielding a zero stress deformation gradient that is close to the observed deformation gradient.\footnote{Importantly, we note that replacing $Z_0$ with $QZ_0$ has no effect on either of the strain tensors.}

We contend that the best choice of $Q$ can be found by performing a polar de-

\footnote{We do not formally define the metric that we use to measure ‘closeness’. Conceptually, two tensors can be said to be ‘close’ if the differences between corresponding elements are very small.}
composition of $Y$, the Eulerian zero stress deformation gradient. As we will see, this leads to some natural simplifications of both the Lagrangian strain and the Eulerian strain. For a given $Z_0$ and a given $F$, there exists a unique $Q$ such that

$$Y = ZF^{-1} = QZ_0 F^{-1}$$

is symmetric.\(^{27}\) We will refer to the $Y$ and $Z$ chosen in this way as the ‘principal zero stress deformation gradients’. Note that the principal Eulerian zero stress deformation gradient is symmetric but the principal Lagrangian zero stress deformation gradient may not be.

Now, consider the effect that this has on the definition of Eulerian strain. In general, we recall that

$$E^E = \frac{1}{2}(I - Y^T Y).$$

In the case where $E^E$ is small, it follows that $Y^T Y$ must be close to being the identity. Thus, $Y$ is close to being an orthogonal tensor. However, we are interested in the case where $Y$ is symmetric positive definite. Importantly, we recall that the only tensor that is both orthogonal and symmetric positive definite is the identity tensor.\(^ {28}\) This leads us to suspect that $Y$ is close to $I$ when the strains are small and $Y$ is chosen to be symmetric. As $Y = ZF^{-1}$, this also implies that $Z$ is close to $F$.

As noted earlier, the argument above is not rigorous. We have not defined what we mean by ‘close’ and we have not proved that there does not exist a different choice of $Q$ that would make $Z$ ‘closer’ to $F$. However, there are a number of

\(^{27}\)In fact, since $Z_0$ and $F^{-1}$ both have positive determinants, the polar decomposition theorem yields the result that $Y$ is symmetric positive definite.

\(^{28}\)It is relatively simple to prove that this is the case. If a tensor $A$ is both symmetric and orthogonal, it follows that $A = A^T = A^{-1}$ and hence $A^2 = I$. Since $A$ is symmetric, it follows that all of its eigenvalues are real. Thus, we conclude that each eigenvalue of $A$ is either 1 or -1. Furthermore, if $A$ is positive definite then all of its eigenvalues are positive. Thus, the only eigenvalue of $A$ is 1. Now, we recall that all symmetric matrices are orthogonally diagonalisable; that is, they can be expressed in the form $QDQ^T$, where $Q$ is the orthogonal tensor of eigenvectors and $D$ is the diagonal tensor of eigenvalues. In our case, we find that $A = QIQ^T$ and thus $A = I$. 
Zero stress state theory and the definition of strain

reasons for us to suspect that choosing \( Y \) and \( Z \) in this manner is useful. In particular, using the principal zero stress deformation gradients is necessary in the analysis of morphoelasticity given in the next chapter.

Furthermore, the principal zero stress deformation gradients can be used to simplify the definitions of strain. As we describe in Section 4.4.3, the Lagrangian infinitesimal strain can be simplified in the case where \( F^T Z \) is symmetric. Interestingly, \( F^T Z \) is symmetric if \( Z F^{-1} \) is symmetric. In order to see this, we note that

\[
F^T Z = F^T Z F^{-1} \quad \text{and} \quad (F^T Z)^T = F^T (Z F^{-1})^T F,
\]

and it follows that \( (F^T Z)^T = F^T Z \) whenever \( (Z F^{-1})^T = Z F^{-1} \).

It is interesting to compare our approach to the principal zero stress deformation gradient with the approaches taken by Lee [107] and Rodriguez et al. [159]. Lee proposed a multiplicative decomposition of the deformation gradient in which the elastic deformation (equivalent to \( Y^{-1} \) in our notation) is symmetric. He argued that this approach is best because it makes it easy to obtain the principal directions of strain.

In contrast, Rodriguez et al. proposed that the growth deformation (equivalent to \( Z \) in our notation) should be chosen to be symmetric. However, this was an

\[29\] It is not important to our central argument, but it is interesting to note that \( F^T Z \) will also be positive definite. We recall that \( Z F^{-1} \) is symmetric positive definite and we note that a tensor is symmetric positive definite if and only if \( x^T A x \) is positive for all nonzero vectors \( x \). Now,

\[
x^T F^T Z x = x^T F^T Z F^{-1} F x = (F x)^T Z F^{-1} (F x).
\]

Hence, the fact that \( Z F^{-1} \) is symmetric positive definite implies that \( F^T Z \) is symmetric positive definite.

\[30\] Note that this implies that \( Y \) will be symmetric. Thus, Lee’s approach is equivalent to ours.
arbitrary choice that was made in order to uniquely define the zero stress state. Rodriguez et al. do not exploit the fact that $Z$ is symmetric in their analysis and they could have stipulated any given rotation for $Z$ without affecting their results. Thus, there is no good reason for us to follow Rodriguez et al. in choosing $Z$ to be symmetric.

### 4.4.2 Eulerian infinitesimal strain

We will develop our infinitesimal strain approximations for Lagrangian and Eulerian coordinates separately. Firstly, consider the definition of Eulerian strain in terms of the Eulerian zero stress deformation gradient, $Y$. As given in equation (4.33), finite strain is defined so that

$$E^E = \frac{1}{2}(I - Y^T Y).$$

By analogy with classical theory, we can use this to obtain the infinitesimal strain approximation given in (4.34):

$$e^E = I - \frac{1}{2}(Y + Y^T).$$

If we use the principal zero stress deformation gradient, this yields

$$e^E = I - Y.$$  \hspace{1cm} (4.35)

It is instructive to compare this definition of $e^E$ with classical infinitesimal strain. Thus, let us consider a classically elastic body with no residual stresses (i.e. a body where the initial state is always a zero stress state). In this case, we find that $Z$ is proper orthogonal; that is, the Lagrangian zero stress deformation gradient can be rotated to obtain the identity tensor. Moreover, if we are using principal zero stress deformation gradients, we find that $Z$ is the proper orthogonal tensor.

---

31Note that this formulation is only appropriate when $Y \simeq I$. 

chosen so that $Y = Z F^{-1}$ is symmetric.

Now, we recall from equation (4.12) that $F$ can be expressed as

$$F = V R,$$

where $V$, called the left stretch tensor, is symmetric and $R$ is proper orthogonal. This means that

$$F^{-1} = R^T V^{-1},$$

where $V^{-1}$ is also symmetric.

This infers that $V^{-1}$ is the symmetric tensor obtained from a polar decomposition of $F$ and thus,

$$Y = V^{-1}.$$

Substituting into our expression for infinitesimal strain yields

$$e^E = I - V^{-1}. \quad (4.36)$$

In contrast, the classical infinitesimal strain tensor, which we defined in equation (4.25), takes the form

$$e_{\text{classical}}^E = \frac{1}{2} \left( H^E + H^{ET} \right),$$

where $H$ is the displacement gradient, which we defined in equation (4.23). This equation can be rearranged to yield the alternative definition,

$$e_{\text{classical}}^E = I - \frac{1}{2} \left( F^{-1} + F^{-T} \right).$$

Now, it is well known that classical infinitesimal strain is only valid for small rotations (i.e. cases where the rotation tensor of the deformation, $R$, is close to the identity) [176]. Interestingly, we find that the case of no rotation (i.e. $R = I$) yields

$$F = F^T = U = V.$$
In the absence of rotations, this means that

$$e^E_{\text{classical}} = I - V^{-1} = e^E.$$  

Thus, the two infinitesimal strains match each other exactly if $R = I$.

Hypothetically, $e^E = I - V^{-1}$ could be used to replace $e^E_{\text{classical}}$ in a general elastic problem. In fact, it could be used in a wider range of situations because it is valid for arbitrary rotations.\(^{32}\) However, $e^E_{\text{classical}}$ has the advantage of being expressed as a linear function of the displacement gradient; it is computationally difficult to obtain $V$ from $F$. Thus, $e^E_{\text{classical}}$ is more widely used than $e^E$.

### 4.4.3 Lagrangian infinitesimal strain

Now, let us consider the Lagrangian strain definition given in (4.30):

$$E^L = \frac{1}{2} (F^T F - Z^T Z).$$

As before, we wish to consider the case where $F \simeq Z$. Thus, we let

$$F = Z + \epsilon,$$

where the tensor $\epsilon$ is small.\(^{33}\)

Substituting (4.37) into (4.30), we find that

$$E^L = \frac{1}{2} (Z^T Z + \epsilon^T Z + Z^T \epsilon - Z^T Z + O(\epsilon^2)).$$

\(^{32}\)Mauget and Perré [125] use an equivalent approach in their theoretical development of an anisotropic constitutive law for wood that is valid for general displacements. Interestingly, they also mention the multiplicative decomposition of the deformation gradient but they do not make use of it.

\(^{33}\)As with ‘closeness’, we do not formally define the metric that we use to define ‘smallness’. One appropriate choice would be to use the spectral norm of the matrix representation of the tensor (see, for example, Schott [165]). In this case, a tensor is small if $||\epsilon||$ is small and we are justified in using formal order notation.
and thus,

\[ e^L = \frac{1}{2} (e^T Z + (e^T Z)^T) = \frac{1}{2} (F^T Z + Z^T F - 2 Z^T Z). \] (4.38)

Equation (4.38) would appear to be significantly more complicated than the original equation (4.30). However, it can be simplified in the case where we use the principal zero stress deformation gradient. As we noted earlier, this has the consequence that the tensor \( F^T Z \) is symmetric. Thus, equation (4.38) becomes

\[ e^L = F^T Z - Z^T Z = (F^T - Z^T) Z. \] (4.39)

Interestingly, this implies that

\[ e^L = F^T e^E Z. \]

Although we will not make use of this relationship in the work that follows, this suggests a possible method for converting between Eulerian and Lagrangian coordinates in the case where deformations are small. In contrast, it can easily be seen from equation (4.30) and (4.31) that

\[ E^L = F^T E^E F. \]

Hence, the relationship between the infinitesimal stress tensors is different from the relationship between the finite stress tensors.

### 4.5 The challenge of defining strain

In classical theories of deformation, strain is defined in terms of the difference between a body’s initial configuration and its current configuration. However,
classical strain is not always a useful measure of deformation in the context of biomechanics. The continual turnover of ECM in living tissues means that it is difficult to clearly define the ‘initial configuration’ of a biological material. Moreover, biological materials are often residually stressed and they sometimes exhibit plastic behaviour that is very different from the flows considered in classical plasticity.

In order to deal with these issues, it is useful to introduce the concept of a zero stress state. As described in Section 4.3.2, the zero stress state is a representation of the local deformation required in order to relieve stress around a point. This approach is able to account for residual stresses because we do not require the local zero stress deformations to be globally compatible. That is, it may be impossible for all of the particles in a body to be simultaneously at their zero stress states without tears or superpositions occurring.

Given a mathematical representation of the zero stress state, the effective strain can be defined in terms of the difference between the zero stress state and the current configuration. This alternative definition of strain then has the useful property that zero strain corresponds to zero elastic stress. Furthermore, it is possible to construct Eulerian formulations of effective strain that do not directly depend on the initial configuration in any way. Thus, our approach enables us to avoid the difficulties that arise from trying to determine the initial configuration of an evolving tissue.

This chapter has presented a mathematical development of zero stress state theory that utilises the multiplicative decomposition of the deformation gradient proposed by Lee [107] and refined by Hoger, Lubarda and others [94, 113, 159]. We addressed the problems that arise from the inherent nonuniqueness of any multiplicative decomposition of the deformation gradient and we showed that our definitions of strain are consistent with physical expectations. However, we

\[\text{It is interesting to compare the definitions of strain given in this chapter with the definitions given by Lubarda, Hoger and others [111, 113, 159]. We have defined strain so that it is always a measure of the deformation from the zero stress state to the current configuration. However, other approaches involve strain tensors that represent the difference between the zero stress}\]
have not yet considered the challenging problem of describing the evolution of the zero stress state. This is dealt with in the following chapter.
Chapter 5

Morphoelasticity

5.1 Introduction

Soft tissue growth, morphogenesis and dermal wound healing are all examples of biological processes that involve changes to the fundamental mechanical structure of a tissue. Such processes are biologically important but are difficult to model mathematically. This is because the mechanical changes are often plastic in nature; that is, the zero stress state of the tissue evolves over time, possibly leading to the development of residual stresses.

In the previous chapter, we described one of the existing techniques for incorporating the concept of a zero stress state into a mathematical model. By considering stress-relieving deformations, we justified the use of a multiplicative decomposition of the deformation gradient into an elastic part and a plastic part, a technique which was originally proposed by Lee [107] and later introduced to biomechanics by Rodriguez et al. [159]. Additionally, our approach enabled us to construct appropriate definitions of strain, both finite and infinitesimal, that take into account this decomposition.
However, the processes of growth and healing are dynamic; they involve changes to the zero stress state of a tissue that occur over time. In contrast, the theory developed in the previous chapter is entirely static. Although it allows us to determine the elastic strain associated with a specified zero stress state and a specified deformation, it cannot be used to prescribe how the zero stress state changes over time.

In this chapter, we consider the problem of specifying an evolution equation for the zero stress state (or equally, an evolution equation for strain) that can be used in biological models. Importantly, we note that the classical theory of plasticity will not be appropriate. Classical plasticity theory is well suited to modelling the highly localised plastic changes that occur in metallic materials, but it cannot be used to describe the continual and global changes to the zero stress state that are associated with tissue growth and wound healing.

In place of classical plasticity, we need a theory of *morphoelasticity*. This term was introduced by Goriely and coworkers to describe processes in which an elastic model is coupled to a continually evolving zero stress state [18, 70, 71]. The first theory of biological morphoelasticity was presented by Rodriguez *et al.* [159] and it was later expanded and modified by Hoger and coworkers (see especially Chen and Hoger [26]) and, more recently, by Goriely and coworkers [18, 70, 71].

We propose a novel theory of morphoelasticity that differs from this earlier work in a number of ways. Unlike the incremental in time model of growth proposed by Goriely and coworkers [70, 71], our model is expressed in terms of a continuous time variable. Also, the tensor that we use to describe the rate of growth is different from the tensors used by Rodriguez *et al.* for the same purpose.\(^1\) Interestingly, we find that our theory can be used to criticise, clarify and justify aspects of the earlier models by giving them a clearer physical interpretation.

We develop our theory of morphoelasticity in several stages. Firstly, in Section

\(^1\)In some cases, our approach yields equations that are equivalent to those proposed by Rodriguez *et al.* [159]. However, our approach is more flexible and general; see Section 5.5.2 for more details.
5.2, we construct a one-dimensional model of a growing tissue. The advantage of using one-dimensional Cartesian coordinates is that we do not need to consider the directional nature of growth. Thus, we can clearly define the rate of growth and construct uncontroversial equations to describe the evolution of the zero stress state. Although simple, the equations of one-dimensional morphoelasticity are potentially very useful. For example, the Tranquillo-Murray model of dermal wound healing [189] is expressed in one-dimensional Cartesian coordinates and it could potentially be extended by incorporating aspects of one-dimensional morphoelasticity theory.

Having described the one-dimensional theory in Section 5.2, we develop a three-dimensional theory of morphoelasticity in Section 5.3. Firstly, in Section 5.3.1, we consider only the net rate of growth; this leads us to obtain equations that are very similar to those developed in Section 5.2. However, three-dimensional growth is different from one-dimensional growth in that it is possible for certain directions of growth to be preferred over others. Throughout the rest of Section 5.3, we work towards developing a clear and physically interpretable equation that would allow us to take into account both the rate and the direction of growth. This work culminates in Section 5.3.4, where we state a complete set of equations that can be used to describe the mechanical evolution of a morphoelastic body.

Next, in Section 5.4, we adapt our equations for the changing zero stress state to obtain evolution equations for strain. Since strain is a more intuitive and widely-used variable than the zero stress deformation gradient, we expect that these equations will be easier to interpret and to apply than equations that refer to the zero stress state. Indeed, when we discuss applications of morphoelasticity theory in Chapter 6, we will generally use the strain evolution form of our equations.

In Section 5.5, we compare our results with two existing theories of morphoelasticity. We show that the theory developed by Rodriguez et al. [159] is comparable to our theory in some ways, but that it encounters difficulties when the principal directions of growth are not parallel to the principal directions of strain. In contrast, we show that the cumulative theory of growth proposed by Goriely and
Ben Amar [70] is very similar to our own work. By using Goriely and Ben Amar’s approach to construct a continuous-time model, we are able to demonstrate that there is a simple relationship between the two different tensorial descriptions of growth.

Lastly, in Section 5.6, we summarise the results of the last two chapters. We recapitulate the new concepts that we have introduced, and we describe how our theory might be applied to real-world problems.

Before we begin, it should be noted that the term ‘growth’ will often be used as a convenient way to refer to the changing fundamental mechanical structure of a tissue. Despite the apparent focus on growth, the theory developed in this chapter is a general morphoelastic theory that could be used to describe growth, resorption, healing or any other biological process that involves changes to the zero stress state of a tissue. However, it is convenient for us to use the word ‘growth’ instead of constantly repeating this list of biological processes, particularly as modelling growth is the main application of morphoelasticity theory in the present literature.

### 5.2 One-dimensional morphoelasticity

Before we consider three-dimensional morphoelasticity, let us explore a one-dimensional problem. In so doing, we will formalise the mathematical concepts that are necessary for developing a general theory of morphoelasticity. Furthermore, consideration of the one-dimensional case leads to some useful results that can be directly applied in mathematical models.

Thus, let us consider a one-dimensional solid body that consists entirely of closely-

---

2In fact, there are even non-biological applications for this theory. As we discuss in Section 6.2, a Maxwell viscoelastic fluid can be modelled using morphoelasticity theory. Thus, the equations developed in this chapter could also be applied to polymer rheology.
packed cells (as shown in Figure 5.1). Each cell in the body has a ‘desired length’ that represents the size of the cell in the absence of applied stresses and body forces. Furthermore, we assume that all of the cells are elastic and compressible. That is, the size of a cell will change according to whether it is under compression or tension.

We assume that the cells are generally of similar size (i.e. the desired length does not vary much between cells). However, the cells are able to grow, shrink, divide and collapse. In accordance with conventional elasticity, we assume that there is a bijective relationship between stress and extension for each individual cell. Thus, we can associate the force that a cell exerts on its neighbours with the difference between the cell’s current length and its desired length. If the desired length of a cell is greater than its current length (as it is in Figure 5.2), the cell will apply outward forces on its neighbours and vice versa.

Given this situation, let \( P \) and \( Q \) be two distinct particles within the solid being considered. In order to avoid the ambiguities that can arise when cells divide or collapse, we will generally assume that \( P \) and \( Q \) are both located on the boundaries between cells (as depicted in Figure 5.1). However, it is not necessary to place formal restrictions on the positions of \( P \) and \( Q \).

Now, let \( X \) and \( X + \Delta X \) represent the initial positions of \( P \) and \( Q \) respectively. Furthermore, let \( x(X, t) \) be the position of particle \( P \) at time \( t \) and let \( \Delta x(X, \Delta X, t) \) be defined so that \( x(X, t) + \Delta x(X, \Delta X, t) \) is the position of \( Q \) at time \( t \). Given that \( \Delta x \) is sufficiently large for there to be cells between \( P \) and \( Q \),

---

3It should be noted that very few tissues are packed with cells in this way. For example, the distribution of fibroblasts within the healthy dermis is relatively sparse [41, 76]. One example of a biological material that contains closely-packed cells is a tumour spheroid (see, for example, the images obtained by Sharma et al. [167]). Although we illustrate our one-dimensional theory by using a body packed with compressible cells, our model can be used much more widely.

4Throughout our work, we will assume that it is possible to clearly define and track particles through the body being studied. Although this is standard practice in continuum mechanics, it is more difficult to consistently define a particle when the underlying substrate is undergoing substantial revision. Some researchers have even incorporated the creation and destruction of particles into their models (see, for example, Skalak et al. [174]), but this seems to be unnecessary.
Figure 5.1: A diagram illustrating a growing one-dimensional body composed of discrete cells. The cells are able to grow and divide, changing the distance between $P$ and $Q$. For instance, there are more cells at time $t + \Delta t$ than there are at time $t$. Each cell is compressible and each cell has a desired length that it reverts to in the absence of physical stresses (not shown in this figure). The sum of these preferred lengths over the interval $PQ$ is called $\Delta z$, the desired distance between $P$ and $Q$. The total elastic force that acts over the interval $PQ$ is related to the difference between $\Delta x$, the actual distance between $P$ and $Q$, and $\Delta z$.

Figure 5.2: In many cases, the desired length of a cell will be different from its actual length at the current time. In this case, the cell is smaller than its desired length and so it will push on the neighbouring cells in the direction indicated by the arrows. In order for the system to be in mechanical equilibrium, the cell indicated must also be experiencing equal and opposite forces from its neighbours.
at time $t$, we can define a function $N(x, \Delta x, t)$ to represent the number of cells between $P$ and $Q$.

Similarly, we can use the desired lengths of the cells between $P$ and $Q$ to define an appropriate measure of the zero stress state. Specifically, let $\Delta z(x, \Delta x, t)$ be defined as the sum of the desired lengths of the cells between $P$ and $Q$. That is, $\Delta z$ represents the distance between $P$ and $Q$ when the elastic stress is identically zero over the interval $[x, x + \Delta x]$. If $\Delta z > \Delta x$, the cells between $P$ and $Q$ are experiencing net compression and vice versa.\(^5\)

Lastly, we wish to specify a relationship between the cell population, $N$, and the desired distance, $\Delta z$. Given that the cells are roughly of equal size, it is reasonable to assume that

$$\Delta z = k N,$$  \hspace{1cm} (5.1)

where $k$ is some constant representing the average cell size.

Now, let us consider the case where the cells between $P$ and $Q$ are proliferating. In particular, we will assume for now that $N$ is growing exponentially. That is,

$$\frac{DN}{Dt} = a N,$$  \hspace{1cm} (5.2)

where $a$ represents the net rate of growth.\(^6\)

Combining equations (5.1) and (5.2), we find that the desired distance between $P$ and $Q$ is growing exponentially at the same rate as the cells:

$$\frac{D\Delta z}{Dt} = a \Delta z.$$  \hspace{1cm} (5.3)

\(^5\)Note that some cells might be compressed and others might be experiencing extension. In the absence of body forces, we would find that the stress between $P$ and $Q$ is uniform; however, we cannot generally assume that this is true.

\(^6\)Note that equation (5.2) involves the material time derivative of $N$ rather than a normal partial derivative. The material derivative is necessary because we are considering the rate of growth as we follow individual cells.
Having obtained equations (5.2) and (5.3), we now wish to consider the continuum limit as $\Delta X$ and $\Delta x$ tend to zero. Currently, our equations are expressed in terms of $N(x, \Delta x, t)$ (a number of cells) and $\Delta z$ (a length). By dividing these terms by length, it will become possible to replace $N(x, \Delta x, t)$ with a measure of cell density and $\Delta z$ with the zero stress deformation gradient.

Thus, we define cell density $n(x, t)$ so that

$$N(x, \Delta x, t) \sim n(x, t) \Delta x$$

(5.4)

as $\Delta x$ tends to zero.

Similarly, we define $Y$ and $Z$ so that

$$Z(X, t) = \lim_{\Delta X \to 0} \frac{\Delta z}{\Delta X} \quad \text{and} \quad Y(x, t) = \lim_{\Delta x \to 0} \frac{\Delta z}{\Delta x}.$$  

Following the work described in the previous chapter, it can be shown that $Z$ and $Y$ are one-dimensional zero stress deformation gradient tensors. The variable $Z$ is equivalent to $Z$ and $Y$ is equivalent to $Y$.

Lastly, we define the deformation gradient, $F$, so that

$$F(X, t) = \lim_{\Delta X \to 0} \frac{\Delta x}{\Delta X} = \frac{\partial x}{\partial X}.$$  

Given these definitions, equation (5.2) becomes

$$\frac{D}{Dt} (n \Delta x) = a n \Delta x,$$

or equivalently,

$$\frac{Dn}{Dt} \Delta x + \frac{D\Delta x}{Dt} n = a n \Delta x.$$
Thus, given that $\Delta x$ is small, we find that

$$\frac{D n}{D t} \Delta x + \frac{D}{D t} (F \Delta X) n = a n \Delta x.$$ 

Since $\Delta X$ represents a length in the Lagrangian configuration, we find that it will be constant to the material derivative. This means that

$$\frac{D n}{D t} \Delta x + \frac{D F}{D t} \Delta X n = a n \Delta x. \quad (5.5)$$

Now, velocity is defined so that $v(x, t) = \frac{D x}{D t}$. Importantly, we note that

$$\frac{D F}{D t} F^{-1} = \frac{D}{D t} \left( \frac{\partial x}{\partial X} \right) \frac{\partial X}{\partial x} = \frac{\partial v}{\partial X} \frac{\partial X}{\partial x} = \frac{\partial v}{\partial x}.$$ 

Also, we recall that the material derivative is defined by

$$\frac{D n}{D t} = \frac{\partial n}{\partial t} + v \frac{\partial n}{\partial x}.$$ 

Using both of these results, equation (5.5) becomes

$$\left( \frac{\partial n}{\partial t} + v \frac{\partial n}{\partial x} + \frac{D F}{D t} F^{-1} n \right) \Delta x = a n \Delta x,$$

or equivalently,

$$\left( \frac{\partial n}{\partial t} + v \frac{\partial n}{\partial x} + \frac{\partial v}{\partial x} n \right) \Delta x = a n \Delta x.$$
Thus,

\[
\frac{\partial n}{\partial t} + \frac{\partial}{\partial x}(n v) = a n. \tag{5.6}
\]

This is precisely the equation that we would expect to obtain for cell density when cells are undergoing exponential growth while being advected with the moving ECM.

Interestingly, we can repeat the procedure that we used to find (5.6) to obtain an equivalent evolution equation for \( Y \). Using equation (5.3) as our starting point, we find that

\[
\frac{\partial Y}{\partial t} + \frac{\partial}{\partial x}(Y v) = a Y. \tag{5.7}
\]

That is, the Eulerian zero stress deformation gradient ‘grows exponentially’ while being ‘advected’ with the moving ECM. Although this is reasonable behaviour for the cell density, this description of the evolving zero stress deformation gradient is highly unintuitive. However, we will find that the advection of the zero stress deformation gradient is a general result that also applies (in a modified form) to the three-dimensional case.

It should be noted that we can also use equation (5.3) to construct an evolution equation for \( Z \). Using the definition of \( Z \), we find that

\[
\frac{D}{Dt}(Z \Delta X) = a Z \Delta X,
\]

which implies that

\[
\frac{DZ}{Dt} \Delta X = a Z \Delta X;
\]

and hence,

\[
\frac{\partial Z}{\partial t} = a Z. \tag{5.8}
\]
Note that $\frac{DZ}{Dt} = \frac{\partial Z}{\partial t}$ because $Z$ is defined in Lagrangian coordinates.

Equations (5.6), (5.7) and (5.8) are only appropriate when a tissue is packed with cells that are all proliferating at the same rate. We now wish to develop general equations that can be used when the rate of growth varies spatially and temporally, possibly in response to chemical and/or mechanical signals.

To obtain equation (5.3), we specified that $\Delta z$ is proportional to the number of cells between $P$ and $Q$ and we assumed that the number of cells grows exponentially. Now, let us consider the case where $\Delta z$ changes at a rate that is an arbitrary function of our model variables.\(^7\) This would yield an equation of the form

$$\frac{D\Delta z}{Dt} = g^*(x, \Delta x, t), \quad (5.9)$$

where $g^*$ is a specified function that represents the rate of change of $\Delta z$. We note that $g^*$ has units of velocity; hence, $g^*$ can be physically interpreted as the rate at which $P$ and $Q$ are pushed apart by the growth between them. If the interval $[x, x + \Delta x]$ is universally stress-free, $g^*$ represents the relative velocity of $Q$ compared to $P$.

Importantly, it should be noted that we define $g^*$ using an Eulerian coordinate system. This definition is physically motivated; we expect the rate of growth to be dependent on the present cell density and other physical variables that are related to the current configuration of the body. Thus, it is more appropriate to express $g^*$ in Eulerian coordinates than Lagrangian coordinates.

Given equation (5.9), we now wish to develop a continuum analogue. Following equation (5.4), in which we expressed the cell population, $N$, in terms of the cell density, $n$, we wish define a ‘density’ equivalent of $g^*(x, \Delta x, t)$. Hence, let $g(x, t)$ be defined so that

$$g^*(x, \Delta x, t) \sim g(x, t) \Delta x$$

\(^7\)For example, the rate of change of $\Delta z$ could be dependent on the number of cells between $P$ and $Q$, the physical stresses between $P$ and $Q$, the oxygen availability etc.
as $\Delta x$ tends to zero.$^8$

The function $g(x, t)$ has units of inverse time and it effectively represents the instantaneous rate of growth (or shrinkage if $g$ is negative) at the point $x$. That is, $g$ can be interpreted as the rate of change of the desired size of an infinitesimal element relative to the current size of that same element.

By permitting $g(x, t)$ to vary through time and space, we are able to deal with the fact that the different regions of a tissue will grow (or shrink) at different rates. This is particularly useful when modelling the growth of plant roots, where most of the growth occurs in a small region near the root tip, or when modelling the growth of a tumour spheroid that contains proliferating, quiescent and necrotic regions.

Given the definition of $g(x, t)$ above, we find that equation (5.9) yields the result that

$$\frac{D}{Dt}(Y \Delta x) = g(x, t) \Delta x,$$

for small values of $\Delta x$. This means that

$$\left(\frac{\partial Y}{\partial t} + \frac{\partial}{\partial x}(Y v)\right) \Delta x = g(x, t) \Delta x,$$

and hence,

$$\frac{\partial Y}{\partial t} + \frac{\partial}{\partial x}(Y v) = g(x, t). \quad (5.10)$$

This is a particularly interesting result. Effectively we have shown that $g(x, t)$ is analogous to the source term of a mass balance equation. As we will see, growth often appears as a contribution to the source term of an evolution equation.

$^8$Note that $g^*(x, \Delta x, t)$ must tend to zero as $\Delta x$ tends to zero in order for the rate of growth to be physically realistic. This is because $g^*$ represents the relative velocity due to growth of two particles separated by an interval of $\Delta x$. Thus, $g^*(x, 0, t)$ must be zero as it is effectively the relative velocity of a particle compared to itself.
In Lagrangian coordinates, equation (5.9) becomes

\[
\frac{D}{Dt}(Z \Delta X) = g(x(X, t), t) \Delta x,
\]

or equivalently,

\[
\frac{\partial Z}{\partial t} \Delta X = g(x(X, t), t) F \Delta X.
\]

Thus,

\[
\frac{\partial Z}{\partial t} = g(x(X, t), t) F. \quad (5.11)
\]

Note that the deformation gradient must be included to take into account the fact that \( g \) is defined with reference to Eulerian coordinates rather than Lagrangian coordinates. If the growth were stated as a rate of change of the desired length per unit of initial length, the inclusion of \( F \) would be unnecessary. However, we generally expect the rate of growth to be stated in terms of the current configuration, not the initial configuration.

Equations (5.10) and (5.11) are respectively the Eulerian and Lagrangian equations that describe the changing zero stress state of a one-dimensional body that grows at a rate of \( g(x, t) \). Given this framework, it is possible to show that equations (5.7) and (5.8) represent the special case where \( g(x, t) \) is directly proportional to \( n(x, t) \) and the evolution of \( n \) is governed by

\[
\frac{\partial n}{\partial t} + \frac{\partial}{\partial x}(n v) = a n.
\]

That is, the cells responsible for growth reproduce at a rate of \( a \) and are advected with the underlying tissue.

This equation for \( n \) will only be appropriate in cases where there is little random motion of cells and the cell density is always well below its carrying capacity. Thus, equations (5.7) and (5.8) become increasingly inaccurate if the growing
tissue being modelled is physically constrained. In a situation like this, we would expect the increase in cell density (and hence the physical growth) to slow down over time.

Before we move on to consider three-dimensional growth, let us consider the special case of a stress-free growing body. Interestingly, it can be shown that residual stresses (i.e. stresses that can only be relieved by tearing, not by deformation) cannot arise in a one-dimensional Cartesian coordinate system. Hence, stress-free growth will arise in any case where we impose stress-free boundary conditions and there are no body forces.

As described earlier, a state of zero stress corresponds to the case where $\Delta z \equiv \Delta x$. That is,

$$Y = \lim_{\Delta x \to 0} \frac{\Delta z}{\Delta x} \frac{\partial x}{\partial x} = 1,$$

and

$$Z = \lim_{\Delta X \to 0} \frac{\Delta z}{\Delta X} \frac{\partial x}{\partial X} = F.$$

Substituting $Y = 1$ into (5.10) leads to the result that

$$\frac{\partial v}{\partial x} = g(x, t); \quad (5.12)$$

similarly, substituting $Z = F$ into (5.11) yields

$$\frac{\partial F}{\partial t} = g(x(X, t), t) F. \quad (5.13)$$

Equations (5.12) and (5.13) are instinctively meaningful descriptions of unconstrained growth. For example, we recall from classical mechanics that the divergence of velocity can be related to the rate of matter production. Since

---

9. This is a result of the fact that all one-dimensional Cartesian vector fields are conservative. For a given $Y$ or $Z$, this means that it is always possible to integrate to obtain a universally-valid zero stress coordinate system.

10. Specifically, we note that the divergence of velocity is equal to the volumetric rate of matter production in the case where the density is constant through time and space.
\textbf{5.3 Three-dimensional morphoelasticity}

\textbf{5.3.1 Volumetric growth}

Let us now consider how equations (5.10) and (5.11) might be extended to three-dimensional problems. Three-dimensional growth is considerably more complicated than one-dimensional growth for two reasons. Firstly, we need to consider the residual stresses that may result when the desired deformation resulting from growth is incompatible with a real deformation. Secondly, and more significantly, we need to deal with the fact that growth in some directions may be preferred over others (e.g. a tissue may grow in a certain direction in response to a chemical stimulus, or to minimise mechanical stresses).

As a first step towards a general description of growth, let us consider the growth of a small volume element. Thus, let \( P \) be a particle in body \( B \) and let \( X \) and \( x \) represent the initial and current positions of \( P \) respectively. Now, consider an elementary tetrahedron \( T_0 \) defined in the initial configuration so that the vertices of \( T_0 \) have position vectors \( X, X + \Delta X_1, X + \Delta X_2 \) and \( X + \Delta X_3 \). If \( \Delta X_i \) are all sufficiently small, the tetrahedron \( T_0 \) in the initial configuration will correspond to a new tetrahedron \( T \) in the current configuration with vertices \( x, x + \Delta x_1, x + \Delta x_2 \) and \( x + \Delta x_3 \), where \( \Delta x_i = F \Delta X_i \).

Now, let \( \Delta V_0 \) represent the volume of \( T_0 \) and let \( \Delta V \) represent the volume of \( T \).
It is a well-known result from continuum mechanics (see, for example, Spencer [176]) that, for \( \Delta V_0 \) and \( \Delta V \) both small,

\[
\Delta V \sim \det (F) \Delta V_0. \tag{5.14}
\]

That is, the determinant of the deformation gradient tensor can be thought of as the ratio between infinitesimal volumes in the current and initial configurations.\(^{11}\)

By analogy with equation (5.14), we introduce the concept of a desired volume, \( \Delta V_Z \). This is defined so that

\[
\Delta V_Z = \det (Z) \Delta V_0 = \det (Y) \Delta V, \tag{5.15}
\]

for infinitesimal volumes \( \Delta V_0 \), \( \Delta V \) and \( \Delta V_Z \).\(^{12}\) The desired volume is analogous to the desired length described in Section 5.2; if the particle \( P \) is under zero stress, it follows that \( \Delta V = \Delta V_Z \). However, the converse is not true; it is not sufficient to show that \( \det (F) = \det (Z) \) or \( \det (Y) = 1 \) in order to have zero stress at the point \( P \).

Now, we define \( g(\mathbf{x}, t) \) to represent the instantaneous rate of volumetric growth in the neighbourhood of \( \mathbf{x} \), analogous to \( g(x, t) \) in the one-dimensional case. Just as \( g(x, t) \) had units of reciprocal time (effectively, length per length per unit time), \( g(\mathbf{x}, t) \) also has units of reciprocal time (volume per volume per unit time) and can be thought of as the rate at which the desired volume grows per unit volume in the current configuration. Thus, the three-dimensional analogue of equation (5.9) is given by

\[
\frac{D \Delta V_Z}{Dt} = g(\mathbf{x}, t) \Delta V. \tag{5.16}
\]

---

\(^{11}\) Throughout this section, we consider only the evolution of infinitesimal volume elements. Hence, equation (5.14) and all related equations can be treated as being equalities rather than approximations.

\(^{12}\) A more formal derivation of equation (5.15) can be obtained by constructing a local zero stress coordinate system in the neighbourhood of \( P \). In this case, we take \( \chi = 0 \) at \( P \) and we find the vertices of the tetrahedron \( T_Z \) in the zero stress coordinate system using either \( Y \) or \( Z \). Following the method described in Spencer [176], we ultimately obtain the result stated.
We can use equation (5.16) to construct an evolution equation for $Y$ in Eulerian coordinates (equivalent to (5.10)) or for $Z$ in Lagrangian coordinates (equivalent to (5.11)). Firstly, let us consider the Eulerian problem. Substituting (5.15) into (5.16) yields the following:

$$\frac{D}{Dt}\left(\det (Y) \Delta V\right) = g(x, t) \Delta V,$$

or,

$$\frac{D}{Dt}\left(\det (Y)\right) \Delta V + \det (Y) \frac{D\Delta V}{Dt} = g(x, t) \Delta V. \quad (5.17)$$

Using equation (5.14) we therefore find that

$$\frac{D\Delta V}{Dt} = \frac{D}{Dt}\left(\det (F) \Delta V_0\right)$$

$$= \frac{D}{Dt}\left(\det (F)\right) \Delta V_0$$

$$= \frac{D}{Dt}\left(\det (F)\right) \det (F^{-1}) \Delta V. \quad (5.18)$$

In order to simplify equation (5.18), we use the well-known result from matrix calculus (see, for example, Schott [165]), which states that

$$\frac{d}{dt}\left(\det (A)\right) = \text{tr}\left(\frac{dA}{dt} A^{-1}\right) \det (A). \quad (5.19)$$

Also, we note that if we introduce velocity, $v$, defined so that $v = \frac{Dx}{Dt}$, we find that

$$\frac{DF}{Dt} F^{-1} = \frac{D}{Dt}\left(\frac{\partial x}{\partial X}\right) \frac{\partial X}{\partial x}$$

$$= \frac{\partial v}{\partial x} \frac{\partial X}{\partial x}$$

$$= \frac{\partial v}{\partial x}. \quad (5.20)$$
This is a well-known result from continuum mechanics (see, for example, Spencer [176]). The tensor $\frac{\partial v}{\partial x} = \frac{DF}{Dt} F^{-1}$ is called the velocity gradient tensor and it is often written as $L$.

Using these results, equation (5.18) becomes

$$\frac{D}{Dt} \left( \det (F) \right) \det (F^{-1}) = \text{tr} \left( \frac{DF}{Dt} F^{-1} \right) \det (F) \det (F^{-1})$$

$$= \text{tr} (L)$$

$$= \text{div} \, v,$$  \hspace{1cm} (5.21)

since

$$\text{tr} (L) = \delta_{ij} \frac{\partial v_i}{\partial x_j} = \frac{\partial v_i}{\partial x_i} = \text{div} \, v.$$

When we substitute equation (5.21) back into equation (5.18), we find that

$$\frac{D\Delta V}{Dt} = \text{div} \, v \Delta V.$$  \hspace{1cm} (5.22)

This is in accordance with the commonly-used fact that $\text{div} \, v = 0$ represents mass conservation in a medium of constant density. Effectively, equation (5.22) states that $\text{div} \, v$ is representative of the physically observed rate of volume growth per unit of current volume.

Now, we can substitute equation (5.22) back into (5.17) in order to find that

$$\left( \frac{D}{Dt} \left( \det (Y) \right) + \det (Y) \text{ div } v \right) \Delta V = g(\mathbf{x}, t) \Delta V,$$

and thus,

$$\frac{\partial}{\partial t} \left( \det (Y) \right) + \text{div} \left( v \det (Y) \right) = g(\mathbf{x}, t).$$  \hspace{1cm} (5.23)

This is analogous to equation (5.10), where we found that the one-dimensional...
zero stress deformation gradient, $Y$, is advected with a velocity of $v$ and grows at a rate of $g(x, t)$. Similarly, equation (5.23) states that the determinant of the three-dimensional deformation gradient is advected with a velocity of $v$ and grows at a rate given by $g(x, t)$. Thus, $g(x, t)$ is equivalent to the source term in a mass balance equation.

Equation (5.23) describes the evolution of the determinant of $Y$. However, this is not sufficient to specify how $Y$ changes over time; $\det(Y)$ is a scalar while $Y$ is a tensor with six independent components, so we would need a further five equations in order to completely determine the evolution of $Y$. Furthermore, although $g(x, t)$ is indicative of the total rate of growth, there is no way to use $g(x, t)$ to describe the fact that growth may be favoured in some directions over others. In Sections 5.3.2 and 5.3.3, we address these two concerns by constructing a tensor equation that is consistent with (5.23) and can be interpreted in terms of the total rate and direction of growth.

The Eulerian description of growth given in (5.23) is significantly more useful than its Lagrangian equivalent. However, for the sake of completeness, let us now develop the Lagrangian form. Substituting $\Delta V_Z = \det(Z) \Delta V_0$ and $\Delta V = \det(F) \Delta V_0$ into equation (5.16), we find that

$$\frac{D}{Dt} \left( \det(Z) \Delta V_0 \right) = g(x(X, t), t) \det(F) \Delta V_0.$$

Now, since $\det(Z)$ is defined in the Lagrangian configuration, material differentiation is equivalent to partial differentiation. Furthermore, $\Delta V_0$ is constant to material differentiation, so we find that

$$\frac{D}{Dt} \left( \det(Z) \right) \Delta V_0 = g(x(X, t), t) \det(F) \Delta V_0$$

$$\frac{\partial}{\partial t} \left( \det(Z) \right) = g(x(X, t), t) \det(F). \quad (5.24)$$

We recall that $Y$ appears in the definition of strain in the form $Y^T Y$. Thus, we can use the polar decomposition theorem to replace a nonsymmetric $Y$ with an equivalent symmetric tensor (i.e. a tensor with six independent components).
This is analogous to equation (5.11). It is interesting to note that the algebraic development of (5.24) is much simpler than the development of the Eulerian equation. However, equation (5.24), like equation (5.11), is awkward to use in practice because the growth rate, $g$, is specified with respect to the current configuration whereas all of the other variables are defined with respect to the initial configuration.

Before we move on to develop a more complete description of three-dimensional growth, it is again informative to consider what happens in the case of stress-free growth. That is, consider the situation where $\det(Y) = 1$ is a valid solution to equation (5.23) throughout the entire body.\(^{14}\) Substituting into equations (5.23) and (5.24), this yields

$$\text{div} \, \mathbf{v} = g(\mathbf{x}, t),$$

in Eulerian coordinates, or

$$\frac{\partial}{\partial t} \left( \det(F) \right) = g(\mathbf{x}(\mathbf{X}, t), t) \det(F),$$

in Lagrangian coordinates.

These two equations are equivalent to the one-dimensional equations (5.12) and (5.13). Importantly, we note that we retain the result that, in the absence of internal stresses, the growth rate is equal to the divergence of velocity. This is consistent with the fact that $\text{div} \, \mathbf{v}$ can be interpreted as being proportional to the rate of matter production in a body of constant density.

\(^{14}\)Note that this may or may not be possible, depending on the constitutive law given for the body, the body forces applied, the variation of the rate of growth through space etc.
5.3.2 The pre-symmetrised growth tensor

In this section, we extend equation (5.23) to construct a possible representation of how the Eulerian zero stress deformation gradient tensor evolves as a tissue grows, taking into account the fact that growth in some directions may be preferred over others. As we will see, the formulation developed in this section needs further refinement in order to be useful, but it is instructive to consider the pre-symmetrised growth tensor discussed here before moving on to the true growth tensor.

Throughout this section, we will exclusively use Eulerian coordinates; thus, it is convenient for us to use the symbols $E$ and $e$ to represent the finite Eulerian strain and infinitesimal Eulerian strain respectively (previously $E^E$ and $e^E$).

To begin with, we note that equation (5.23) can also be written in the form

$$\frac{D}{Dt} \left( \det (Y) \right) + \det (Y) \text{ div } v = g(x, t). \quad (5.25)$$

Now, we use the result given in equation (5.19) to find an alternative expression for $\frac{D}{Dt} (\det(Y))$ as follows:

$$\frac{D}{Dt} (\det (Y)) = \text{tr} \left( \frac{DY}{Dt} Y^{-1} \right) \det (Y)$$

$$= \text{tr} \left( Y^{-1} \frac{DY}{Dt} \right) \det (Y), \quad (5.26)$$

since $\text{tr}(AB) = \text{tr}(BA)$.

Recalling that $\text{div } v = \text{tr}(L)$, we substitute (5.26) into (5.25) to find that

$$\text{tr} \left( \det (Y) Y^{-1} \frac{DY}{Dt} + \det (Y) L \right) = g(x, t). \quad (5.27)$$

Thus, the total rate of volumetric growth at a particular point may be expressed
as the trace of the tensor $\Gamma$ defined by

$$\Gamma = \det(Y) Y^{-1} \frac{DY}{Dt} + \det(Y) L. \quad (5.28)$$

For reasons that are discussed later, we call the tensor $\Gamma$ defined above the pre-symmetrised growth tensor and we contend that it is, in certain circumstances, a physically meaningful measure of growth. Importantly, $\Gamma$ not only describes the total rate of growth but, as we will show, it can also be used to define preferred directions of growth.

Before we do this, it should be pointed out that we have made a large conceptual leap from saying that the trace of $\Gamma$ represents the rate of volumetric growth to saying that $\Gamma$ itself is physically meaningful. There is a large family of tensors with the same trace as the pre-symmetrised growth tensor and it is possible that another tensor in this family would be a more useful measure of growth. However, several properties of $\Gamma$ suggest that it is a useful measure of the rate of growth.

In order to investigate the pre-symmetrised growth tensor, let $\Gamma_Z$ and $\Gamma_v$ be defined by

$$\Gamma_Z = \det(Y) Y^{-1} \frac{DY}{Dt} \quad \text{and} \quad \Gamma_v = \det(Y) L,$$

so that $\Gamma = \Gamma_Z + \Gamma_v$. Now, let $d\mathbf{x}$ be an infinitesimal vector in the current configuration with magnitude $dx$ and direction $\hat{x}$. Similarly, let $d\Omega$ be an infinitesimal surface element in the current configuration with area $dA$ and unit normal $\hat{n}$, and define the infinitesimal vector $d\mathbf{n} = dA \hat{n}$. Given these definitions, we can construct the bilinear forms

$$B_Z(d\mathbf{n}, d\mathbf{x}) = d\mathbf{n}^T \Gamma_Z d\mathbf{x} \quad \text{and} \quad B_v(d\mathbf{n}, d\mathbf{x}) = d\mathbf{n}^T \Gamma_v d\mathbf{x}.$$ 

Let us firstly consider $B_Z(d\mathbf{n}, d\mathbf{x})$. We recall from elementary matrix algebra that the inverse of a matrix is equal to the adjoint divided by the determinant.
That is,

\[ A^{-1} = \frac{1}{\det(A)} \text{adj}(A) = \frac{1}{\det(A)} (\text{cof}(A))^T, \]

where \( \text{cof}(A) \) represents the matrix constructed from the cofactors of \( A \) (i.e. the determinants of the minors of \( A \)). Thus, we can use the definition of \( \Gamma_Z \) to find that \( B_Z(dn, dx) \) may be rewritten as

\[
B_Z(dn, dx) = dn^T \det(Y) Y^{-1} \frac{DY}{Dt} dx = dn^T (\text{cof}(Y))^T \frac{DY}{Dt} dx = (\text{cof}(Y) dn) \cdot \left( \frac{DY}{Dt} dx \right). \tag{5.29}
\]

Now, it can be shown\(^{15}\) that if \( c = a \times b \), then

\[
\text{cof}(A) c = (A a) \times (A b). \tag{5.30}
\]

That is, if \( c \) is normal to a planar region \( R \) in one coordinate system, then \( \text{cof}(A) c \) is normal to the planar area which is the image of \( R \) under \( A \). Furthermore, if the magnitude of \( c \) corresponds to the area of \( R \), we find that the magnitude of \( \text{cof}(A) c \) corresponds to the area of the image of \( R \).

Now, let \( d\Omega' \) be an area element in a locally constructed\(^{16}\) zero stress coordinate system such that \( d\Omega' \) is the image of \( d\Omega \) under \( Y \). Using the result above, we find that \( \text{cof}(Y) dn \) is normal to \( d\Omega' \) and \( || \text{cof}(Y) dn || \) corresponds to the area

\(^{15}\)It is possible to prove this result by considering a general \( a, b \) and \( A \) and then expanding both sides of equation (5.30) algebraically. It is also possible to prove a more general case of (5.30) by considering the definition of a determinant as an exterior product and using results from differential geometry (see, for example, Darling [42]).

\(^{16}\)We recall that a zero stress coordinate system cannot generally be constructed if residual stresses exist. However, for a given point \( P \), it is always possible to construct a zero stress coordinate system which is consistent in the neighbourhood of \( P \). Thus, we can use the concept of a zero stress coordinate system as long as we restrict ourselves to infinitesimal vectors, as we do here.
of $d\Omega'$. This gives us a physical interpretation for the expression on the left hand side of the dot product in equation (5.29).

Furthermore, $\frac{D \mathbf{Y}}{Dt} \, d\mathbf{x}$ can be thought of as being like a velocity in the local zero stress coordinate system. We recall that if $d\mathbf{x}$ is the displacement between two particles in the current configuration, then $\mathbf{Y} \, d\mathbf{x}$ is the displacement between the same two particles in some locally-constructed zero stress coordinate system. Thus, $\frac{D \mathbf{Y}}{Dt} \, d\mathbf{x}$ defines the relative ‘velocity’ in the zero stress coordinate system of two neighbouring particles in the current configuration.

Hence, equation (5.29) means that $B_Z(d\mathbf{n}, d\mathbf{x})$ can be interpreted as the dot product of a velocity in the zero stress coordinate system with a normal vector to a surface in the zero stress coordinate system: $B_Z(d\mathbf{n}, d\mathbf{x})$ is, in some sense, a flux associated with growth.

Now, consider $B_v$. We note that $B_v(d\mathbf{n}, d\mathbf{x})$ can be expressed in the form

$$B_v(d\mathbf{n}, d\mathbf{x}) = \det(\mathbf{Y}) \, d\mathbf{n}^T \mathbf{L} \, d\mathbf{x} = \det(\mathbf{Y}) \cdot (d\mathbf{n} \cdot (\mathbf{L} \, d\mathbf{x})).$$

Thus, $B_v$ is proportional to the dot product of a velocity in the Eulerian coordinate system with a normal vector to a surface in the Eulerian coordinate system. It is also a flux, but this time representative of flux in real space rather than in zero stress space.

Having demonstrated that this is the case, it is still not clear why $B_Z$ and $B_v$ can be added to yield a meaningful result. Moreover, we have not fully justified the flux interpretation of $B_Z$ and $B_v$. In order to understand this better, consider two planar area elements in Eulerian space, $dA_1$ and $dA_2$, which are defined so that iff $\mathbf{x}$ is the position vector of a point in $dA_1$ then $\mathbf{x} + d\mathbf{x}$ is the position vector of a point in $dA_2$. That is, the two areas are parallel and have the same area, but are separated by a vector of $d\mathbf{x}$. It is possible for them to contain overlapping regions. Since the areas are parallel with the same area, it is possible to define a common normal $d\mathbf{n}$ such that $d\mathbf{n}$ is perpendicular to the two planar elements.
and has a length representative of their common area.

We now consider the equivalents of $dA_1$ and $dA_2$ in some locally constructed zero coordinate system; $dA_1$ and $dA_2$ will correspond to new planar area elements $dA_1^Z$ and $dA_2^Z$. These will have the property that there exists a vector $dn_Z$ such that $dn_Z$ is normal to both of them, and both new planar elements have an area of $||dn_Z||$. Furthermore, it is possible to construct a vector $dx_Z$ representing the distance between the two planes, such that $x_Z + dx_Z$ is the position vector of a point in $dA_2^Z$ iff $x_Z$ is the position vector of a point in $dA_1^Z$. Following the previous work, we recognise that $dn_Z = \text{cof}(Y) dn$ and $dx_Z = Y dx$.

![Figure 5.3: A diagram showing a general prism in the current state and the zero stress state. The vector $dn$ is perpendicular to the faces of the prism while the vector $dx$ is parallel with the length of the prism. As shown, the zero stress state may be rotated and deformed relative to the current state.](image)

Now, consider the prism $C$ which has $dA_1^Z$ and $dA_2^Z$ as opposite faces. Given $dn_Z$ and $dx_Z$ defined as above, we find that the volume of $C$ is given by

$$V_C = (dn_Z) \cdot (dx_Z) = dn^T \text{cof}(Y)^T Y dx. \tag{5.32}$$

Thus, for a fixed material prism defined in Eulerian coordinates, it is possible to describe how the volume of the prism evolves over time as the zero stress state changes. However, since $C$ is simply an infinitesimal volume element, we note that any change in the $V_C$ can be specified completely by considering the material
rate of change of $\det Y$. We can see this clearly since

$$V_C = d n^T \ \text{cof}(Y)^T X d x$$

$$= d n^T \ \det(Y) Y^{-1} Y d x$$

$$= \det(Y) d n^T d x.$$ 

Despite this, equation (5.32) is still very useful. In particular, we note that the material rate of change of $V_C$ takes the form

$$\frac{D V_C}{D t} = \frac{D}{D t} (d n_Z) \cdot (d x_Z) + (d n_Z) \cdot \frac{D}{D t} (d x_Z),$$

and hence it is possible to separate the change in the volume of $V_C$ into two components, one representative of the changes to the size and orientation of $d A^Z_1$ and $d A^Z_2$ and the other representative of the changes to the distance between them.

Importantly, consider what happens if we neglect the changes to $d n_Z$ and consider only the changes to $d x_Z$. That is, we are keeping the surfaces $d A^Z_1$ and $d A^Z_2$ fixed and we are considering only the flux of material over them that occurs in response to growth. This is illustrated for a simple case in Figure 5.4 below.

Expressed algebraically, we are effectively considering what happens if we only include the second part of equation (5.33). That is,

$$\begin{bmatrix} \frac{D V_C}{D t} \\ d n^Z \text{ held fixed} \end{bmatrix} = (d n_Z) \cdot \frac{D}{D t} (d x_Z)$$

$$= d n^T \ \text{cof}(Y)^T \frac{D}{D t} (Y d x)$$

$$= \det(Y) d n^T Y^{-1} \left( \frac{D Y}{D t} d x + Y L d x \right)$$

$$= \det(Y) \left( d n^T Y^{-1} \frac{D Y}{D t} d x + d n^T L d x \right)$$
Figure 5.4: Consider the case where $dn_Z$ and $dx_Z$ are both horizontal, as represented by the two arrows in the diagram above. Now, let the continuous rectangle above represent the prism $C$ at time $t$ and let the dashed rectangle represent $C$ at time $t + dt$. As we see, the there has been both stretching in the horizontal direction and stretching in the vertical direction. That is, $dA^2_1$ and $dA^2_2$ have grown in area and moved apart; the total amount of growth because of this can be represented by the difference in areas between the two rectangles. If we restrict ourselves to considering only the change in volume as a result of the relative movement of the two end planes (i.e. we treat the area and orientation of $dA^2_1$ and $dA^2_2$ as fixed), we are left with only the growth depicted in blue. Importantly, the volume of this blue region does not depend only on the sizes of $dn_Z$ and $dx_Z$ (as the difference between the total volumes does), but also on their orientations.

$= dn^T \Gamma dx$. (5.34)

Thus, we see that for any given values of $dn$ and $dx$, it is possible to obtain the growth volume corresponding to the blue regions in Figure 5.4 using the pre-symmetrised growth tensor. This gives us a justification for adding the two fluxes represented by $B_Z$ and $B_v$ as well as giving us a physical interpretation of the bilinear form associated with $\Gamma$. Effectively, $dn^T \Gamma dx$ is the material rate of change of the desired volume of the prism $C$ defined earlier, except that we do not include changes to desired volume associated with the stretching or rotation of the two principal faces.

If we consider a right prism (i.e. the case where $dn$ is parallel to $dx$), we find that $dn^T \Gamma dx$ is effectively a quadratic form associated with growth along the length of the prism. That is, for any given direction $\hat{x}$, we expect that $\hat{x}^T \Gamma \hat{x}$ can be interpreted as a representation of the rate of growth in the direction $\hat{x}$. It
follows from this that if $\Gamma$ is symmetric, the eigenvectors of $\Gamma$ are the principal directions of growth, and the eigenvalues of $\Gamma$ are representative of the rates of growth in these principal directions.

5.3.3 A unique representation of the growth tensor

We have found that the pre-symmetrised growth tensor can be physically interpreted as a measure of growth. However, there are still some remaining problems. Firstly, we find that $\Gamma$ changes significantly if the rotation in the zero stress state deformation gradient is changed. To see that this is the case, let $Y(x, t)$ be a zero stress deformation gradient tensor. It follows from the arguments in Section 4.3.2 that $\hat{Y}(x, t) = Q(x, t)Y(x, t)$ will also be a zero stress deformation gradient tensor where $Q$ is proper orthogonal for all values of $x$ and $t$.

Thus, if $\Gamma = \det(Y)Y^{-1}\frac{DY}{Dt} + \det(Y)L$ is the pre-symmetrised growth tensor associated with a zero stress deformation gradient of $Y$, we can also construct an alternative growth tensor, $\hat{\Gamma}$, using $\hat{Y}$. In this case,

$$\hat{\Gamma} = \det(\hat{Y})\hat{Y}^{-1}\frac{D\hat{Y}}{Dt} + \det(\hat{Y})L$$

$$= \det(Y)Y^{-1}Q^{T}\left(Q\frac{DY}{Dt} + \frac{DQ}{Dt}Y\right) + \det(Y)L$$

$$= \det(Y)Y^{-1}\frac{DY}{Dt} + \det(Y)Y^{-1}Q^{T}\frac{DQ}{Dt}Y + \det(Y)L$$

$$= \Gamma + \det(Y)Y^{-1}SY, \quad (5.35)$$

where $S = Q^{T}\frac{DQ}{Dt}$ is a tensor function of space and time.

Given that $Q(x, t)$ is orthogonal for all values of $x$ and $t$, it can easily be shown
that $S$ is always skew-symmetric. Conversely, if $S$ is skew-symmetric for all values of $x$ and $t$ and $Q(x, 0)$ is proper orthogonal, it is possible to show that

$$\frac{DQ}{Dt} = QS$$

implies that $Q(x, t)$ is always proper orthogonal. Thus, we find that the set of possible choices for $S$ is identical to the set of skew-symmetric tensors. For a given choice of $Y$, any tensor of the form

$$\Gamma = \det(Y) Y^{-1} \frac{DY}{Dt} + \det(Y) L + \det(Y) Y^{-1} SY,$$

(5.36)

is a valid growth tensor as long as $S$ is skew-symmetric.

Out of each family of tensors defined by equation (5.36), we would like to choose a single tensor that best represents the growth of our tissue. We recall from Section 4.4 that it is convenient for us to define the zero stress state using the principal zero stress deformation gradient. In an Eulerian coordinate system, this means that we choose the rotation in $Y$ so that $Y$ is symmetric.

\footnote{This can be seen from the fact that

$$0 = \frac{DI}{Dt} = \frac{D}{Dt} (Q^T Q) = Q^T \frac{DQ}{Dt} + \frac{DQ^T}{Dt} Q = S + S^T.$$}

Thus, $S^T = -S$ and $S$ is skew-symmetric.

\footnote{To see that this is true, consider the case where $Q$ is a function of $t$ alone (following a material point) and $Q(t^*)$ is known to be proper orthogonal. We wish to show that $S$ being skew-symmetric implies that $Q(t^* + \Delta t)$ is proper orthogonal, where $\Delta t$ is very small. By discretising the differential equation above, we find that small $\Delta t$ implies that

$$Q(t^* + \Delta t) = Q(t^*) (I + S(t^*) \Delta t),$$

Now, consider $(Q(t^* + \Delta t))^T Q(t^* + \Delta t)$:

$$(Q(t^* + \Delta t))^T Q(t^* + \Delta t) = (I - S \Delta t) (Q(t^*))^T Q(t^*) (I + S \Delta t)$$

$$= (I - S \Delta t) (I + S \Delta t)$$

$$= I + O(\Delta t^2).$$

It follows from this that $Q(t^* + \Delta t)$ is orthogonal. Furthermore, the fact that the determinant of $Q$ is a continuous function of $t$ implies that $Q(t^* + \Delta t)$ will also be proper orthogonal.}
Given that $Y$ is symmetric, we now wish to determine the best choice of $S$ that will lead to a unique and interpretable choice of $\Gamma$. At a first glance, it would seem appropriate to impose $S = 0$ and thus reduce the complexity of the definition of $\Gamma$. However, by considering the case of classical elasticity (i.e. zero growth), we will see that there is actually a better choice of $S$ that leads to a growth tensor that matches our physical intuition.

In classical elasticity, the material coordinate system effectively acts as a zero stress coordinate system. That is, $F^{-1} = \frac{\partial X}{\partial x}$ can be thought of as a tensor which maps infinitesimal vectors in the current configuration to corresponding infinitesimal vectors in a zero stress coordinate system. Thus, $F^{-1}$ satisfies the definition of a valid zero stress deformation gradient and could potentially be substituted for $Y$ in equation (5.36).

We have, however, stated that we would prefer $Y$ to be symmetric and it is not generally true that $F^{-1}$ will be symmetric. We therefore use the polar decomposition theorem to define $Y$ so that $Y = QF^{-1}$ is symmetric and $Q$ is a proper orthogonal tensor. It follows that

$$Y^{-1} \frac{DY}{Dt} = F \frac{DF^{-1}}{Dt} + FKF^{-1}$$

where $K = Q \frac{DQ^T}{Dt}$ is skew-symmetric. The tensor $K$ effectively represents the evolving rotation of the local zero stress coordinate system that is necessary in order for $Y$ to be symmetric.

Now, let us substitute $Y = QF^{-1}$ into equation (5.36) and observe the form that $\Gamma$ takes. Firstly, we find that

$$\Gamma_{\text{zero growth}} = \det(F^{-1}) \left( F \frac{DF^{-1}}{Dt} + L + FKF^{-1} + FQ^T S Q F^{-1} \right).$$

19Note that $F, Q$ and $Y$ may vary over time.
We recall from matrix calculus (see, for example, Schott [165]) that
\[
\frac{dA^{-1}}{dt} = -A^{-1} \frac{dA}{dt} A^{-1},
\]
where \( A \) is an invertible matrix. Also, we note that
\[
(Q^T S Q)^T = Q^T S^T Q = -Q^T S Q.
\]
Thus, if \( S \) is a general, undetermined skew-symmetric tensor, so is \( S^*_s = Q^T S Q \).

Using these results, we find that
\[
\Gamma_{\text{zero growth}} = \det(F^{-1}) \left( -\frac{DF}{Dt} F^{-1} + L + F K F^{-1} + F S^*_s F^{-1} \right).
\]

Furthermore, we recall from the derivation of equation (5.21) that \( \frac{DF}{Dt} F^{-1} = L \). Thus,
\[
\Gamma_{\text{zero growth}} = \det(F^{-1}) \left( -L + L + F (K + S^*_s) F^{-1} \right)
= \det(F^{-1}) F (K + S^*_s) F^{-1}.
\]

Since the pre-symmetrised growth tensor is representative of growth and we are considering the case of zero growth, we would prefer to choose \( \Gamma \) so that it is equal to the zero tensor. Hence, the most appropriate choice of \( S \) is a solution to \( K + Q^T S Q = 0 \). As we will see later (as a special case of a more general result), there is a unique skew-symmetric tensor \( S \) that will fulfill this requirement.

From equation (5.38) we see clearly that it is not generally appropriate to pick \( S = 0 \). Furthermore, although equation (5.38) allows us to determine the correct choice for \( S \) in the case of zero growth, we are left with the question of what the most appropriate value of \( S \) is in the case of general growth.

Given the physical interpretation of the pre-symmetrised growth tensor from Section 5.3.2, there are certain advantages to choosing \( S \) so that \( \Gamma \) is symmetric.
For example, consider the value of $\hat{n} \cdot \Gamma \hat{x}$ when $\hat{n}$ and $\hat{x}$ are both unit vectors; this may be considered to be representative of the flux associated with growth in the $\hat{x}$ direction as measured across a plane with normal $\hat{n}$.

If $\Gamma$ is symmetric, we know that the eigenvalues of $\Gamma$ are all real and the eigenspaces of $\Gamma$ are orthogonal. Thus, we can write $\hat{n}$ and $\hat{x}$ as $\hat{n} = n_i g_i$ and $\hat{x} = x_i g_i$, where $g_i$ are orthogonal eigenvectors\(^{20}\) of $\Gamma$ and $n_i n_i = x_i x_i = 1$. If $\lambda_i$ represents the eigenvalue of $\Gamma$ associated with $g_i$, we find that

$$\hat{n} \cdot \Gamma \hat{x} = \lambda_i n_i x_i.$$ 

Thus, if $\Gamma$ has at least one positive (negative) eigenvalue, it follows that the maximum (minimum) value of $\hat{n} \cdot \Gamma \hat{x}$ occurs when $\hat{n} = \hat{x}$ and $\hat{x}$ is an eigenvector of $\Gamma$ associated with the largest (most negative) eigenvalue of $\Gamma$.

That is, in the case of positive growth there always exists a direction such that the flux associated with growth in that direction across a plane normal to that direction is maximal. Furthermore, as an eigenvector, this principal direction of growth is clearly and closely related to the symmetric choice of the growth tensor.

More generally, if we consider the quadratic form $\hat{x} \cdot \Gamma \hat{x}$, we clearly see that the orthogonal eigenvectors of $\Gamma$ can be associated with the directions of fastest, intermediate and slowest growth, with the rate of growth in the direction of an eigenvector indicated by the associated eigenvalue of $\Gamma$. This is analogous to the way in which the eigenvectors of the symmetric $Y$ correspond to the principal axes of strain and the magnitudes of these strains can be calculated from the eigenvalues of $Y$. Thus, if growth occurs preferentially in the direction of least stress/strain, we expect that $\Gamma$ will have the same eigenvectors as $Y$, but that the largest eigenvalues of $\Gamma$ correspond to the smallest eigenvalues of $Y$.

This strongly suggests that it is preferable to choose $\Gamma$ to be symmetric. A symmetric growth tensor can be clearly interpreted in terms of its eigenvalues and

\(^{20}\)Even if $\Gamma$ has repeated eigenvalues, it is possible to construct orthogonal eigenvectors using the Gram-Schmidt process.
eigenvectors and this leads to the possibility of constructing phenomenological expressions for the rate of growth in terms of other model variables. Moreover, we will see that choosing $\Gamma$ to be symmetric ensures that the growth tensor is zero in the case of zero growth.

An even better reason for choosing $\Gamma$ to be symmetric can be found in the theories of plasticity that draw on the multiplicative decomposition of the deformation gradient. As described by Lubarda [111, 112] and Dafalias [39], an important concept in elastoplasticity theory is the idea of plastic spin. Plastic spin is effectively a measure of the rate of rotation of a given choice of zero stress state relative to the rate of rotation of the current state. Given that the zero stress state is only unique up to arbitrary rotations, it is always possible to choose the changing zero stress state so that the plastic spin is zero. This choice of zero stress state is called the ‘isoclinic’ configuration (i.e. the configuration with the same inclination as the current state) [39, 111].

The isoclinic configuration is particularly useful when specifying an evolution equation for the zero stress state or an associated variable. This is because the isoclinic configuration is defined so that it negates the effect of arbitrary plastic spin for variables defined with respect to the zero stress configuration. In biological problems, we generally expect that the fibres that define the mechanical structure of the tissue are embedded in the zero stress state. Hence, the evolution of fibre behaviour (and subsequently the evolution of the zero stress state) is best represented by considering an isoclinic configuration. A more detailed description of this argument can be found in Dafalias [39].

It should be emphasised that isoclinicity refers to the dynamics of the changing zero stress state, not to the choice of zero stress configuration in a static problem. Thus, we can remove the need for directly constructing an isoclinic configuration.

---

[21] Dafalias [39] develops a strong argument that the isoclinic configuration is only useful if we can make the constitutive assumption that the underlying crystalline/fibrous structure of the body is embedded in the zero stress state. For other constitutive assumptions (for example, the case of a crystal lattice that is not directly associated with the zero stress state), a variety of different ‘constitutive spins’ can be used to obtain a meaningful type of observer independence for developing evolution equations.
by developing time derivatives that are equivalent to what would be observed in an isoclinic configuration. For an arbitrary zero stress configuration, this can easily be achieved by using a corotational time derivative based on the plastic spin.

That is, we can introduce isoclinicity by replacing the material time derivative \( \frac{DY}{Dt} \) with a corotational derivative of the form

\[
\frac{D}{Dt}Y = \frac{DY}{Dt} - \Omega_p Y,
\]

where \( \Omega_p \) is the plastic spin.\(^{22}\) As given by Lubarda [112], the plastic spin is the unique skew-symmetric tensor that satisfies the equation

\[
\text{skw} \left( \frac{DY}{Dt} - Y - 1 Y \Omega_p Y \right) = \text{skw} \left( L \right).
\]

Using identity (5.37), this can be rearranged to yield

\[
\text{skw} \left( Y^{-1} \frac{DY}{Dt} + L - Y^{-1} \Omega_p Y \right) = 0. \tag{5.39}
\]

Now, replacing \( \frac{DY}{Dt} \) with \( \frac{D}{Dt}Y \) in the definition of \( \Gamma \) given in (5.28) yields

\[
\Gamma = \det \left( Y \right) Y^{-1} \frac{DY}{Dt} + \det \left( Y \right) L - \det \left( Y \right) Y^{-1} \Omega_p Y.
\]

We note that the fact that \( \Omega_p \) is skew-symmetric means that this expression is in the same form as equation (5.36). Hence, the \( \Gamma \) given above satisfies the requirements for being a growth tensor. Moreover, equation (5.39) implies that this choice of growth tensor is symmetric. Thus, using a corotational time-derivative based on an isoclinic configuration is equivalent to choosing the skew-symmetric tensor \( S \) in equation (5.36) that ensures that \( \Gamma \) is symmetric. This gives us further confirmation that this is an appropriate way to define \( S \).

\(^{22}\)Note that the corotational derivative given here takes the same form as a vector corotational derivative. This is because \( Y \) is a two-point tensor and it is only necessary to rotate the coordinate that lies in the zero stress configuration.
Having established the value of making $\Gamma$ symmetric, we will now show that for any point in time and space there will exist a unique skew-symmetric tensor, $S$, which makes the growth tensor symmetric. Equation (5.36) states that

$$\Gamma = \det(\mathbf{Y}) \mathbf{Y}^{-1} \frac{D\mathbf{Y}}{Dt} + \det(\mathbf{Y}) \mathbf{L} + \det(\mathbf{Y}) \mathbf{Y}^{-1} \mathbf{S} \mathbf{Y}.$$  

Since $\Gamma$ is symmetric, $\mathbf{Y}$ is symmetric positive definite and $\mathbf{S}$ is skew-symmetric, we can take the transpose of this equation to find that

$$\Gamma = \det(\mathbf{Y}) \frac{D\mathbf{Y}}{Dt} \mathbf{Y}^{-1} + \det(\mathbf{Y}) \mathbf{L}^T - \det(\mathbf{Y}) \mathbf{S} \mathbf{Y}^{-1}.$$  

If we premultiply the first of these equations by $\mathbf{Y}$ and postmultiply the second of these equations by $\mathbf{Y}$, we find that

$$\frac{1}{\det\mathbf{Y}} (\mathbf{Y} \Gamma) = \frac{D\mathbf{Y}}{Dt} + \mathbf{Y} \mathbf{L} + \mathbf{S} \mathbf{Y}$$

$$\frac{1}{\det\mathbf{Y}} (\Gamma \mathbf{Y}) = \frac{D\mathbf{Y}}{Dt} + \mathbf{L}^T \mathbf{Y} - \mathbf{Y} \mathbf{S}.$$  

Taking the difference of these equations and rearranging yields:

$$\mathbf{S} \mathbf{Y} + \mathbf{Y} \mathbf{S} = \mathbf{L}^T \mathbf{Y} - \mathbf{Y} \mathbf{L} + \frac{1}{\det\mathbf{Y}} (\mathbf{Y} \Gamma - \Gamma \mathbf{Y}). \tag{5.40}$$

Given that $\mathbf{S}$ is our unknown, this is in the form of the Sylvester equation (see, for example, Antoulas [7]). Importantly, we note that since $\mathbf{Y}$ is symmetric positive definite, it is impossible for $\mathbf{Y}$ to have any eigenvalues in common with $-\mathbf{Y}$. Thus, it follows from the theory of Sylvester equations that a unique solution to (5.40) will exist.

Now, although we used the fact that $\mathbf{S}$ is skew-symmetric to construct equation (5.40), it would be wise to ensure that the unique solution obtained when solving (5.40) is skew-symmetric. Otherwise, it may be possible that no appropriate solution exists and we have actually derived a contradiction.

---

23Interestingly, Hoger and Carlson [93] also use the Sylvester equation in their derivation of an expression for the time derivative of the square root of a tensor.
To see that this \( S \) is skew-symmetric, we note that equation (5.40) is in the form

\[
X A + A X = W,  \tag{5.41}
\]

where \( X \) is our unknown, \( A \) is symmetric positive definite and \( W \) is skew-symmetric. Taking the transpose of this general form, we note that

\[
A X^T + X^T A = W^T = -W.
\]

This can then be added to equation (5.41) to find that

\[
(X + X^T) A + A (X + X^T) = 0.
\]

Since this is again in the appropriate form of a Sylvester equation and \( A \) shares no eigenvalues with \( -A \), we find that there is a unique solution for \( X + X^T \).

Thus, since \( X + X^T = 0 \) is clearly a solution to the equation, it must be the only solution. Hence, we can conclude that \( X \) is necessarily skew-symmetric.

From this work, it follows that for a given \( Y \) and \( L \), there exists one and only one skew-symmetric tensor \( S \) such that

\[
\det (Y) Y^{-1} \frac{DY}{Dt} + \det (Y) L + \det (Y) Y^{-1} SY
\]

is symmetric. That is, there is a unique symmetric tensor of the form

\[
G = \det (Y) \left( Y^{-1} \frac{DY}{Dt} + L + Y^{-1} SY \right),  \tag{5.42}
\]

where \( S \) is skew-symmetric.

We define the symmetric tensor \( G \) in equation (5.42) to be the principal pre-

---

24 As an aside, we recall that in the case of zero growth we were looking for a skew-symmetric tensor \( S \) which would allow us to recover \( \Gamma = 0 \). As the zero tensor is symmetric, the result above tells us that there does indeed exist a unique \( S \) which leads to \( \Gamma = 0 \) in this case. Furthermore, it also establishes the fact that there are no symmetric tensors other than the zero tensor which could conceivably be obtained from the definition of \( \Gamma \) in the case of zero growth.
symmetrised growth tensor (or, more simply, the growth tensor). In a model of tissue growth, we would generally specify $G$ and then solve the resultant equations for $\frac{DY}{Dt}$ and $L$.

Unfortunately, equation (5.42) is not simple to use because of the difficulty involved in solving (5.40) to obtain $S$. This is compounded by the fact that $L$ is generally not known when we first attempt to find $S$; hence, a complicated scheme would be needed to find the $\frac{DY}{Dt}$, $L$ and $S$ that satisfy all of the relevant equations. Therefore, we wish to eliminate $S$ from equation (5.42).

If we premultiply equation (5.42) by $Y$ and postmultiply by $Y^{-1}$ we find that

$$YGY^{-1} = \det(Y) \left( \frac{DY}{Dt} Y^{-1} + YLY^{-1} + S \right).$$

Adding equation (5.43) to its own transpose, it follows that

$$YGY^{-1} + Y^{-1}GY = \det(Y) \left( Y^{-1} \frac{DY}{Dt} Y^{-1} + YLY^{-1} + Y^{-1}LY \right),$$

or

$$\det(Y) \text{ sym} \left( \frac{DY}{Dt} Y^{-1} + YLY^{-1} \right) = \text{sym} \left( YGY^{-1} \right),$$

where $\text{sym}(A) = \frac{1}{2} (A + A^T)$ represents the symmetric part of $A$.

Equation (5.44) is more useful than equation (5.42) because we have circumvented the need to find $S$. In a computational problem, using (5.42) instead of (5.44) would make it necessary to solve a Sylvester equation at every time step. Although we expect that there will still be computational challenges involved with implementing equation (5.44), these are not as great as the challenges that would be encountered with (5.42).

It is interesting to note that equations (5.42) and (5.44) will be valid regardless of whether or not we use the principal zero stress deformation gradient.\(^{25}\)

\(^{25}\)A formal proof of this result can be found at the beginning of Appendix D. Effectively, replacing $Y$ with $QY$ changes the choice of $S$ needed to obtain a symmetric growth tensor.
However, the value of \( Y \) that we obtain from solving the model equations will only be uniquely defined if \( Y \) is the principal zero stress deformation gradient. Unless stated otherwise, it is implied that \( Y \) refers to the principal zero stress deformation gradient whenever we use equation (5.42) or equation (5.44).

Lastly, we note that it is possible to simplify equation (5.44) in the case where the principal directions of growth are parallel to the principal directions of strain. This should arise regularly in physical situations; for example, we would often expect the growth of a tissue to be directed so that the greatest growth occurs in the direction of least strain and the least growth occurs in the direction of greatest strain. When the principal directions of growth and strain match, we find that \( Y \) and \( G \) have the same eigenvectors and hence commute. Thus, it follows that

\[
\det(Y) \text{ sym}\left( \frac{DY}{Dt} Y^{-1} + Y L Y^{-1} \right) = G
\]

is an alternative to equation (5.44) in this case.

### 5.3.4 Incorporating morphoelasticity into a model

Equation (5.44) describes how the Eulerian zero stress deformation gradient changes in response to an imposed direction-dependent rate of growth, represented by \( G \). However, this equation will only make sense in the context of a larger mechanical model. We note that equation (5.44) contains \( Y \) and \( L \); in a practical problem, these will both be unknowns, but (5.44) on its does not contain enough information to enable us to find both of them. Thus, we will now take a step backwards and investigate how (5.44) might be incorporated into a larger mathematical model.

To begin with, let us consider the number of dependent variables that we have in a typical problem and thus consider the number of equations that we need. Generally, we would specify the growth tensor (possibly as a function of other

while \( G \) itself is unaffected.
parameters) and seek solutions for the zero stress deformation gradient, \( Y \); the strain, \( E \); the stress, \( \sigma \); the velocity gradient, \( L \); and the velocity, \( v \), at every point in the body.\(^{26}\)

Now, consider a full three-dimensional problem. If we use the principal zero stress deformation gradient, we find that \( Y \) is symmetric and thus, \( Y \) will have six independent components. Similarly, the elastic strain tensor and the Cauchy stress tensor must also be symmetric; hence, they each have six independent components. In contrast, the velocity gradient, \( L \), has no special structure and will have nine independent components. Lastly, velocity is a vector, so it can be treated as having three independent components. In total, we find that \( Y \), \( E \), \( \sigma \), \( L \) and \( v \) have thirty independent components; it follows that we require the equivalent of thirty independent scalar equations in order for a mechanical problem to be well-posed.

One of the equations that we intend to use is equation (5.44). The terms on both sides of this equation are symmetric tensors; thus, (5.44) is equivalent to six scalar equations. Secondly, we will use results obtained in the previous chapter to define strain in terms of the zero stress deformation gradient. If each of the six independent components of strain is a function of the components of \( Y \), this gives us a further six scalar equations.

Similarly, we need a constitutive law relating the components of stress, \( \sigma \), to the components of strain, \( E \); this will again be equivalent to six scalar equations. We also need a force balance equation (see, for example, Spencer [176]) that relates the divergence of stress to body forces and to acceleration. As force is a vector, the force balance equation will be equivalent to three scalar equations. Lastly, we note that the components of \( L \) can be defined in terms of the components of \( v \), since \( L_{ij} = \frac{\partial v_i}{\partial x_j} \). This will give us the thirty equations that we require.

\(^{26}\)In Section 4.3.3, it was pointed out that it is possible to use the Eulerian zero stress deformation gradient to avoid any need for reference to the initial configuration. One of the consequences of this is that it is not necessary to solve for displacement when considering a fully Eulerian mechanical problem with an evolving zero stress state. Theoretically, it should be possible to use the velocity solution to recover displacements, but this would probably be difficult.
It is interesting to compare this with a typical problem in classical elasticity, in which there would be twenty-four unknowns (the six components of strain, the six components of stress, the nine components of the displacement gradient and the three components of displacement) and fifteen equations (six equations for the definition of strain in terms of displacement, six constitutive equations relating strain and stress, three force balance equations and nine equations for the definition of the displacement gradient). This is very similar to our formulation, but with a number of major differences.

Most obviously, our equations use the velocity and the velocity gradient, while classical elasticity is expressed in terms of displacement and the displacement gradient. We are able to do this for two reasons. Firstly, we define the strain in terms of the zero stress deformation gradient, not the displacement gradient. Secondly, we define the zero stress deformation gradient using an evolution equation. As this relates the changing zero stress deformation gradient to the velocity gradient, we find that there is no need to include displacement in our system of equations.

The second major difference between classical elasticity and our model is that we have six additional unknowns and we thus require six more equations. This is because it is no longer possible to define the strain in terms of the displacement gradient. Instead, we define strain in terms of the zero stress deformation gradient, which has no analogue in classical elasticity theory. The evolution of the zero stress deformation gradient gives us the six new equations that we need for a well-defined system.

Now, we can state the complete systems of equations that we would use to model the growth of an elastic body. As described above, this will consist of a zero stress state evolution equation (i.e. equation (5.44)), a strain definition, a stress-strain relationship, a force balance equation and the velocity gradient definition.

If finite deformations away from the zero stress state occur, it is necessary to use the full strain definition from equation (4.33). Noting that $Y$ is symmetric, this
becomes

\[ E = \frac{1}{2} (I - Y^2). \]

However, if \( Y \) is close to the identity, it is appropriate to use a small strain approximation. Since we are using the principal zero stress deformation gradient, we can use equation (4.35):

\[ e = I - Y. \]

There are a variety of different forms that the stress-strain definition could take. In the case of finite deformations, it would probably be necessary to construct a strain energy function, \( W(E) \) or some equivalent, and use this to define the stress tensor.\(^{27}\) Following the example of Goriely and Ben Amar [18, 70], we find that the Cauchy stress tensor for an incompressible material will take the form

\[ \sigma = Y^{-1} \frac{\partial W}{\partial Y^{-1}} - q I, \]

where \( q \) is the hydrostatic pressure and \( W \), the strain energy, is defined as a function of \( Y^{-1} \). The analysis of such systems rapidly becomes very difficult and is beyond the scope of the present work.

Instead, we will restrict our analysis to the case of linear elasticity. More specifically, we find that if the material being considered is isotropic and linearly elastic, we can use the isotropic Hookes law:

\[ \sigma = \lambda \text{tr} (e) I + 2 \mu e, \]

where \( \lambda \) and \( \mu \) are the Lamé coefficients.

Next, consider the force balance equation. If there are no body forces and inertial

\(^{27}\)An account of the use of strain energy in classical finite elasticity can be found in Spencer [176]. There are, however, a number of subtleties involved in the definition of strain energy in the case where the zero stress state is permitted to evolve. For a detailed treatment of strain energy in this case, see especially the seminal paper by Lubarda and Hoger [113]. Other relevant papers include the work of Hoger [92] and the recent work of Goriely and Ben Amar [70].
effects are ignored, we find that this takes the form
\[ \text{div } \sigma = 0, \]
or, in Einstein notation,
\[ \frac{\partial \sigma_{ij}}{\partial x_i} = 0. \]
Lastly, \( L \) is defined so that
\[ L = \frac{\partial v}{\partial x}, \]
giving us the final nine equations that we require.

In the case where the strains are small and the isotropic Hooke’s law is appropriate, we thus obtain the following system:

\[ \det (Y) \text{ sym} \left( \frac{DY}{Dt} Y^{-1} + Y L Y^{-1} \right) = \text{sym} \left( Y G Y^{-1} \right), \quad (5.46) \]
\[ e = I - Y, \quad (5.47) \]
\[ \sigma = \lambda \text{ tr} (e) I + 2 \mu e, \quad (5.48) \]
\[ \text{div } \sigma = 0, \quad (5.49) \]
\[ L = \frac{\partial v}{\partial x}. \quad (5.50) \]

Equations (5.46) to (5.50) can be simplified further when some additional symmetry of the system is known. For example, in the case of spherical symmetry, we find that \( v \) and \( L \) each have only one nonzero component (relating to the radial velocity), while \( Y, e \) and \( \sigma \) each have two independent components (the radial strain/stress and the circumferential strain/stress).

In fact, it is even possible to make some simplifications when no special symmetry is known. In Section 5.4.2, we use the assumption of small strains to combine equations (5.46) and (5.47), yielding a single equation for the evolution of the strain tensor, \( e \).
Equations (5.46) to (5.49) illustrate how the equations developed in this chapter could be incorporated into a model of tissue growth. For given boundary conditions and a specified growth tensor (which may vary through time and space), these equations can be solved to yield predictions of the motion and stresses associated with growth.

Similarly, we could use the equations to construct more complicated systems. For example, new variables like cell density and oxygen concentration could be introduced into the model. Then, stress could be a function of cell density (as with many dermal wound healing models) and the growth tensor could even depend on stress, strain, cell density and other model variables. We discuss some of these extensions and their applications in Chapter 6.

5.4 Evolution equations for strain

5.4.1 One-dimensional strain evolution

If we are using the principal zero-stress deformation gradient we find that the finite and infinitesimal Eulerian strain tensors are defined as follows:

\[ E = \frac{1}{2} (I - Y^2) , \]

and

\[ e = I - Y . \]

Now, we can rearrange these equations to obtain expressions for \( Y \) in terms of \( E \) or \( e \):\(^{28}\)

\[ Y = \sqrt{I - 2 E} , \]  \hspace{1cm} (5.51)

\(^{28}\)Note that the square root operation in equation (5.51) will be uniquely defined as we require that \( Y \) is positive definite.
or,

\[ Y = I - e. \]  

(5.52)

In this section, we substitute equations (5.51) and (5.52) into our zero stress state evolution equations in order to obtain equations which describe how the Eulerian strain changes in response to growth. As in the previous section, the full three-dimensional case is significantly more complicated than the one-dimensional case; for example, great care needs to be taken with the tensorial square root in equation (5.51). Thus, we will firstly consider how strain evolves in a one-dimensional Cartesian formulation.

In the one-dimensional case, the tensors in equations (5.51) and (5.52) can be replaced by scalars. That is,

\[ Y = \sqrt{1 - 2E}, \]  

(5.53)

and,

\[ Y = 1 - e. \]  

(5.54)

From equation (5.10), we recall that the evolution of \( Y \) is given by

\[ \frac{\partial Y}{\partial t} + \frac{\partial}{\partial x}(vY) = g(x, t), \]

where \( g \) represents the rate of growth and \( v \) represents velocity.

Now, if we substitute the expression for \( Y \) from equation (5.54) into equation (5.10), we obtain an evolution equation for infinitesimal strain as follows:

\[ -\frac{\partial e}{\partial t} + \frac{\partial}{\partial x}(v - ve) = g(x, t), \]

or, equivalently,

\[ \frac{\partial e}{\partial t} + \frac{\partial}{\partial x}(ve) = \frac{\partial v}{\partial x} - g(x, t). \]  

(5.55)
Thus, Eulerian infinitesimal strain is effectively advected with the moving tissue and is ‘created’ at a rate determined by the imbalance between the velocity gradient and the rate of growth.

Despite the fact that equation (5.55) has an unusual form, we find that it matches our intuitive expectations for the physics of the system. For example, we find that growth in the absence of velocity makes the strain more negative (i.e. a growing solid that is constrained will feel more compressed over time). Similarly, we note that \( e = \frac{\partial u}{\partial x} \) is consistent with \( g = 0 \) (i.e. we can recover classical elasticity in the case of zero growth).

Having obtained the infinitesimal strain formulation, let us now use equation (5.51) to construct a similar evolution equation for finite strain. Substituting into (5.10), we find that

\[
\frac{\partial}{\partial t} \left( \sqrt{1 - 2E} \right) + \frac{\partial}{\partial x} \left( v \sqrt{1 - 2E} \right) = g(x, t).
\]

Thus,

\[
\frac{-1}{\sqrt{1 - 2E}} \frac{\partial E}{\partial t} + \frac{\partial v}{\partial x} \sqrt{1 - 2E} - \frac{v}{\sqrt{1 - 2E}} \frac{\partial E}{\partial x} = g(x, t),
\]

and we can show that

\[
\frac{\partial E}{\partial t} + \frac{\partial v}{\partial x} (1 - 2E) - v \frac{\partial E}{\partial x} = g(x, t) \sqrt{1 - 2E};
\]

ultimately yielding

\[
\frac{\partial E}{\partial t} + \frac{\partial}{\partial x} (v E) = (1 - E) \frac{\partial v}{\partial x} - g(x, t) \sqrt{1 - 2E},
\]

or, equivalently,

\[
\frac{DE}{Dt} + 2E \frac{\partial v}{\partial x} = \frac{\partial v}{\partial x} - g(x, t) \sqrt{1 - 2E}.
\]
Equation (5.56) is similar to equation (5.55) but there are some interesting differences. In particular, we note that the rate of growth in equation (5.56) is multiplied by $\sqrt{1 - 2E}$. This is difficult to interpret physically; in effect, the growth term is modified in order to take into account the conversion factor between the current volume and the desired volume.

Equation (5.56) is most likely to be useful when the residual strains are large. However, in many such cases it is desirable to have an explicit representation of the zero stress deformation gradient that can be used in the definition of strain energy.\(^{29}\) In our work we will generally focus on applications where the strains are always small and equation (5.55) can be used. This avoids the problem of developing a constitutive law that is valid at large strains.

However, we do consider one example of finite strain incorporated into a morphoelastic model. In Section 6.2, we develop a morphoelastic model of a one-dimensional Maxwell material that uses equation (5.56) in conjunction with a linear constitutive law relating stress and strain. As we will see, this leads to some interesting results that could not be obtained using infinitesimal strain.

Lastly, it should be noted that there are other ways of describing the evolution of the zero stress state that we have not considered. For example, it is relatively simple to construct an equation for the evolution of the Lagrangian strain using equation (5.11). As the Lagrangian strain is of limited usefulness to us, we will not consider this in further detail.

### 5.4.2 Three-dimensional strain evolution

As given in equations (5.51) and (5.52), it is possible to define the principal Eulerian zero stress deformation gradient tensor, $Y$, in terms of the finite and

\(^{29}\)See, for example, the strain energy function developed by Ben Amar and Goriely [18].
infinitesimal strain tensors as follows:

\[ Y = \sqrt{I - 2E}, \]

and,

\[ Y = I - e. \]

We now wish to substitute these expressions into the three-dimensional growth equation, (5.44).

We note that equation (5.44) involves the material time derivative of \( Y \). Thus, using equation (5.51) to construct an evolution equation for \( E \) would involve finding

\[ \frac{D}{Dt} \left( \sqrt{I - 2E} \right), \]

the derivative of the square root of a tensor. The problem of finding such a derivative was dealt with by Hoger and Carlson [93] as they considered the problem of expressing the material derivative of the stretch tensors in terms of the evolution of the Cauchy-Green tensors. Unfortunately, Hoger and Carlson’s results are very algebraically dense and they cannot be used to construct a straightforward equation for the evolution of finite strain.

Despite this difficulty, it is possible to use an alternative method to construct an evolution equation for finite strain. The derivation of the finite strain evolution equation is described in Appendix D. Ultimately, we find that

\[ \frac{D\mathbf{E}}{Dt} = \text{sym} \left( (I - 2\mathbf{E}) \mathbf{L} - \frac{1}{\sqrt{\det(I - 2\mathbf{E})}} (I - 2\mathbf{E}) \mathbf{G} \right). \]  

(5.57)

Equation (5.57) is an interesting result that deserves further analysis. As we describe in Appendix D, this equation is observer-independent (i.e. it is unaffected by a rotation of the coordinate axes). Furthermore, it is easy to show that
equation (5.57) is consistent with the classical definition of strain when $G = 0$.

Importantly, we note that equation (5.57) could easily be incorporated into a model of biological growth. In fact, it is reasonable to expect that (5.57) will be less computationally demanding to solve than equation (5.44).\(^{30}\) At this stage, we have not constructed any models that incorporate three-dimensional evolving finite strain, nor have we attempted to solve (5.57) numerically. Further work is needed in order to obtain practical results.

In this section, we will focus on combining equations (5.44) and (5.52) to develop an evolution equation for the infinitesimal strain, $e$.\(^{31}\) Although the strain evolution equation that we develop is not linear, it is considerably simpler than equations (5.44) and (5.57).

Now, we recall that equation (5.44) is given by

$$\det (Y) \ \text{sym} \left( \frac{DY}{Dt} \ Y^{-1} + Y \ L \ Y^{-1} \right) = \text{sym} \left( Y \ G \ Y^{-1} \right).$$

If we make the substitution $Y = I - e$, but do not do any further simplification, we will obtain a complicated evolution equation for $e$ that is not particularly useful. However, the infinitesimal strain tensor, $e$, must be small in order for $Y = I - e$ to be valid. Thus, we can use asymptotic methods to construct a simple equation for the strain evolution.\(^{32}\)

\(^{30}\)We note that equation (5.57) appears to lend itself easily to the construction of an algorithm that uses explicit (Euler) time-stepping. Although such an approach may suffer from numeric instabilities and there are problems with ensuring that the force balance equation is satisfied (see later notes), the construction of an explicit algorithm presents us with a simple option for validating our model numerically. In contrast, it is always necessary to solve a nonlinear system in order to step forwards in time using equation (5.44).

\(^{31}\)Alternatively, it is possible to obtain the evolution equation for infinitesimal strain by simplifying equation (5.57). Both approaches yield the same result but using the zero stress state evolution equation is somewhat simpler.

\(^{32}\)As in Section 4.4, we do not formally define the metric that we use to measure the ‘smallness’ of a tensor for asymptotic purposes. However, we require that the tensor norm we use be submultiplicative. For a submultiplicative norm, $||e^n|| \leq ||e||^n$ and it follows that we can use conventional order notation. That is, we write that $A \sim O(e)$ if $||A|| \sim O(||e||)$ as $e \to 0$. Given this definition and given that the norm is submultiplicative, it follows that $A \sim O(e)$ implies that $A^n \sim O(e^n)$. One appropriate tensor norm with this property is the spectral norm.
For example, we note that (5.44) involves $Y^{-1}$ and $\det(Y)$. Now, it is well known that the series

$$(I - A)^{-1} = I + A + A^2 + \ldots$$

converges when the eigenvalues of $A$ all have a magnitude less than unity (see, for example, Tylavsky and Sohie [191]). Thus,

$$Y^{-1} = (I - e)^{-1} = I + e + O(e^2),$$

since $e$ is assumed to be infinitesimal.

Similarly, by expanding the determinant of $I - e$ and neglecting quadratic and higher order terms, it can easily be seen that

$$\det(Y) = \det(I - e) = 1 - \text{tr}(e) + O(||e||^2).$$

Now, $Y$ is not the only dependent variable in equation (5.44); we must also take into account the velocity gradient. Although $Y = I - e$ by definition, $L$ must be represented with an asymptotic series. That is, let

$$L \sim L_0 + L_1 + \ldots,$$

where $L_i = O(e^i)$.

Taking $G$ to be a constant of order $I$ and substituting these results into equation (5.44), we find that

$$\left(1 - \text{tr}(e)\right) \text{sym} \left(-\frac{De}{Dt}(I + e) + (I - e)(L_0 + L_1)(I + e)\right)$$

$$= \text{sym} \left((I - e)G(I + e)\right) + O(e^2).$$

(5.58)

of the matrix representation of the tensor.
Collecting the leading order terms, this yields

\[ \text{sym} \left( L_0 \right) = \text{sym} \left( G \right) = G. \quad (5.59) \]

Interestingly, we note that equation (5.59) does not involve strain. In Section 5.3.4 we demonstrated that a well-defined model of tissue growth needs five equations: a zero stress evolution equation, a strain definition, a stress-strain constitutive law, a force balance equation and a velocity gradient definition. At leading order, we find that we have replaced the zero stress evolution equation and the strain definition with a new equation that relates the velocity gradient to \( G \).

In order to understand what this means, consider what happens if all of the dependent variables in the full model (other than \( e \)) are expressed as asymptotic series. That is,

\[ v \sim v_0 + v_1 + \ldots, \]

and

\[ \sigma \sim \sigma_0 + \sigma_1 + \ldots, \]

where \( v_i = O(e^i) \), and similarly for \( \sigma_i \).

Now, the velocity gradient definition equation takes the form

\[ L = \frac{\partial v}{\partial x}. \]

Applying our asymptotic expansion, this means that

\[ L_0 = \frac{\partial v_0}{\partial x}. \quad (5.60) \]

Since equation (5.59) does not involve strain, we find that we can use equations
(5.59) and (5.60) to completely determine $L_0$ and $v_0$. Now, $L_0$ has nine independent components and $v_0$ has three independent components. However, we can see that (5.59) is equivalent to six equations and (5.60) has nine equations. Thus, we have an over-determined system and we may not be able to find a solution.

To see why this is the case, we note that the we have constructed our equations on the assumption of small strains. Thus, the leading order system corresponds to the case where the strain, and hence the stress, is zero. This means that solving for $L_0$ effectively involves finding a velocity gradient that allows the body to be stress free at every point. It is interesting to compare this with the static problem of finding a deformation gradient, $F$, that corresponds to a universal state of zero stress.

In Section 4.3.2, we discussed the fact that, for a given $Z$, it is not always possible to construct $F$ so that $\sigma = 0$ everywhere. This is because $F = \frac{\partial x}{\partial X}$, which means that each row of $F$ must define a conservative vector field. In contrast, there is no such restriction on the zero stress deformation gradient, $Z$. Since zero stress corresponds to the case where $F = QZ$, it follows that it is not generally possible to choose $F$ so that $\sigma = 0$ throughout the body being considered.

Similarly, we note that $L_0 = \frac{\partial v_0}{\partial x}$ and hence each row of $L_0$ must also define a conservative vector field. However, equation (5.59) states that the symmetric part of $L_0$ is equal to the growth tensor, $G$. Finding $L_0$ when given $G$ is equivalent to the problem of finding a deformation gradient corresponding to zero stress if $Z$ is specified. Although some choices of $G$ make it possible to find $L_0$, this is not generally the case because $\text{curl} \ L_0$ must be zero. Some growth rates are inconsistent with a stress-free body; these necessarily lead to the development of residual strains.

One way of addressing this problem would be to stipulate that $G$ must be compatible with a real velocity gradient in order for the asymptotic analysis to be valid. However, our work has been motivated by the desire to construct an approach to elasticity that is valid when growth leads to residual stress. Thus, we
would like our strain evolution equation to be able to deal with the case of small residual strains (i.e. the case where $G$ is close to being the symmetric part of a compatible velocity gradient). In order to do this, we propose that $G$ can be expressed as the sum of a compatible part and an incompatible part. That is,

$$G = G_c + G_i,$$

where $G_c$ and $G_i$ are both symmetric, $\text{curl curl } G_c = 0$, and $G_i$ is as small as possible. By $\text{curl curl } G$, we imply that both the row-curl and the column-curl are taken. As these two operators commute, the order of taking the two different curls is unimportant. The double curl is necessary because $G$ is associated with the symmetric part of $L$ rather than $L$ itself.\(^{33}\)

In order to continue our asymptotic analysis, we require that $G_i = O(e)$. In this case, we find that (5.58) becomes

$$\left(1 - \text{tr} \left( e \right) \right) \, \text{sym} \left(- \frac{D e}{D t} (I + e) + (I - e) \left( L_0 + L_1 \right) (I + e) \right) = \text{sym} \left( (I - e) \left( G_c + G_i \right) (I + e) \right) + O(e^2), \quad (5.61)$$

and our leading order balance becomes

$$\text{sym} \left( L_0 \right) = G_c. \quad (5.62)$$

Unlike the previous leading order balance, we can be sure that this has a solution corresponding to a valid velocity gradient.\(^{34}\) Thus, with appropriate boundary conditions, we can use (5.60) and (5.62) to obtain solutions for $L_0$ and $v_0$.

\(^{33}\)The feasibility of this decomposition can be assessed using the incompatibility tensor, $\eta = \text{curl curl } G$, described by Skalak et al. [175]. If $\eta$ is small, then $G$ is almost compatible and it should be possible to construct a decomposition where $G_i$ is small.

\(^{34}\)In fact, there will be a family of solutions for the velocity gradient. This is because finding $L_0$ given $G_c$ is equivalent to the problem of finding $F$ given a compatible $Z$. There will be infinitely many solutions for $L_0$ because introducing a spatial rotation changes $L_0$ without affecting $G_c$. It is necessary to use the boundary conditions in order to uniquely determine the appropriate choice of $L_0$. 

In order to define the strain evolution, we need to move on to consider the first order balance. Collecting terms of order $e$ in (5.61), we find that

$$\operatorname{sym} \left( -\frac{De}{Dt} - \operatorname{tr}(e) L_0 - eL_0 + L_0 e + L_1 \right) = \operatorname{sym} \left( -eG_e + G_e e + G_i \right).$$

Now, if $A$ and $B$ are both symmetric, then $BA - AB$ will be skew-symmetric. Similarly, if $A$ is symmetric and $B$ is skew-symmetric, then $BA - AB$ will be symmetric. Thus, we can exploit the fact that $e, G_e$ and $G_i$ are all symmetric to find that

$$\frac{De}{Dt} + \operatorname{tr}(e) \operatorname{sym}(L_0) + e \operatorname{skw}(L_0) - \operatorname{skw}(L_0) e = \operatorname{sym}(L_1) - G_i, \quad (5.63)$$

where $\operatorname{skw}(A)$ represents the skew-symmetric part of $A$, so that $\operatorname{skw}(A) = \frac{1}{2} (A - A^T)$.

Since we can evaluate $L_0$ by solving (5.62) with appropriate boundary conditions, we find that equation (5.63) is linear in $e$ and $L_1$. Also, since both sides of (5.63) are symmetric tensors, it will be equivalent to six scalar equations.

Now, consider the other equations that are needed for a well-defined system. From the first order term in the definition of the velocity gradient, we find that

$$L_1 = \frac{\partial v_1}{\partial x}.$$ 

Since $L_1$ has no special structure, this is equivalent to nine linear scalar equations.

Next, we can use the stress-strain relationship to express $\sigma$ as a function of $e$, yielding the equivalent of six scalar equations. The fact that we are considering

\footnotesize{35} As noted above, $\operatorname{curl} L_0$ must be zero. Hence, there will be a unique $L_0$ that satisfies equation (5.62) as well as the given boundary conditions. That is, it is possible to construct $L_0$ in its entirety from $\operatorname{sym}(L_0)$.

\footnotesize{36} Note that we used the assumption that $e$ is small to construct asymptotic expansions for $L, \sigma$ and $v$, but we did not construct an asymptotic expansion for $e$. If we are using linear elasticity, $\sigma$ will be a linear function of $e$ and it follows that there will only be one nonzero term in the asymptotic expansion of $\sigma$. Thus, we can say that $\sigma$ is a function of $e$, rather than considering $\sigma_0, \sigma_1, etc.$
small strains means that we expect this relationship to be linear. Lastly, we need a force balance equation. This will be equivalent to three scalar equations and, if there are no nonlinear body force terms, we find that these equations are also linear.

Thus, we find that we have the same number of equations as unknowns. Furthermore, all of our equations appear to be linear. If we use the isotropic Hooke’s law and neglect all body forces and inertial effects, our complete system takes the form

$$\text{sym}(L_0) = G_c,$$

$$L_0 = \frac{\partial v_0}{\partial x},$$

$$\frac{D e}{D t} + \text{tr}(e) \text{sym}(L_0) + e \text{skw}(L_0) - \text{skw}(L_0) e = \text{sym}(L_1) - G_i,$$

$$L_1 = \frac{\partial v_1}{\partial x},$$

$$\sigma = \lambda \text{tr}(e) I + 2\mu e,$$

$$\text{div} \sigma = 0.$$  

Although all of these equations are effectively linear, there is an additional problem that we have not discussed. Our analysis depends on expressing $G$ as the sum of a compatible part, $G_c$, and a small incompatible part, $G_i$. However, we have not described how this decomposition can be done in practice.

---

37 As before, we note that the these assumptions could be replaced with constitutive laws that are more appropriate for the material being modelled. However, in our analysis we will limit ourselves to using the isotropic Hooke’s law and no body forces.
One approach would be to consider the double curl of $G$ as follows:

$$\text{curl} \text{ curl} G = \text{curl} \text{ curl} \left( G_c + G_i \right)$$

$$= \text{curl} \text{ curl} G_i,$$

since $\text{curl} \text{ curl} G_c = 0$ for compatibility. Thus, the problem of finding $G_i$ is equivalent to the problem of inverting the double curl operator.

For a given $G$, solving

$$\text{curl} \text{ curl} G_i = \text{curl} \text{ curl} G$$

will yield a large family of solutions for $G_i$, even when we require that $G_i$ is symmetric. In order to determine which decomposition of $G$ is most appropriate, we note that the divergence of stress must be equal to the body force (in this case assumed to be zero). Not all strains lead to stresses that are consistent with this requirement and it is important to ensure that $G_i$ is constructed in such a way that an appropriate solution for $e$ can be found. Determining whether a given decomposition of $G$ yields consistent results would appear to be a difficult problem in the most general case and it is beyond the scope of this thesis.\textsuperscript{38}

One way of avoiding this difficulty is to recombine equations (5.62) and (5.63) as follows. Equation (5.62) is equivalent to

$$0 = \text{sym} \left( L_0 \right) - G_c.$$

Adding this to equation (5.62), we find that

$$\frac{De}{Dt} + \text{tr} \left( e \right) \text{ sym} \left( L_0 \right) + e \text{ skew} \left( L_0 \right) - \text{skew} \left( L_0 \right) e$$

$$= \text{sym} \left( L_0 \right) - G_c + \text{sym} \left( L_1 \right) - G_i,$$

\textsuperscript{38}We also seek to make $G_i$ as small as possible. It is unclear as to how this interacts with the problem of choosing $G_i$ so that it is ultimately consistent with the force balance equation.
or equivalently,
\[
\frac{D e}{D t} + \text{tr}(e) \text{ sym}(L_0) + e \text{ skew}(L_0) - \text{skw}(L_0) e = \text{sym}(L) - G + O(e^2).
\]

Now, since \( L = L_0 + O(e) \), we find that
\[
\text{tr}(e) \text{ sym}(L_0) = \text{tr}(e) \text{ sym}(L) + O(e^2),
\]
and similar results will apply to \( e \text{ skew}(L_0) \) and \( \text{skw}(L_0) e \). Thus, if we neglect terms of order \( e^2 \), we find that
\[
\frac{D e}{D t} + \text{tr}(e) \text{ sym}(L) + e \text{ skew}(L) - \text{skw}(L) e = \text{sym}(L) - G. \tag{5.64}
\]

Unlike equation (5.63), we note that this equation is not linear because we do not explicitly solve for \( L_0 \) before moving on to consider \( L_1 \).

Interestingly, we note that the terms on the left hand side of (5.64) are of order \( e \), while the terms on the right hand side of (5.64) are of order \( I \). Although this difference in order seems strange, we recall that \( \text{sym}(L) - G \) must be small in order for our analysis to be possible. Thus, there is no inherent contradiction in equation (5.64) as it stands, although we need to be careful to ensure that \( \text{sym}(L) - G \) does not become large without us noticing the change.

If using equation (5.64), our complete system of equations for modelling growth will take the form
\[
\frac{D e}{D t} + \text{tr}(e) \text{ sym}(L) + e \text{ skew}(L) - \text{skw}(L) e = \text{sym}(L) - G, \tag{5.65}
\]

\[
\sigma = \lambda \text{ tr}(e) I + 2 \mu e, \tag{5.66}
\]

\[
\text{div } \sigma = 0, \tag{5.67}
\]

\[
L = \frac{\partial v}{\partial x}. \tag{5.68}
\]
This is a significant improvement over the original system of equations described in Section 5.3.4. Most importantly, we note that the strain evolution equation does not include a term of the form $\frac{DY}{Dt} Y^{-1}$; in fact, we have circumvented the need to compute any inverses or determinants. Thus, we expect this model to be significantly less computationally intense than the original model.\(^{39}\)

However, equation (5.64) is an unusual equation and it deserves to be investigated further. In particular, we wish to ensure that we can use (5.64) to recover the one-dimensional strain evolution equation and also that (5.64) yields physically sensible results in the case of zero growth. Finally, we would like to ensure that (5.64) is observer-independent; we do not formally demonstrate this here, but the process is analogous to the analysis of finite strain evolution described in Appendix D.\(^{40}\)

\(^{39}\)Although it may be less computationally demanding than other morphoelastic models, the numeric implementation of the system given above still requires very careful treatment. For example, consider the case where $G$ is a specified function of space and time. Let us assume that $\varepsilon$, $\sigma$, $v$ and $L$ are known at $t = t_0$ and we wish to find $\varepsilon(x, t_0 + \Delta t)$, $\sigma(x, t_0 + \Delta t)$, etc. using an appropriate numeric scheme.

At first, it would appear possible to use equation (5.65) to obtain an approximate updated value for $\varepsilon$ using explicit (or implicit) time stepping. However, substituting this updated strain into equation (5.66) would yield an updated stress that may or may not satisfy equation (5.67). Furthermore, it is unclear as to how updated values of $L$ and $v$ can be obtained for the next timestep.

As in Section 3.4.2, we encounter difficulties because the force balance equation does not depend on velocity. Therefore, it may be necessary to differentiate equation (5.67) with respect to time (in some way) in order to obtain an evolution equation for the velocity gradient. Careful thought needs to be given to the problem of constructing a numeric framework that is appropriate for morphoelastic problems, but this is beyond the scope of this thesis.

\(^{40}\)It is interesting to note that equation (5.64) can be rewritten in the form

$$\frac{\partial \varepsilon}{\partial t} + \text{tr}(\varepsilon) \text{sym}(L) = \text{sym}(L) - G,$$

where $\frac{\partial}{\partial t}$ represents the Jaumann time derivative. The Jaumann time derivative of strain (or stress) is a particularly useful construct in continuum mechanics because it is inherently observer independent [44, 111]. However, it is not the only tensor time derivative with this property. Denn [44] gives a general expression for a large family of observer independent time derivatives, which includes the Jaumann derivative and the Oldroyd (upper convected) derivative as special cases. Interestingly, the entire left hand side of equation (5.64) is in an appropriate form to be expressed as one of Denn’s tensor time derivatives. It remains to be seen whether this tensor time derivative can be given a clear physical interpretation in general use.
Firstly, let us consider the simplification of (5.64) in the one-dimensional case. Since all one-dimensional tensors are symmetric, it follows that

$$\text{skw}(\mathbf{L}) = 0, \quad \text{and} \quad \text{sym}(\mathbf{L}) = \frac{\partial v}{\partial x};$$

where $v$ is the one-dimensional velocity and $x$ is the one-dimensional Eulerian spatial coordinate. Similarly, we find that $\mathbf{e} = \text{tr}(\mathbf{e}) = e$ are all equivalent representations of the one-dimensional strain, while $\mathbf{G} = g$ is equal to the rate of volumetric growth.

Thus, equation (5.64) becomes

$$\frac{D\mathbf{e}}{Dt} + e \frac{\partial v}{\partial x} = \frac{\partial v}{\partial x} - g,$$

or equivalently,

$$\frac{\partial \mathbf{e}}{\partial t} + \frac{\partial}{\partial x}(ve) = \frac{\partial v}{\partial x} - g,$$

which, as we expected, is identical to equation (5.55).

Now, consider what happens to the three-dimensional equation when there is no growth; that is, when $\mathbf{G} = \mathbf{0}$. In this case,

$$\frac{D\mathbf{e}}{Dt} + e \text{ skw}(\mathbf{L}) - \text{skw}(\mathbf{L}) \mathbf{e} = \text{sym}(\mathbf{L}). \quad (5.69)$$

Since there is no growth, we expect this equation to be valid for an elastic (as opposed to morphoelastic) material. Thus, we would like to determine whether classical infinitesimal strain, given by

$$\mathbf{e}_{\text{classical}} = \text{sym} \left( \frac{\partial \mathbf{u}}{\partial x} \right),$$

is consistent with equation (5.69).

However, we recall from Section 4.4.2 that our definition of infinitesimal strain is different from the classical definition, even in the case where the initial con-
configuration is always a zero stress state. Specifically, we note that our use of the principal zero stress deformation gradient allows us to construct a definition of strain which is valid for arbitrary physical rotations of the body being considered. In contrast, classical infinitesimal strain can only be used when the rotations are small.

Thus, we will investigate equation (5.69) by considering two different cases: firstly, we will consider a body which is being elastically deformed but not rotated (which we expect will lead to classical infinitesimal strain) and then we will consider a strained body undergoing a rigid rotation.

Now, it is well-known that the skew-symmetric part of the velocity gradient is representative of the rate of rotation (see, for example, Spencer [176]). If we are concerned with the case where there is elastic deformation but no rotation, this means that $\text{skw} \( L \) = 0$. Thus, equation (5.69) becomes

$$\frac{D}{Dt} e + \text{tr}(e) \text{sym} \( L \) = \text{sym} \( L \). \quad (5.70)$$

Now, consider the material derivative of classical strain. Since $\frac{\partial u}{\partial x} = I - F^{-1}$, we find that

$$\frac{D}{Dt} \left( \frac{\partial u}{\partial x} \right) = - \frac{DF^{-1}}{Dt}$$

$$= F^{-1} \frac{DF}{Dt} F^{-1}$$

$$= \left( I - \frac{\partial u}{\partial x} \right) L.$$ 

Thus,

$$\frac{D}{Dt} (e_{\text{classical}}) = \frac{D}{Dt} \left( \text{sym} \left( \frac{\partial u}{\partial x} \right) \right)$$

$$= \text{sym} \( L \) - \text{sym} \left( \frac{\partial u}{\partial x} \right) L.$$
For small deformations, we note that $\frac{\partial \mathbf{u}}{\partial x} \mathbf{L}$ is one asymptotic order smaller than $\mathbf{L}$. Thus, we find that equation (5.70) is leading order equivalent to the result obtained from taking the material derivative of classical infinitesimal strain.

Lastly, we wish to describe what happens when a strained body is rotated through space. Let us then consider a body defined so that at time $t = 0$ the deformation gradient at a given material particle is $\mathbf{F}(0) = \mathbf{F}_0$ and the strain is $\mathbf{e}(0) = \mathbf{e}_0$. If we apply a time-varying rotation to the body, we find that the deformation gradient at the particle being considered becomes

$$
\mathbf{F}(t) = \mathbf{Q}(t) \mathbf{F}_0,
$$

where $\mathbf{Q}$ is a proper orthogonal tensor for all values of $t$ and $\mathbf{Q}(0) = \mathbf{I}$.

We note that rotating a body in space is equivalent to rotating the Eulerian coordinate system. That is,

$$
\mathbf{x}_{\text{new}} = \mathbf{Q} \mathbf{x}_{\text{old}},
$$

where $\mathbf{x}_{\text{new}}$ is an Eulerian vector at time $t$ and $\mathbf{x}_{\text{old}}$ is the corresponding vector at time 0. Also, we recall that strain is a one-point tensor; we can see this clearly from the fact that the quadratic form $d\mathbf{x}^T \mathbf{e} d\mathbf{x}$ is physically meaningful when $d\mathbf{x}$ and $d\mathbf{x}^T$ are both vectors in Eulerian space. Now, for the strain tensor to be physically sensible, we expect that

$$
d\mathbf{x}_{\text{new}}^T \mathbf{e}(t) d\mathbf{x}_{\text{new}} = d\mathbf{x}_{\text{old}}^T \mathbf{e}_0 d\mathbf{x}_{\text{old}}.
$$

Thus, we expect that the strain at any given time will be given by

$$
\mathbf{e}(t) = \mathbf{Q}(t) \mathbf{e}_0 \mathbf{Q}^T(t).
$$

(5.71)

We now wish to determine whether or not equation (5.71) is consistent with equation (5.69). Firstly, we note that

$$
\mathbf{L} = \frac{D\mathbf{F}}{Dt} \mathbf{F}^{-1} = \frac{D\mathbf{Q}}{Dt} \mathbf{F}_0 \mathbf{F}_0^{-1} \mathbf{Q}^T = \frac{D\mathbf{Q}}{Dt} \mathbf{Q}^T.
$$
Furthermore, it follows from the work described in Section 5.3.3 that $L$ is skew-symmetric and thus $\text{sym}(L) = 0$. Hence, equation (5.69) can be simplified to give
\[ \frac{De}{Dt} + eL - Le = 0. \]

Now, consider what happens when we take the material time derivative of $e$ as defined in equation (5.71). In this case, we would find that
\[
\frac{De}{Dt} = \frac{DQ}{Dt} e_0 Q^T + Q e_0 \frac{DQ^T}{Dt}
\]
\[
= \frac{DQ}{Dt} Q^T Q e_0 Q^T + Q e_0 Q^T Q \frac{DQ^T}{Dt}
\]
\[
= Le - eL,
\]

since $Q \frac{DQ^T}{Dt} = (\frac{DQ}{Dt} Q^T)^T$ and $\frac{DQ}{Dt} Q^T = L$ is skew-symmetric. Thus, equation (5.71) is consistent with our physical expectations for the rotation of a strained body, giving us further confirmation that (5.64) is the correct equation for describing the evolution of infinitesimal strain.

5.5 Comparison with other mechanical treatments of growth

5.5.1 The Lagrangian growth equation

The growth equation that we have developed has the potential to be used to model a wide range of biomechanical phenomena. However, it is only one of a number of similar theories in the existing mechanical and biological literature. Having fully developed our mathematical treatment of growth in a general form, we will now compare and contrast it with other related theories.
Specifically, we will concentrate on two of the prominent theories of biomechanical remodelling in the current literature: the continuous growth model described by Rodriguez et al. [159] and the cumulative growth model given by Goriely and Ben Amar [70]. It should be noted, however, that these are not the only models of growth and remodelling that have been developed.

For example, Cook [33] modelled wound plasticity by assuming that the eigenvectors of the strain tensor are unaffected by plastic changes, but the eigenvalues evolve over time. This led him to construct an evolution equation for the zero stress Cauchy-Green tensor that reflects both remodelling plasticity (i.e. changes due to cells replacing stressed fibres with unstressed fibres) and mechanical plasticity (i.e. changes that occur when the strain crosses a plasticity boundary). Interestingly, it is possible to show that we can obtain a very similar formulation to Cook’s if we assume an appropriate form for $G$. However, we do not discuss this here.

Another model of growth in the current literature is the recent work by Ambrosi and coworkers [3, 4, 5]. In their main paper, Ambrosi and Guana [3] use thermodynamic arguments to propose a possible equation for the evolution of the zero stress state. In our work, however, we have consciously avoided the concept of strain energy because of the complexity and ambiguity that it introduces. Thus, there are some fundamental difficulties involved in directly comparing the two approaches. Despite this, it would be interesting to perform a detailed comparison of our model with the models proposed by Ambrosi and coworkers. However, this is beyond the scope of the present work.

In the previous chapter, we discussed the approach taken by Rodriguez et al. [159] in their model of the evolving zero stress state. They separated the deformation gradient into an elastic component, $F_e$, and a growth (or plastic) component, $F_g$. Converting to our notation, $F_e$ is equivalent to $Y^{-1}$ and $F_g$ is equivalent to $Z$.

\footnote{Consideration of entropy leads to an inequality that the evolving strain energy must satisfy. Unfortunately, this inequality does not generally yield a unique way of specifying the rate of growth [3, 4].}
Rodriguez et al. used their representation of the zero stress state to construct a definition of strain (see Appendix C for further details). Furthermore, they proposed a model for the evolution of $F_g$ over time; that is, they developed a theory of morphoelasticity that we will now compare with the theory that we developed in Section 5.3.

However, our theory describes the evolution of $Y$ over time; in contrast, Rodriguez et al. were concerned with the evolution of $F_g = Z$ over time. Indeed, all of the authors who have considered the evolution of the zero stress state in a biological context have expressed their results in terms of the evolution of $Z$ or a related tensor (see, for example, Cook [33], Lubarda and Hoger [113], and Goriely and coworkers [70, 71]). Therefore, we firstly need to construct Lagrangian equivalents for equations (5.42) and (5.44) before we can compare our results with the existing theories.

From equation (5.42), we recall that

$$\det (Y) \left( Y^{-1} \frac{DY}{Dt} + L + Y^{-1} S Y \right) = G,$$

where $G$ is the growth tensor and $S$ is a skew symmetric tensor chosen so that $Y$ is always symmetric.

Now, equation (5.42) is physically meaningful because of the results that we obtain when we premultiply both sides of the equation by $dn^T$, an infinitesimal vector that is normal to opposite faces of a small prism, and postmultiply both sides of the equation by $dx$, another infinitesimal vector that represents the distance between those faces. However, $dn$ and $dx$ are defined in the Eulerian coordinate system and we are seeking a Lagrangian form of equation (5.42).

Thus, let us consider infinitesimal vectors $dX$ and $dN$ in Lagrangian space, such that they can be used to define a small prism in the same manner as $dx$ and $dn$. If the Lagrangian prism and the Eulerian prism correspond to the same set of
particles, we will find that

\[ dx = F \, dX, \]

and that

\[ dn = \text{cof} \left( F \right) \, dN = \det \left( F \right) \, F^{-T} \, dN, \]

using the results that we obtained in section 5.3.2.

This means that the Eulerian equation for growth,

\[ dn^T \left[ \det \left( Y \right) \left( Y^{-1} \frac{DY}{Dt} + L + Y^{-1} S \, Y \right) \right] \, dx = dn^T \, G \, dx, \]

can be rewritten as a Lagrangian equation as follows:

\[ dN^T \, \text{det} \left( F \right) \, F^{-1} \left[ \det \left( Y \right) \left( Y^{-1} \frac{DY}{Dt} + L + Y^{-1} S \, Y \right) \right] \, F \, dX = dN^T \, \text{det} \left( F \right) \, F^{-1} \, G \, F \, dX. \]

Omitting \( dN \) and \( dX \) and replacing \( Y \) with \( Z \, F^{-1} \), we find that

\[ \text{det} \left( F \right) \, F^{-1} \left[ \frac{\text{det} \left( Z \right)}{\text{det} \left( F \right)} \left( F \, Z^{-1} \frac{D}{Dt} \left( Z \, F^{-1} \right) + L + F \, Z^{-1} \, S \, Z \, F^{-1} \right) \right] \, F = \text{det} \left( F \right) \, F^{-1} \, G \, F. \quad (5.72) \]

Now,

\[ F \, Z^{-1} \frac{D}{Dt} \left( Z \, F^{-1} \right) = F \, Z^{-1} \frac{DZ}{Dt} \, F^{-1} + F \, Z^{-1} \, Z \, \frac{DF^{-1}}{Dt}; \]
and furthermore,

$$F Z^{-1} Z \frac{D F^{-1}}{D t} = F \frac{D F^{-1}}{D t}$$

$$= -F F^{-1} \frac{D F}{D t} F^{-1}$$

$$= -L.$$  

Thus, equation (5.72) becomes

$$\det (F) F^{-1} \left[ \frac{\det (Z)}{\det (F)} \left( F Z^{-1} \frac{D Z}{D t} F^{-1} - L + L + F Z^{-1} S Z F^{-1} \right) \right] F = \det (F) F^{-1} G F,$$

or, equivalently,

$$\frac{\det (Z)}{\det (F)} \left( Z^{-1} \frac{D Z}{D t} + Z^{-1} S Z \right) = F^{-1} G F. \quad (5.73)$$

Equation (5.73) is equivalent to (5.42), the Eulerian definition of the growth tensor. Furthermore, we can use (5.73) to construct a Lagrangian equation analogous to (5.44). By premultiplying both sides of (5.73) by $Z$ and postmultiplying both sides by $Z^{-1}$, it is possible to isolate the skew-symmetric tensor, $S$. Thus, we find that

$$\frac{\det (Z)}{\det (F)} \text{sym} \left( \frac{D Z}{D t} Z^{-1} \right) = \text{sym} \left( (Z F^{-1}) G (Z F^{-1})^{-1} \right). \quad (5.74)$$

We note that $Z F^{-1} = Y$. Hence, we will again find that if the principal directions of growth are parallel to the principal directions of strain, equation (5.74)
can be simplified to yield

\[
\frac{\det(Z)}{\det(F)} \text{sym}\left(\frac{DZ}{Dt}Z^{-1}\right) = G. \tag{5.75}
\]

Interestingly, we can use (5.74) to recover equation (5.11), the one-dimensional Lagrangian growth equation. In the one-dimensional case, we can replace \(Z\), \(F\) and \(G\) with the scalars \(z\), \(f\) and \(g\). Furthermore, when \(A\) is one-dimensional, we note that \(\det(A) = A\) and \(\text{sym}(A) = A\). Thus, we find that equation (5.74) becomes

\[
\frac{DZ}{Dt} = gF,
\]

which is identical to (5.11).

Before we compare (5.74) with the equations proposed by other authors, it is important to consider what would happen if we used a choice of \(Z\) other than the principal zero stress deformation gradient. As we discussed in Section 5.3.3 for the Eulerian case, any rotation in the zero stress deformation gradient can be absorbed into \(S\). Since \(S\) has been completely eliminated from equation (5.74), this means that the validity of the equation is unaffected by the choice of \(Z\).

This is important because many of the existing models of morphoelasticity specify that \(Z\) is symmetric, while we have chosen our principal zero stress deformation gradient so that \(Y = ZF^{-1}\) is symmetric. Choosing \(Z\) to be symmetric has some advantages when using equation (5.74). Most importantly, it is much easier to track the six independent components of \(Z\) when these are represented by the six components of a symmetric tensor.

However, the disadvantage of using a symmetric \(Z\) instead of the principal zero stress deformation gradient is that equation (5.75) is no longer appropriate.\(^{42}\)

This is because the simplification used to obtain (5.75) relies on the fact that

\[^{42}\text{If the physical constraints of our problem are such that the principal Lagrangian zero stress deformation gradient is always symmetric, then equation (5.75) can still be used. However, this will not generally be the case.}\]
\( Z F^{-1} \) is a symmetric matrix with the same eigenvectors as the strain tensor.

### 5.5.2 Rodriguez et al. and the rate of growth tensor

The first significant contribution to the theory of morphoelasticity was made by Rodriguez et al. [159]. As mentioned in Section 2.4, Rodriguez et al. proposed two different measures of how the zero stress state evolves over time: the rate of growth stretch tensor, \( \dot{U}_g \), and the rate of growth tensor, \( D_g \).

The rate of growth stretch tensor is defined by Rodriguez et al. as the material time derivative of \( U_g \), where \( U_g \) is the right stretch tensor associated with the growth deformation gradient. Hence, \( U_g \) has the property that

\[
F_g = R_g U_g,
\]

where \( F_g \) is the growth deformation gradient (equivalent to our \( Z \)), \( R_g \) is proper orthogonal and \( U_g \) is symmetric.

In turn, Rodriguez et al. defined the rate of growth tensor, \( D_g \), to be the symmetric part of the velocity gradient associated with the evolving \( U_g \). That is,

\[
D_g = \text{sym} \left( \dot{U}_g \left( U_g^{-1} \right) \right).
\]

Rodriguez et al. also gave an equation for \( \dot{U}_g \) in terms of \( D_g \) and \( U_g \), but it is significantly more complicated than the definition of \( D_g \) given above. However, the existence of this equation indicates that, for a given \( U_g \), there is a one-to-one correspondence between \( D_g \) and \( \dot{U}_g \).

Because of this one-to-one correspondence, it is always possible to convert from one representation of the rate of growth to the other. However, it is much easier to obtain a solution for \( F_g \) if \( \dot{U}_g \) is known, compared to the case where \( D_g \) is known. Thus, Rodriguez et al. proposed that growth problems should preferably
be expressed in terms of $\dot{U}_g$ rather than in terms of $D_g$.

Although it is more mathematically convenient to use $\dot{U}_g$ than $D_g$, we also need to be able to give a physical interpretation for any measure of growth that we use; it does not make sense to stipulate $\dot{U}_g$ unless this is physically meaningful. However, Rodriguez et al. do not seem to have given this much consideration. In one of their examples, they specified that the components of $\dot{U}_g$ were linearly dependent on the components of stress, but they did not give any justification for why this should be the case. In fact, Rodriguez et al.’s stress tensor is defined in a different coordinate system from $\dot{U}_g$, meaning that such a formulation would generally be invalid.\textsuperscript{43}

Rodriguez et al. were able to physically interpret some aspects of $D_g$; specifically, they were able to relate the trace of $D_g$ to the rate of volumetric growth. However, they were not able to construct any similar results for $\dot{U}_g$. This suggests that it may be more physically meaningful to specify $D_g$ in growth problems than $\dot{U}_g$; perhaps a more correct approach to stress-dependent growth would be to make the components of $D_g$ linear functions of strain.

In order to investigate this further, we will compare the work of Rodriguez et al. with our own theory of morphoelasticity. Importantly, we recall that the growth tensor, $G$, was developed as a physically meaningful measure of the rate and direction of growth. As discussed in Sections 5.3.2 and 5.3.3, the eigenvectors of $G$ correspond to the principal directions of growth. Furthermore, each eigenvalue of $G$ indicates the rate of desired volume growth per unit current volume in the direction of its associated eigenvector.

Let us now compare $\dot{U}_g$ and $D_g$ with $G$. If $Z$ is a Lagrangian zero stress defor-

\textsuperscript{43}Rodriguez et al. comment on the fact that it may be necessary to rotate the stress tensor so that it is defined in the same frame as the rate of growth stretch tensor. However, we should still be very careful with stress tensors and growth tensors that are defined in different coordinate systems. For example, in the Lagrangian equation (5.73) we note that we use $F^{-1} G F$ where the Eulerian equation (5.28) has $G$. This is like converting between the Cauchy stress tensor and the Second Piola-Kirchoff stress tensor; we find that the growth tensor is rotated \textit{and deformed} as we go from one coordinate system to the other.
utation gradient and \( Q \) is proper orthogonal, we recall that \( \dot{Z} = QZ \) is also a valid zero stress deformation gradient. Hence, equation (5.76) can be rearranged to yield

\[
U_g = R_g^T F_g = R_g^T Z,
\]

and it follows that \( U_g \) is a valid Lagrangian zero stress deformation gradient. We also note that \( U_g \) will not generally be the principal zero stress deformation gradient; \( U_g \) corresponds to the choice of \( Z \) that makes \( Z \) symmetric, not the choice of \( Z \) that makes \( ZF^{-1} \) symmetric. Thus, \( \dot{U}_g \) and \( D_g \) can be expressed in our notation as

\[
\dot{U}_g = \frac{DZ}{Dt},
\]

and

\[
D_g = \text{sym}\left( \frac{DZ}{Dt} Z^{-1} \right),
\]

where \( Z \) is chosen to be symmetric.

Now, we recall from equation (5.74) that

\[
\frac{\det(Z)}{\det(F)} \text{sym}\left( \frac{DZ}{Dt} Z^{-1} \right) = \text{sym}\left( (ZF^{-1}) G (ZF^{-1})^{-1} \right),
\]

regardless of the choice of \( Z \). Thus, we find that

\[
\frac{\det(U_g)}{\det(F)} D_g = \text{sym}\left( (U_g F^{-1}) G (U_g F^{-1})^{-1} \right). \tag{5.78}
\]

Interestingly, this equation involves \( D_g \) but not \( \dot{U}_g \). This suggests that \( D_g \) is a more physical measure of growth than \( \dot{U}_g \).

Indeed, if we are in a situation where \( G \) commutes with \( U_g F^{-1} \); that is, where the principal zero stress deformation gradient is symmetric and the principal directions of growth are parallel with the principal directions of strain; we find that

\[
\frac{\det(U_g)}{\det(F)} D_g = G. \tag{5.79}
\]
In this case, the eigenvectors of $D_g$ correspond to the principal directions of growth/strain and the eigenvalues of $D_g$ are indicative of the rate of growth associated with each principal direction.

We note that the eigenvalues of $D_g$ are proportional to, but different from, the eigenvalues of $G$. This is because $G$ and $D_g$ are defined in different coordinate systems. We can see this clearly by comparing the traces of these two tensors. From the original definition of the growth tensor, we recall that the trace of $G$ is equal to $g$, the rate of desired volume growth (i.e. the rate of volume growth in a locally constructed zero stress coordinate system) per unit of current volume.

In contrast, we can take the trace of both sides of equation (5.79)\(^4\) to find that

$$\text{tr} \left( D_g \right) = \frac{\det(F) g}{\det(U_g)}.$$  

We recall from (5.14) that the determinant of $F$ is the ratio of infinitesimal volumes in the current and initial configurations. Similarly, we find that $\det(U_g)$ is the ratio of infinitesimal volumes in the zero stress and initial configurations. Thus, it follows that $\text{tr}(D_g)$ is equal to the rate of volume growth in the zero stress configuration per unit volume in the zero stress configuration.

If $G$ does not commute with $U_g F^{-1}$, the interpretation of $D_g$ is much less clear. Although the trace of $D_g$ will still be representative of the rate of volumetric growth, it is difficult to see how the eigenvalues or eigenvectors of $D_g$ should be understood in context of growth. Similarly, it is unclear as to whether $\dot{U}_g$ can be assigned a physical interpretation, regardless of whether or not $G$ commutes with $U_g F^{-1}$.

Thus, we conclude that the two measures of growth described by Rodriguez et al.\(^4\)

\(^4\)In fact, we can take the trace of both sides of equation (5.78) and obtain the same result, since

$$\text{tr} \left( \text{sym} \left( ABA^{-1} \right) \right) = \text{tr} \left( ABA^{-1} \right) = \text{tr} \left( B \right).$$
al. are less than ideal. Although $D_g$ can be clearly interpreted in some circumstances, it is less appropriate in cases where the principal directions of growth are not parallel to the principal directions of strain, or in situations where $U_g$ is not the principal zero stress deformation gradient. The rate of growth stretch tensor would appear to be even less informative. Despite the fact that Rodriguez et al. prefer it over $D_g$, we suspect that it is not physically meaningful in isolation and that it is only usable in cases where the evolution of $U_g$ is already well known.

Having said this, the achievement of Rodriguez et al. in introducing a mathematically coherent theory of morphoelasticity to biomechanics cannot be overstated. By demonstrating the usefulness of the multiplicative decomposition of the deformation gradient and showing how it could be used to describe the genesis of residual stresses in an elastic tissue, Rodriguez et al. laid the foundation for all subsequent work in the area of biological morphoelasticity; their paper remains one of the definitive works in this area to the current day.

5.5.3 Goriely and Ben Amar’s cumulative theory of growth

One of the most significant recent developments in morphoelasticity is the cumulative theory of growth proposed by Goriely and Ben Amar [70]. As we will see, this theory differs from the work of Rodriguez et al. (and our own morphoelastic theory) because it focuses on the changes that occur over a small timestep; Goriely and Ben Amar do not explicitly construct differential equations.

Thus, we seek to develop a differential equation that reflects Goriely and Ben Amar’s cumulative approach before we attempt to compare their work with our own theory of morphoelasticity. Firstly, consider the mathematical approach taken by Goriely and Ben Amar. Following the example of Rodriguez et al., they used a multiplicative decomposition of the deformation gradient to represent the
plasticity associated with growth. In their notation,

$$F = AG,$$

where $A$ is representative of the elastic part of the deformation (equivalent to $F_e$ or $Y^{-1}$) and $G$ is representative of the growth/plastic part of the deformation (equivalent to $F_g$ or $Z$). To avoid confusion with our own growth tensor, $G$, we will use $Z$ to represent the Lagrangian zero stress deformation gradient throughout the rest of this section, but we will retain Goriely and Ben Amar’s definition of $A$.

Goriely and Ben Amar began their analysis by considering a body that is free of residual stresses at $t = 0$. That is, we find that $Z = I$ is a valid zero stress deformation gradient throughout the body at $t = 0$. After a small time step, $\Delta t$, $Z$ evolves so that

$$Z(X, \Delta t) = Z_1(X),$$

and $Z_1$ represents a small deformation (i.e. $Z$ is close to being proper orthogonal).

This new zero stress deformation gradient, $Z_1$, may or may not be compatible with a real deformation. Furthermore, there may be body forces, applied surface stresses and other mechanical considerations that may contribute to the presence of a nonzero elastic strain. Taking all of these into account, it is possible to find the elastic deformation, $A_1$, so that internal forces are in balance and

$$F(X, \Delta t) = A_1 Z_1,$$

where $F$ is a valid deformation gradient.

Goriely and Ben Amar then consider what happens after a second time step. Using $F(X, \Delta t)$ as a new reference configuration, they propose that

$$F(X, 2\Delta t) = A_2 Z_2 F(X, \Delta t) = A_2 Z_2 A_1 Z_1,$$

where $Z_2$ is representative of the growth that takes place between $t = \Delta t$ and
\[ t = 2\Delta t \] and \( A_2 \) is representative of the elastic deformation needed in order to take any new incompatibilities or applied forces into account.

Unfortunately, the tensors \( Z_2 \) and \( A_2 \) are not as easy to interpret as \( Z_1 \) and \( A_1 \). As illustrated in Figure 2 of Goriely and Ben Amar’s paper, \( Z_2 \) represents the deformation from \( B_1 \), the current state of the body at \( t = \Delta t \), to \( V_2 \), a virtual state that reflects any new incompatibilities introduced by growth. However, \( V_2 \) is not a zero stress state of the body, neither does it correspond to a real, physical deformation. Instead, it is a virtual configuration of the body that cannot be physically interpreted unless some information about \( B_1 \) is known.

Despite this, \( Z_2 \) is still a useful measure of growth; in fact, we will show that \( Z_2 \) is closely related to our own growth tensor, \( G \). The main reason for the usefulness of \( Z_2 \) is that the deformation from \( B_1 \) to \( V_2 \) can be related to the deformation from \( V^{(1)} \), the zero stress state at \( t = \Delta t \), to \( V^{(2)} \), the zero stress state at \( t = 2\Delta t \). Thus, even though \( Z_2 \) represents a deformation from the current state of the body to a physically meaningless virtual state, it can be used to construct an equation for \( Z(X, 2\Delta t) \). In order to obtain such an equation, Goriely and Ben Amar made use of the earlier work by Hoger et al. \[94\] on modelling the zero stress state of residually strained materials.\[45\]

45In their paper, Hoger et al. discuss the case where the reference state of a tissue (i.e. the state which is used to define Lagrangian coordinates) is residually strained. As Hoger et al. wish to retain the property that \( F = \frac{\partial x}{\partial X} = F_e F_g \), this means that \( F_e \) can no longer be used to fully define strain; instead, strain depends not only on \( F_e \) but on \( V_0 \), a new tensor introduced to represent the residual strain in the reference configuration.

Our use of stress-relieving deformations to define the zero stress state allows us to avoid this complexity; the initial residual strain and any residual strains that develop over time are all incorporated into \( Z \). However, the fact that Hoger et al. include an explicit representation of the initial residual strain makes their work very useful to Goriely and Ben Amar. In the notation of Hoger et al., Goriely and Ben Amar effectively update the initial residual strain, \( V_0 \), after every timestep. They then use Hoger et al.’s methods to analyse the elastic and plastic parts (\( F_e \) and \( F_g \)) of the small deformation that occurs in the next timestep.

It is impossible to use the zero stress state theory that we developed in the previous chapter in a similar manner. Our representation of the zero stress state has the advantage that all of the residual strains are combined into one tensor. However, we cannot directly follow Goriely and Ben Amar’s methods using our representation of the zero stress state to show that they are equivalent.
Interestingly, one part of their analysis required Goriely and Ben Amar to construct a polar decomposition of $Z_2$ in the form

$$Z_2 = V_2 R_2,$$

where $V_2$ is symmetric and $R_2$ is proper orthogonal. The purpose of using this decomposition was to take into account the possibility that the principal axes of growth are different from the principal axes of strain. Thus, we expect that Goriely and Ben Amar’s equations will be an improvement on those developed by Rodriguez et al. [159], which run into difficulties if the axes of growth are different from the axes of strain.

Ultimately, Goriely and Ben Amar were able to show that

$$Z(X, 2\Delta t) = A_1^{-1} R_2^T Z_2 A_1 Z_1,$$

where $R_2$ is defined as above. Having obtained this result, Goriely and Ben Amar were easily able to extend their work to the general case of a small deformation from time $t$ to time $t + \Delta t$. We summarise their results as follows:

Let the current deformation gradient be expressed as the product of an elastic deformation, $A(t)$, and a zero stress deformation gradient, $Z(t)$, so that

$$F(t) = A(t) Z(t).$$

After a short period of time, $\Delta t$, we introduce a growth change, $Z_*$, and an elastic change, $A_*$, so that

$$F(t + \Delta t) = A_* Z_* A(t) Z(t).$$

Now, we construct a polar decomposition of $Z_*$ so that

$$Z_* = V_* R_*,$$

\[\text{For convenience and clarity, we omit the} \ X \ \text{dependence of} \ F, \ A \ \text{etc. from the equations given here.}\]
where $\mathbf{R}_s$ is proper orthogonal and $\mathbf{V}_*$ is symmetric. In this case, we find that the method developed by Goriely and Ben Amar tells us that $\mathbf{Z}(t + \Delta t)$ is given by

$$\mathbf{Z}(t + \Delta t) = \mathbf{A}^{-1}(t) \mathbf{R}_s^T \mathbf{Z}_* \mathbf{A}(t) \mathbf{Z}(t),$$

or equivalently,

$$\mathbf{Z}(t + \Delta t) = \mathbf{A}^{-1}(t) \mathbf{R}_s^T \mathbf{V}_* \mathbf{R}_s \mathbf{A}(t) \mathbf{Z}(t). \quad (5.80)$$

Now, consider what happens in the limit as $\Delta t$ goes to zero. Since

$$\lim_{\Delta t \to 0} \mathbf{Z}(\mathbf{X}, t + \Delta t) = \mathbf{Z}(\mathbf{X}, t),$$

equation (5.80) indicates that

$$\lim_{\Delta t \to 0} \mathbf{V}_* = \mathbf{I}.$$ 

Note that it makes sense for $\mathbf{V}_*$ to approach $\mathbf{I}$ as $\Delta t$ approaches zero, because $\mathbf{V}_*$ is a measure of the deformation and we expect there to be no deformation over an infinitesimally small time interval.

Now, consider the tensor $\mathbf{W}$, defined so that

$$\mathbf{W} = \lim_{\Delta t \to 0} \frac{\mathbf{V}_* - \mathbf{I}}{\Delta t}.$$ 

Since $\mathbf{W}$ is the derivative of $\mathbf{V}_*$ with respect to $t$, the theory of Taylor series tells us that we can use $\mathbf{W}$ to describe the behaviour of $\mathbf{V}_*$ when $\Delta t$ is very small. That is,

$$\mathbf{V}_* \sim \mathbf{I} + \mathbf{W} \Delta t, \quad (5.81)$$

and it follows that $\mathbf{W}$ is a measure of the speed and direction of growth. Thus, we expect to find that $\mathbf{W}$ is related to our own growth tensor, $\mathbf{G}$.

Considering very small values of $\Delta t$ and substituting (5.81) into (5.80), we find
that
\[ Z(t + \Delta t) \sim Z(t) + A^{-1}(t) R_s^T W R_s A(t) Z(t) \Delta t, \]
or equivalently,
\[ \frac{Z(t + \Delta t) - Z(t)}{\Delta t} \sim A^{-1}(t) R_s^T W R_s A(t) Z(t). \]

Thus, in the limit as \( \Delta t \to 0 \), we find that
\[ \frac{\partial Z}{\partial t} = A^{-1}(t) R_s^T W R_s A(t) Z(t), \]
and hence,
\[ \frac{\partial Z}{\partial t} Z^{-1} = A^{-1} R_s^T W R_s A. \] (5.82)

Now, we can compare this result with (5.74), the Lagrangian growth equation that we developed in Section 5.5.1. Since we are working in Lagrangian coordinates, we note that \( \frac{\partial Z}{\partial t} = D Z Dt \). Also, we recall that \( A = Y^{-1} = (ZF)^{-1} \). Hence, equation (5.74) can be written as
\[ \frac{\det(Z)}{\det(F)} \text{sym} \left( \frac{\partial Z}{\partial t} (Z)^{-1} \right) = \text{sym} (A^{-1} G A). \] (5.83)

This is almost identical to equation (5.82). If we define \( G \) using \( W \) and \( R_s \) so that
\[ G = \frac{\det(Z)}{\det(F)} R_s^T W R_s, \]
we will find that equation (5.83) is always satisfied. Importantly, we note that
\[ (R_s^T W R_s)^T = R_s^T W^T R_s. \]

Hence, the fact that \( W \) is symmetric implies that \( G \) is symmetric.

There are a number of apparent differences between (5.82) and (5.83) that are worthy of further consideration. Most obviously, we note that equation (5.82)
equates two entire tensors, while (5.83) only equates the symmetric parts of the
two tensors. Thus, equation (5.82) would appear to be a stronger result than
equation (5.83).

In order to understand this, we recall that our development of equation (5.83)
took into account the fact that any rotation of the zero stress coordinate system
is physically meaningless. Thus, if \( Z(t) \) satisfies equation (5.83), so will \( \hat{Z}(t) = Q(t) Z(t) \). If we wish to solve (5.83) to obtain a unique solution for \( Z \), we need to
specify which of the many equivalent choices for \( Z \) is preferred; for example, we
could stipulate that \( Z \) is always symmetric or that \( Z \) is the principal zero stress
deformation gradient. In contrast, equation (5.82) does not give us the freedom
to choose between a family of equivalent \( Z \) tensors. For a given \( Z_* \), equation
(5.82) yields a unique value for \( Z(t + \Delta t) \), despite the fact that \( QZ \) would also
represent the same physical situation.

It is unclear as to whether the unique \( Z_L \) obtained from Goriely and Ben Amar’s
method has any special properties. Possibly, using equation (5.82) yields the
same results as requiring that \( Z \) is symmetric in some conditions. It may even be
true that (5.82) leads us to recover the principal zero stress deformation gradient.
However, this would not generally seem to be the case.

A second point that should be commented on is the relationship between \( G \) and
\( W \). Firstly, we note that \( G \) is defined in terms of \( R_*^T W R_* \). Now, \( W \) is defined
so that
\[
W = \lim_{\Delta t \to 0} \frac{V_* - I}{\Delta t}.
\]

Thus,
\[
R_*^T W R_* = \lim_{\Delta t \to 0} \frac{R_*^T V_* R_* - I}{\Delta t}
= \lim_{\Delta t \to 0} \frac{U_* - I}{\Delta t},
\]

where \( U_* \) is the right stretch tensor associated with \( Z_* \).
This means that our definition of $G$ can be rewritten as

$$G = \frac{\det(Z)}{\det(F)} \hat{W},$$

(5.84)

where

$$\hat{W} = \lim_{\Delta t \to 0} \frac{U_* - I}{\Delta t}.$$  

(5.85)

It is physically sensible that $G$ depends only on the right stretch associated with $Z_*$, since this is equivalent to saying that $G$ is unaffected by rotations in the zero stress coordinate system.

Now, let us compare equation (5.84) with equation (5.79), which showed the relationship between $D_g$, the rate of growth tensor used by Rodriguez et al., and $G$. Since $Z$ and $U_g$ have the same determinant, we find that $\hat{W} = D_g$; both tensors represent a rate of volume growth in the zero stress configuration per unit of desired volume in the zero stress configuration.

However, equation (5.79) was only valid when $G$ commutes with $U_g F^{-1}$. In contrast, equation (5.84) is always valid. Thus, the tensor $\hat{W}$ is effectively a corrected version of $D_g$; whenever $D_g$ has a clear physical interpretation, $D_g = \hat{W}$, but $\hat{W}$ is still meaningful when it is difficult to interpret $D_g$.

Equations (5.84) and (5.85) effectively allow us to obtain $G$ from $Z_*$. As a final note, let us describe how we could reverse this process and obtain $Z_*$ from $G$. Given a short timestep, $\Delta t$, we can consider the leading terms of a Taylor expansion of $U_*$ to find that

$$U_* = I + \hat{W} \Delta t + O(\Delta t^2).$$

We note that the error term is quadratic in $\Delta t$, but we will omit it in the following steps. Using the definition of $G$ from (5.84), we find that

$$U_* = I + \frac{\det(F)}{\det(Z)} G \Delta t.$$
Since $U_*$ is the right stretch tensor of $Z_*$, it follows that $U_*$ is a valid choice of $Z_*$, as is any tensor of the form

$$Z_* = Q + \frac{\det(F)}{\det(Z)} Q G \Delta t,$$

(5.86)

where $Q$ is proper orthogonal.

Despite the apparent differences between Goriely and Ben Amar’s approach to morphoelasticity and our own approach, we have shown that the equations that they obtained are equivalent to ours. Thus, the two approaches can be used to justify each other; the work that we described earlier in this chapter can be used to clarify the physical meaning of $Z_*$, while Goriely and Ben Amar’s work provides us with further confirmation that $G$ is a meaningful measure of growth.

Lastly, we note that the equivalence of the two approaches provides us with a possible avenue for developing a numeric method to solve equation (5.74). Since equation (5.86) is equivalent to (5.74), we could use $G$ to construct $Z_*$ at every timestep and use (5.80) to move forwards in time.\(^{47}\) However, we would need to be very careful in doing this. Not only does this have the potential for propagating errors, but we lose the ability to specify that $Z$ is always the principal zero stress deformation gradient unless an additional rotation step is included. Further work is necessary in order to determine whether there is an easy way of adapting equation (5.80) to incorporate this.

### 5.6 Summary of zero stress state theory and morphoelasticity

In this chapter and the previous one, we carefully developed a mathematical framework for modelling the evolving zero stress state of a material. Before we

\(^{47}\)We recall that the zero stress state
go on to consider applications of this work, it behooves us to take a step backwards and consider what we have achieved. In this section, we summarise our important results and discuss the physical meaning of the variables that we have introduced.

**Static zero stress state theory**

Firstly, let us review what we mean by a zero stress state. In an elastic solid, we expect that the stress at any given point can be expressed as a function of the deformation gradient at that point. Furthermore, we assume that it is always possible to find a deformation gradient tensor that corresponds to a stress of zero at the point being considered. Any deformation gradient tensor with this property is called a zero stress deformation gradient tensor.

The zero stress deformation gradient is not a unique measure of the zero stress state. In general, if $Z$ is a zero stress deformation gradient and $Q$ is proper orthogonal then $QZ$ will also be a zero stress deformation gradient tensor. This is because a point at zero stress will still be at zero stress following a rigid rotation. However, under normal circumstances we expect this to be the only ‘nonuniqueness’ in the definition of the zero stress deformation gradient. That is, it is easy for us to use $Z$ to construct a unique measure of the zero stress state (e.g. $\Theta = Z^T Z$).

Now, the deformation of a body is generally measured using the strain. Strain can best be understood by considering the changing distance between two neighbouring particles, $P$ and $Q$. Using equations (4.16) and (4.17), we find that the classical strain tensor relates the current distance between $P$ and $Q$ to the initial distance between $P$ and $Q$. Thus, a strain of zero corresponds to the case where

---

48Actually, it should be noted that neither of these assumptions are universally true. For example, it is not immediately obvious that each deformation gradient should correspond to a unique stress. In a rigid body, we would find that a deformation gradient equal to the identity tensor could correspond to any stress. Similarly, it may not be possible to find a deformation gradient that corresponds to zero stress. As we describe in Appendix B, further conditions are needed before we can be confident that the zero stress state exists.
the distance between $P$ and $Q$ is the same as what it was originally.

However, we would prefer to define strain so that zero strain corresponds to zero stress. That is, we wish to define strain so that it relates the current distance between $P$ and $Q$ to the desired distance between $P$ and $Q$ (i.e., the distance between $P$ and $Q$ in the zero stress state). This leads us to develop the improved definitions of strain given in equations (4.31) and (4.30).

Having developed appropriate definitions for finite strain, we next considered the case of small strains. Although the zero stress deformation gradient is not uniquely defined, our analysis of small strains led us to construct the principal zero stress deformation gradient. The principal zero stress deformation gradient is uniquely defined so that it is close to the actual deformation gradient. If a point is at zero stress, for example, we find that the principal zero stress deformation gradient is equal to the actual deformation gradient.

We contend that the principal zero stress deformation gradient is the best way to represent the zero stress state of a material. Previously, most researchers have used the right stretch tensor associated with the zero stress deformation gradient as a unique representation of the zero stress state. However, the useful properties of the principal zero stress deformation gradient strongly suggest that it is preferable.

Using the principal zero stress deformation gradient, we were able to construct simple expressions for infinitesimal strain. Unlike the classical definitions of infinitesimal strain, our definitions are frame invariant. That is, we find that our infinitesimal strain remains valid through a general rotation of the coordinate system. This means that our definition of infinitesimal strain can be used in some cases where classical infinitesimal strain is inappropriate.

\[49\] Interestingly, our definition of Lagrangian strain is subtly different from the definition given by Rodriguez et al. [159]. This is because we are careful to avoid using a coordinate system that might be ill-defined.
Morphoelasticity

The results described above give us ways to define the strain with reference to the zero stress state. Having obtained these relationships, we wish to construct an evolution equation to describe how the zero stress state changes over time (i.e. morphoelasticity). Importantly, we require that any parameters we introduce must have clear physical interpretations. It is more important to develop equations that relate interpretable quantities than it is to develop simple equations.

Effectively, we seek to relate the changing zero stress state to an appropriate representation of the desired rate of growth.\(^{50}\) By referring to the desired rate of growth, we emphasise the fact that growth does not necessarily imply a change of physical size. If a growing material is physically constrained, growth will lead to the development of physical stresses. In most cases, growth will simultaneously cause changes to the physical size of a body and its internal stress distribution.

In order to construct a morphoelastic model of growth, we need to consider two things: the net rate of growth and any directional bias to the growth. In one dimension, consideration of net growth leads us to construct complete evolution equations for the zero stress deformation gradient. However, the three-dimensional equation for net growth only yields information about the determinant of the zero stress deformation gradient. In order to obtain a complete evolution equation, we need to take directed growth into consideration.

Although the representation of net growth is clear and unambiguous, it is difficult to find an appropriate way to model directional dependence. In Sections 5.3.2 and 5.3.3, we develop a novel tensorial representation of growth that lends itself naturally to physical interpretation. We contend that our tensor, \(G\), is the most natural way to represent the evolution of the zero stress state. However, as

---

\(^{50}\) We recall that any process that involves a changing zero stress state can be modelled using morphoelasticity. This includes growth, resorption and any plastic deformation. For convenience, we continue to use the term growth to cover all of these.
our theoretical development is not completely rigorous, we cannot be absolutely confident that this is the case.

The growth tensor, $G$, can be related to the desired rate of volume growth of an infinitesimally small prism. Specifically, $G$ allows us to calculate the volumetric growth of a prism where two opposite faces are held at a fixed size but the distance between them is allowed to evolve with the changing zero stress state. By effectively considering a rectangular prism, we are able to separate the growth into three independent and orthogonal components.

Importantly, the eigenvalues and eigenvectors of $G$ have physical interpretations. The eigenvectors of $G$ correspond to the principal directions of growth. That is, one of the eigenvectors of $G$ will be the direction of most growth (or least shrinkage) and one of the other eigenvectors will be the direction of least growth (or most shrinkage). The eigenvalue associated with a given eigenvector represents the rate of growth in that direction.\(^{51}\) As the sum of the eigenvalues, the trace of $G$ therefore represents the net rate of growth.

When $G = 0$, we find that there is no growth and we recover the equations of classical elasticity. Alternatively, when $G = \frac{2}{3}I$, we would find that the growth occurs equally in all directions. The net rate of volumetric growth in this case would be given by $g$. In biological problems, we would expect $G$ to be dependent on a variety of different factors. These could include stress/strain, oxygen concentration, growth factor concentration, nutrient concentration, cell density etc. In the following chapter, we will discuss a number of forms that $G$ might take in practical applications of this theory.

Interestingly, we have developed three different ways of using the tensor $G$ to describe the mechanical evolution of a material. Equation (5.44) shows how the evolution of $Y$, the Eulerian zero stress deformation gradient, can be related to $G$. Similarly, equation (5.74) gives the evolution of $Z$, the Lagrangian zero stress deformation gradient, in terms of $G$. Unfortunately, both of these equations are

\(^{51}\)Positive eigenvalues indicate growth while negative eigenvalues indicate shrinkage.
highly nonlinear and would be difficult to apply computationally.

In contrast, equations (5.57) and (5.64) describe how the Eulerian finite and infinitesimal strain tensors evolve in response to a growth. Effectively, these equations combine the analysis of the evolving zero stress state described in this chapter with the definitions of strain developed in the previous chapter. These equations are the two most important results developed in this thesis.

In general, we expect that equation (5.64) will be easier to analyse and apply computationally than equation (5.57). However, it is more restricted. Although (5.57) is valid for all deformations, equation (5.64) depends on the assumption that the growth is directed so that the residual strain is minimised. Fortunately, we expect that this will often be true for biological growth. It is reasonable to assume that cells will preferentially divide in the direction of least compressive stress.

The theory of morphoelasticity that we present is very useful for describing the changes to the fundamental mechanical structure of a biological material. However, it is important for us to be aware of the fact that it cannot account for all mechanical changes. In particular, we note that morphoelasticity only deals with the evolution of strain; it is still necessary to specify a constitutive law that relates strain to stress. Thus, if internal remodelling causes a tissue to become stiffer, we will need to model this change using other methods (e.g. by changing the Young’s modulus). Similarly, our theory does not deal with the anisotropy of a tissue’s mechanical response. Instead, it is useful for describing any anisotropy in the evolution of the zero stress state.

In conclusion, we have developed a mathematical theory of morphoelasticity that leads to practical and interpretable evolution equations for the zero stress state and strain. These equations incorporate a novel tensorial representation of growth that we believe is superior to previous representations. The next step is to show that these equations lead to results that are consistent with our physical expectations. Furthermore, we would like to use insights that we have obtained from
this work to construct an improved mechanical model of dermal wound healing. This brings us to the topics that we will consider in Chapter 6.
Chapter 6

Applications of zero stress state theory and morphoelasticity

6.1 Introduction

The theory of morphoelasticity that we have developed allows us to describe the evolution of the zero stress state in response to a prescribed rate of growth. In a biological context, a changing zero stress state can be used to model a variety of processes that involve internal remodelling of a tissue’s mechanical structure. Such processes include soft tissue growth, arterial remodelling and aneurysms, wound contracture, morphogenesis and solid tumour growth.

In this chapter we apply morphoelasticity theory to construct models of physical processes. We consider a range of different applications, allowing us to demonstrate the flexibility and versatility of our theory. By making different assumptions about the dependence of growth on other physical parameters, it is possible to model a wide range of biological processes.

Moreover, morphoelasticity is not limited to biological applications. In fact, we
recall that the multiplicative decomposition of the deformation gradient was originally proposed by Lee [107] to describe metal plasticity. Any process that involves a continuous change to the zero stress state of a material can be modelled using morphoelasticity. Importantly, this means that morphoelasticity can be incorporated into models of nonbiological materials that exhibit plastic or viscoelastic behaviour.

Although our emphasis is on biomechanical modelling, we begin this chapter by considering a nonbiological application. In Section 6.2, we show that morphoelasticity can be used to obtain the classical model of a Maxwell material (a linear viscoelastic fluid). In many ways, it is surprising that we can model a Maxwell material in this way. Such materials are fluids, but our theory of morphoelasticity developed from considering a growing solid. As we will see, it is possible to represent the viscous flow of a Maxwell material by making the rate of ‘growth’ dependent on stress.

Similarly, we expect that biological growth will be dependent on a variety of parameters, including stress. In Section 6.3, we describe a model for the stress-dependent growth of a soft tissue. By assuming that growth is directed to minimise stress, we construct a simple equation in which the growth tensor is expressed as a function of the stress tensor. Interestingly, we find that this equation involves only two parameters: the total rate of growth and the directional sensitivity of growth to stress.

As a case study, we consider the growth of a tumour spheroid. As described in recent reviews [8, 160], tumour growth is a highly productive area for mathematical modelling. Although we do not fully develop our morphoelastic model, we are able to obtain some interesting and general preliminary results. Our model yields predictions about the stress distribution within a tumour spheroid that can

---

1In the past, there has been some debate as to whether biological growth is dependent on stress or strain [36, 159]. However, recent studies indicate that it is most reasonable to assume that cells respond to stress rather than strain. For example, it has been shown that fibroblasts change their gene expression when exposed to mechanical stress through fluid flow (i.e. where the strain rate governs the stress rather than the strain) [116].
potentially be investigated experimentally.

Then, in Section 6.4, we present a novel mechanochemical model of dermal wound healing. Following our analysis of the Tranquillo-Murray model of dermal wound healing [189], we do not use a viscoelastic constitutive law to describe the mechanical behaviour of the ECM. Instead, we use an elastic constitutive law and we model the changing mechanical structure of the wound using a morphoelastic evolution equation for strain.

In some ways, our model is similar to the models proposed by Cook [33] and Olsen et al. [136, 137, 138, 139, 140, 141]. Like Olsen, we include two species of fibroblast: ‘normal’ fibroblasts and myofibroblasts. Like Cook, we model the evolution of the zero stress state in response to ECM turnover. However, our model also has some distinct differences from the previous work. Most importantly, we propose a mathematical model of wound healing in which the rate of fibroblast-myofibroblast differentiation is dependent on the local stress.

Unfortunately, we have not reached the stage of running numeric simulations of our wound healing model. In Section 6.4.5, we discuss some preliminary results that we have obtained and we consider how we should proceed towards obtaining numeric results. Although the model that we construct appears promising, we cannot yet be confident that it will yield an accurate description of dermal wound healing. Further validation and analysis using numeric simulations is necessary.

Lastly, in Section 6.5 we consider other possible applications of morphoelasticity. The theory developed in the previous two chapters is very powerful and could potentially be used in a variety of different contexts. We also discuss the numeric challenges that are involved in incorporating morphoelasticity into mathematical models.
6.2 A one-dimensional Maxwell material

Firstly, let us consider a one-dimensional\(^2\) Maxwell material. A Maxwell material (also called a Maxwell fluid) is a linear viscoelastic material constructed from a series of infinitesimal Maxwell elements. Each Maxwell element consists of a Hookean spring and a Newtonian dashpot connected in series, as depicted in Figure 6.1.

\[ \frac{1}{k} \frac{D\sigma}{Dt} + \frac{1}{\eta} \sigma = \frac{D\varepsilon}{Dt}, \]

where \( t \) represents time, \( \sigma \) represents stress, \( \varepsilon \) represents the apparent strain, \( k \) is the coefficient of elasticity, and \( \eta \) is the coefficient of viscosity. In one spatial dimension, the apparent strain, \( \varepsilon \), is equivalent to the displacement gradient, \( \frac{\partial u}{\partial x} \).

In the limit as \( \eta \to \infty \), we find that the dashpots of the Maxwell elements become infinitely stiff. As a result, the Maxwell material effectively behaves as an elastic solid. Similarly, in the limit as \( k \to \infty \), we find that the Maxwell material behaves as a viscous fluid.

In rheology (see, for example, Bird et al. [20]), the degree to which a material behaves as a solid or fluid is measured using a dimensionless parameter called the Deborah number, \( De \). The Deborah number is defined as the ratio of the

\(^2\)A three-dimensional theory of Maxwell materials also exists, but it is harder to reconcile this with the strain evolution equations developed in Section 5.4. For now, we will focus on dealing with the subtleties of the one-dimensional theory.
time scale of viscous relaxation to the time scale of observation. For a Maxwell material, this takes the form $D\dot{e} = \frac{n}{kT}$, where $T$ is the time scale of observation. At large Deborah numbers, a Maxwell material behaves mainly as a solid, whereas at small Deborah numbers it behaves mainly as a fluid.\(^3\)

Another interesting property of a Maxwell material is that the zero stress deformation gradient can evolve over time.\(^4\) For a Maxwell element to be unstressed, the spring must be at its natural length. However, the dashpot in the element has no memory of its initial position and no set length corresponding to zero stress. Hence, any dashpot movement can be thought of as indicating a change to the zero stress state.

Thus, we expect that we will be able to model a Maxwell material using the theory of morphoelasticity from Chapter 5. Importantly, we note that any change to the zero stress state is associated with the movement of a dashpot. Thus, we expect that the ‘rate of growth’ of the Maxwell material is related to the viscous rate of strain. Having incorporated viscosity into the strain evolution equation, we can assume a linear constitutive law that relates the stress to the effective strain.\(^5\)

Now, we recall from equation (5.56) that the evolution of finite strain in response to growth takes the form

$$\frac{DE}{Dt} + 2E \frac{\partial v}{\partial x} = \frac{\partial v}{\partial x} - g \sqrt{1 - 2E},$$

where $E$ represents strain, $v$ represents velocity and $g$ is the rate of growth.

---

\(^3\)Interestingly, we note that the Deborah number can be decreased arbitrarily by taking $T$ to be large. On a long enough time scale, any Maxwell material can be treated as a viscous fluid. This is in accordance with the principle espoused by Bingham, the founder of rheology, who famously used a quote thought to originate with Heraclitus: ‘πάντα ῥεῖ’ (i.e. ‘all things flow’) [20].

\(^4\)This can also be seen from the fact that substituting $\sigma = 0$ into equation (6.1) yields $\frac{D\varepsilon}{Dt} = 0$. Thus, it is possible to have zero stress when the apparent strain is nonzero. In morphoelasticity, this would indicate a situation where the zero stress deformation gradient is no longer equal to the identity tensor.

\(^5\)Note that the effective strain, $E$, is different from the apparent strain, $\varepsilon$, that we used in equation (6.1). The apparent strain does not take into account the evolving zero stress state.
To simulate a Maxwell material, we will consider the case where the rate of growth is given by the viscous rate of strain.\textsuperscript{6} Given the discussion above, this would appear to be appropriate. However, at this stage we cannot be confident that it is correct to model a Maxwell material in this way.

Making the assumption that the rate of growth is equal to the viscous rate of strain, we find that
\[ g = \frac{1}{\eta} \sigma, \]
where \( \eta \) is the coefficient of viscosity. Interestingly, we note that equation (6.2) can be interpreted to mean that the zero stress state changes in response to any stress. This is consistent with the fact that a Maxwell material is a fluid and cannot sustain a shear stress.

For the elastic behaviour of the material, we will assume that stress is directly proportional to effective strain. That is,
\[ \sigma = k E, \]
where \( k \) is analogous to the coefficient of elasticity. Note that we have assumed that stress is proportional to finite strain, not infinitesimal strain. Normally, we would assume a nonlinear constitutive law at large strains but this is an unnecessary complication.

Using these assumptions, we find that equation (5.56) yields the following evolution equation for \( \sigma \):
\[ \frac{1}{k} \left( \frac{D\sigma}{Dt} + 2\sigma \frac{\partial v}{\partial x} \right) + \frac{\sigma}{\eta} \sqrt{1 - 2 \frac{\sigma}{k}} = \frac{\partial v}{\partial x}. \]
(6.3)

This can be nondimensionalised as follows. Let \( L \) be a natural length scale for the

\textsuperscript{6}We note that the rate of growth and the rate of strain both have the same dimension (namely, inverse time). Thus, there is no dimensional inconsistency with this assumption.
problem being considered,\(^7\) and similarly let \( T \) be a natural time scale. Hence, we can define dimensionless variables, \( \sigma^*, v^*, x^* \) and \( t^* \), so that

\[
\sigma = \eta \frac{T}{\sigma^*}, \quad v = \frac{L}{T} v^*, \quad x = L x^* \quad t = T t^*.
\]

Substituting into equation (6.3) and rearranging, we find that

\[
D_e \left( \frac{D\sigma^*}{Dt^*} + 2 \sigma^* \frac{\partial v^*}{\partial x^*} \right) + \sigma^* \sqrt{1 - 2 D_e \sigma^*} = \frac{\partial v^*}{\partial x^*},
\]

(6.4)

where \( D_e = \frac{\eta}{\kappa T} \) is the Deborah number.

Note that we nondimensionalised \( \sigma \) with respect to the characteristic viscous stress over the time scale being considered. We expect that this will be most appropriate when the Deborah number is small (i.e. when the observed behaviour is mainly viscous). For a large Deborah number, it would be more appropriate to nondimensionalise \( \sigma \) using the elastic stress. That is, we can define

\[
\sigma = k \sigma^{**},
\]

where \( \sigma^{**} \) is dimensionless. Using this nondimensionalisation, we find that

\[
\frac{D\sigma^{**}}{Dt^*} + 2 \sigma^{**} \frac{\partial v^*}{\partial x^*} + \frac{1}{D_e} \sigma^{**} \sqrt{1 - 2 \sigma^{**}} = \frac{\partial v^*}{\partial x^*}.
\]

(6.5)

Having constructed these dimensionless equations, there are a number of interesting cases to consider. Firstly, let us assume that the Deborah number is small. This corresponds to the case where the time scale of observation is sufficiently long for the fluid behaviour of the material to be apparent. Thus, we expect the small \( D_e \) solution of equation (6.4) to correspond to viscous flow with a small correction representing the elastic behaviour.

In order to show that this is what happens, let us formally consider the case

\(^7\)As it happens, the choice of \( L \) does not affect the nondimensionalisation. Hence, any arbitrary length scale could be used.
where $0 < De \ll 1$ and propose asymptotic expansions for $\sigma^*$ and $v^*$. That is, we let

$$\sigma^* \sim \sigma_0^* + De \sigma_1^* + \ldots,$$

and

$$v^* \sim v_0^* + De v_1^* + \ldots,$$

where $\sigma_0^*$, $v_0^*$, $\sigma_1^*$ etc. are all of order unity

Substituting into equation (6.4) and using a binomial expansion for $\sqrt{1 - 2De \sigma^*}$, we find that

$$De \left( \frac{D\sigma_0^*}{Dt^*} + \ldots + 2 (\sigma^* + \ldots) \left( \frac{\partial v_0^*}{\partial x^*} + \ldots \right) \right) + (\sigma_0^* + De \sigma_1^* + \ldots) (1 - De \sigma_0^* + \ldots) = \left( \frac{\partial v_0^*}{\partial x^*} + De \frac{\partial v_1^*}{\partial x^*} + \ldots \right).$$

Collecting the terms of order unity, this yields the result that

$$\sigma_0^* = \frac{\partial v_0^*}{\partial x^*}. \quad (6.7)$$

Given that $\sigma$ was nondimensionalised with respect to $\frac{\eta}{T}$ and $\frac{\partial v}{\partial x}$ was nondimensionalised with respect to $\frac{1}{T}$, we find that equation (6.7) is equivalent to saying that

$$\sigma \sim \eta \frac{\partial v}{\partial x} + \frac{\eta}{T} O(De).$$

That is, the stress experienced by the material is a small perturbation away from the stress predicted by Newtonian viscosity, with the size of the correction governed by the size of the Deborah number.

This is consistent with what we would expect from a Maxwell material. Thus, we have obtained some evidence that supports our assertion that choosing $g = \frac{1}{\eta} \sigma$ is an appropriate way to model a Maxwell material using morphoelasticity.
Now, consider the first order terms from equation (6.6):

\[
\frac{D\sigma_0^*}{Dt^*} + 2\sigma_0^* \frac{\partial v_0^*}{\partial x^*} + \sigma_1^* - (\sigma_0^*)^2 = \frac{\partial v_1^*}{\partial x^*}.
\] (6.8)

We could potentially use this equation to find the first correction to stress and velocity associated with the elastic behaviour of the material.

Interestingly, equation (6.7) can be used to show that

\[
\sigma_0^* \frac{\partial v_0^*}{\partial x^*} = (\sigma_0^*)^2.
\]

Thus, equation (6.8) can be rewritten as

\[
\frac{D\sigma_0^*}{Dt^*} + \sigma_0^* \frac{\partial v_0^*}{\partial x^*} + \sigma_1^* = \frac{\partial v_1^*}{\partial x^*}.
\] (6.9)

Now, consider what happens when we multiply equation (6.9) by $De$ and add it to equation (6.7). This yields

\[
De \left( \frac{D\sigma_0^*}{Dt^*} + \sigma_0^* \frac{\partial v_0^*}{\partial x^*} \right) + \sigma_0^* + De\sigma_1^* = \frac{\partial v_0^*}{\partial x^*} + De \frac{\partial v_1^*}{\partial x^*},
\]

or equivalently,

\[
De \left( \frac{D\sigma_0^*}{Dt^*} + \sigma_0^* \frac{\partial v_0^*}{\partial x^*} \right) + \sigma^* = \frac{\partial v_0^*}{\partial x^*} + O(De^2).
\]

Interestingly, this is the stress evolution equation that we would have obtained if we had used the infinitesimal strain evolution equation \(i.e.\) equation (5.55)) at the beginning of our analysis. The infinitesimal strain evolution equation is valid in the case where the effective strain is small. Furthermore, we recall that a small effective strain implies that most of the observed deformation is associated with ‘growth’ (in this case, viscous flow). Thus, it makes sense that the infinitesimal

---

8By this, we mean the strain associated with elastic stress. This is in contrast to the apparent strain, $\varepsilon$, described earlier.
strain evolution equation that we would write down for a Maxwell material is associated with a small Deborah number.

This suggests one point where we may encounter difficulties if we were to try to extend this work to a three-dimensional model. In three dimensions, we were only able to obtain an evolution equation for infinitesimal strain, not finite strain. Thus, any three-dimensional analysis would be limited to the case where the Deborah number is small. It may not be possible to use the morphoelastic theory that we have developed to construct a three-dimensional model of a Maxwell material that is valid at all Deborah numbers.

Having considered the case where $De$ is small, let us now move on to the case where $De$ is large. That is, we are dealing with a time scale that is sufficiently short for the material being considered to behave elastically, with a small modification for viscous effects. In this case, it is more appropriate for us to use the nondimensionalisation in equation (6.5). Thus, let us propose asymptotic series for $\sigma^{**}$ and $v^*$ as follows:

$$\sigma^{**} \sim \bar{\sigma}_0 + \frac{1}{De} \bar{\sigma}_1 + \ldots;$$

$$v^* \sim \bar{v}_0 + \frac{1}{De} \bar{v}_1 + \ldots.$$

Substituting into equation (6.5) and considering only the leading order terms, we find that

$$\frac{D\bar{\sigma}_0}{Dt^*} + 2 \bar{\sigma}_0 \frac{\partial \bar{u}_0}{\partial x^*} = \frac{\partial \bar{v}_0}{\partial x^*}. \quad (6.10)$$

Importantly, we note that this is consistent with dimensionless stress\(^9\) taking the form

$$\bar{\sigma}_0 = \frac{\partial \bar{u}_0}{\partial x^*} - \frac{1}{2} \left( \frac{\partial \bar{u}_0}{\partial x^*} \right)^2,$$

where $\bar{u}_0$ is the leading order approximation of displacement.\(^{10}\) That is, equation

\(^9\)Since we have nondimensionalised stress with respect to the coefficient of elasticity in this case, the dimensionless stress is effectively identical to the elastic strain.

\(^{10}\)In order to show that this is the case, we need to use the fact that $\frac{D}{Dt} \left( \frac{\partial \bar{u}_0}{\partial x^*} \right) = \frac{\partial \bar{v}_0}{\partial x^*} \left( 1 - \frac{\partial \bar{u}_0}{\partial x^*} \right)$. This is analogous to the result for $\frac{D}{Dt} \left( \frac{\partial \bar{v}_0}{\partial x^*} \right)$ that we described in Section 5.4.
Applications of zero stress state theory and morphoelasticity

(6.10) implies that
\[ \sigma \sim k \tilde{E} + k O(De^{-1}), \]
where \( \tilde{E} = \frac{\partial u}{\partial x} - \frac{1}{2} \left( \frac{\partial u}{\partial x} \right)^2 \) is the classical finite strain (see, for example, Roberts [158]) and \( De \) is large.

Thus, taking a large Deborah number leads us to construct an asymptotic series where the leading order term corresponds to classical elasticity. In a practical problem, we could use higher order terms from equation (6.5) to obtain the viscous corrections. However, we will not consider this here.

Although we have shown that our formulation yields viscous behaviour as \( De \to 0 \) and elastic behaviour as \( De \to \infty \), we have not yet demonstrated any clear relationship between our equations and the classical Maxwell constitutive law. Quite clearly, equation (6.3) is different from equation (6.1) and the two will not always yield the same results. However, let us consider what happens when \( \sigma \) and \( \frac{\partial v}{\partial x} \) are both small.

Formally, we are interested in the case where \( De \sim O(1) \) and \( \sigma^* \) and \( \frac{\partial v}{\partial x^*} \) are both of order \( \delta \) where \( 0 < \delta \ll 1 \). As before, we propose asymptotic series for \( \sigma^* \) and \( \frac{\partial v}{\partial x^*} \). This time, our series are in powers of \( \delta \) and contain no zeroth order term:

\[ \sigma^* \sim \delta \tilde{\sigma}_1 + \delta^2 \tilde{\sigma}_2 + \ldots; \]
\[ v^* \sim \delta \tilde{v}_1 + \delta^2 \tilde{v}_2 + \ldots. \]

Substituting into equation (6.4) and collecting the terms of order \( \delta \), we find that

\[ De \frac{D\sigma_1}{Dt^*} + \dot{\sigma}_1 = \frac{\partial \tilde{v}_1}{\partial x^*}. \]

This is very similar to the result that we would obtain if we nondimensionalised equation (6.1). The only significant difference is that we have \( \frac{\partial \tilde{v}_1}{\partial x} \) where the Maxwell formulation has \( \frac{D\varepsilon}{Dt} \). Furthermore, we recall that these two are very close when \( \frac{\partial u}{\partial x} \) is small. Thus, it would appear that the classical Maxwell constitutive
law is a linear approximation to the full model given in equation (6.3).

In order to understand why this is the case, we note that there are three significant differences between (6.1) and (6.3):

- The right-hand side of equation (6.3) is written as \( \frac{\partial v}{\partial x} \), not \( \frac{D\varepsilon}{Dt} \),
- the elastic contribution to the strain rate in equation (6.3) contains an additional term of the form \( \frac{2\sigma}{k} \frac{\partial v}{\partial x} \), and
- the viscous contribution to the strain rate in equation (6.3) is multiplied by \( \left(1 - \frac{2\sigma}{k}\right)^{\frac{1}{2}} \).

By careful consideration of a Maxwell element, it is clear that \( \frac{\partial v}{\partial x} \) is to be preferred over \( \frac{D\varepsilon}{Dt} \). Similarly, we note that the addition of \( \frac{2\sigma}{k} \frac{\partial v}{\partial x} \) is necessary in order to obtain an expression for the velocity gradient associated with elasticity that is valid at large deformations. Thus, we find that the first two differences listed above can be easily explained and both lead to the conclusion that (6.3) is more valid than (6.1) for large displacement gradients.

However, the third difference is more interesting. Consider, for example, the response of equation (6.3) when a constant and homogeneous stress is applied (i.e. a creep test). After an initial jump in the displacement gradient, we find that

\[
\frac{\partial v}{\partial x} = \frac{\sigma}{\eta \sqrt{1 - 2\frac{\sigma}{k}}}. \tag{6.11}
\]

This behaviour is different from what we would predict. If we consider a Maxwell material as a series of springs and dashpots, it would seem sensible for the time-dependent creep response to depend entirely on the behaviour of the dashpots.

\[11\] If the particle \( P \) with position \( x(t) \) is at the left-hand end of a Maxwell element and the particle \( Q \) with position \( x(t) + \Delta x(t) \) is at the right-hand end of a Maxwell element, we find that

\[
\Delta x(t) = \Delta x_{\text{spring}} + \Delta x_{\text{dashpot}}.
\]

By taking the material derivative of this expression with respect to time and considering the limit as \( \Delta x \) goes to zero, we can conclude that \( \frac{\partial v}{\partial x} \) should be expressed as the sum of an elastic component and a viscous component.
That is, we would expect that \( \frac{\partial v}{\partial x} = \frac{\sigma}{\eta} \).

Why does our morphoelastic model predict an additional dependence on the elastic behaviour of the material?

To answer this question, we note that our morphoelastic model depended on the assumption that the strain evolves according to equation (5.56), but the relationship between stress and strain does not change over time. That is, the instantaneous elastic response of a small segment of the deformed material is identical to the instantaneous elastic response of a segment of the undeformed material of equal size.

However, this is different from the behaviour of a series of springs and dashpots. Consider a single Maxwell element that initially consists of a spring and dashpot of equal length. If we were to apply a tensile stress over a long period of time and then release it, we would find that the spring returns to its original length but the dashpot remains at its extended length. Thus, the length of the dashpot in the deformed Maxwell element is now considerably greater than the length of the spring.

This means that instantaneously increasing the length of the deformed Maxwell element by a factor of \( \alpha \) requires more force than instantaneously increasing the length of the undeformed element by the same amount.\(^{12}\) Thus, we find that the instantaneous elastic coefficient of the entire element (spring and dashpot) has changed.

---

\(^{12}\)For example, consider the case where the deformed Maxwell element is two-thirds dashpot and one-third spring, while the undeformed element consists of a spring and a dashpot of equal length. If we wish to instantaneously double the length of the undeformed element, the spring needs to triple in length. This is because the dashpot takes time to change its length and can be treated as being rigid to instantaneous deformations.

However, in order to double the length of the deformed element, it is necessary for the spring to quadruple in length. Since the force experienced by a Hookean spring is proportional to extension, it follows that it takes a greater force to double the length of the deformed element than the undeformed element.
For a series of springs and dashpots, this means that a large viscous deformation can change the stiffness of the material. This is a direct consequence of internal changes to the ratio of spring-length to dashpot-length. In contrast, our model of a Maxwell material always has the same instantaneous elastic properties. In effect, new springs are laid down as the dashpots extend in order to maintain the coefficient of elasticity.

If a tensile stress is applied over a period of time, these new springs will form and then extend. Thus, it is reasonable for a morphoelastic model of a Maxwell material to predict greater extension than a spring-and-dashpot model. This is exactly what we see in equation (6.11); our constitutive law predicts a greater velocity gradient in the creep test than the classical law does.\(^\text{13}\)

In conclusion, we find that the morphoelasticity theory developed in Chapter 5 can be used to construct a model of a Maxwell material. Interestingly, the morphoelastic model would appear to be more general than the classical model; unlike Maxwell’s constitutive law, our constitutive law remains mathematically valid at large deformation gradients and large velocity gradients.

Of course, this does not necessarily imply that our model is more useful in practical situations. A material that behaves as a Maxwell material at small deformations may behave very differently at large deformations. In such a situation, it is probable that neither the classical Maxwell model nor our morphoelastic model would be able to accurately describe the physically observed behaviour.

Lastly, it is interesting to note that morphoelasticity theory was developed to explain the behaviour of growing solids. However, a Maxwell material is a viscoelastic fluid, not a viscoelastic solid. Because the growth term is proportional to stress, we find that our morphoelastic ‘solid’ is actually able to behave as a

\(^{13}\text{We also note that equation (6.11) predicts that the velocity is singular when }\frac{\sigma}{E} = \frac{1}{2}. \text{ Since } \frac{\sigma}{E} = E \text{ is the elastic strain, we note that this will never be a problem. Given the definition of Eulerian finite strain from equation (4.33), it is possible to show that } Y \to 0 \text{ (and hence } Z \to 0) \text{ as } E \to \frac{1}{2}. \text{ As a result, } E = \frac{1}{2} \text{ corresponds to the unphysical situation where the desired distance between two neighbouring particles is zero (i.e. they would be superimposed in the zero stress state).}\)
fluid. Despite the fact that morphoelasticity theory was derived using principles from solid mechanics, we have shown that it has more general application within the field of continuum mechanics.

### 6.3 Modelling growth

#### 6.3.1 Dependence of the growth tensor on stress

As we discussed in Section 2.4, morphoelasticity has always been intimately associated with models of biological growth and remodelling. In their original paper, Rodriguez et al. \cite{Rodriguez2001} considered a problem where a bone specimen is remodelled to accommodate a stress. In more recent times, Ambrosi and coworkers \cite{Ambrosi2008, Ambrosi2009, Goriely2010} and Goriely and coworkers \cite{Goriely2011, Goriely2012, Goriely2013} have made significant progress with morphoelastic models of growth.

In this section, we will demonstrate how the strain evolution equation developed in Section 5.4.2 can also be used to model tissue growth. As a first step, we need to construct a plausible expression for the growth tensor, $G$, that is consistent with our physical expectations of growth. Hypothetically, the equations given in the previous chapter (e.g. equation (5.44)) could be used to model any process that involves internal remodelling of the zero stress state. However, these equations are very complicated and many of them involve unspecified functional relations. By making some reasonable assumptions about biological growth, we find that we are able to obtain equations that are much simpler, but still have the flexibility to model a variety of different forms of growth.

Firstly, let us assume that the net rate of volumetric growth is known at every point in the growing tissue. The net rate of growth, which we represent by $g(x, t)$, may be dependent on other model variables.\footnote{For example, most models of tumour growth (including the classical model of tumour}
this functional relationship is known.

As described in the previous chapter, the trace of the growth tensor is equal to the net rate of volumetric growth. That is, we find that

$$\text{tr}(G) = g(x, t).$$

Clearly, this is not sufficient to completely specify $G$ in a three-dimensional problem; we need to make some additional assumptions about the directional bias of growth.

One possibility would be to assume that growth is isotropic. In this case, we would find that all of the eigenvalues of $G$ are equal and that every vector is an eigenvector. Thus,

$$G = \frac{g(x, t)}{3} I.$$

This is simple, but it is not physically realistic. If a cell is highly compressed in one direction and less compressed in all other directions, it is reasonable to expect that it will prefer not to divide in the direction of most compression. Assuming isotropic growth can lead to unreasonably large stresses.\(^{15}\)

Thus, let us assume that growth is directed so that the principal directions of growth are parallel to the principal directions of stress. Importantly, this means that we are effectively assuming that stress is the only factor that influences the direction of growth. If we were modelling a biological process where growth is directed to follow a chemical gradient (or some other nonmechanical cue), this assumption would be invalid.\(^{16}\)

\(^{15}\)For example, Jones et al. [97] assumed the isotropic growth of a tumour spheroid and they found that the internal stresses of the spheroid diverged to infinity over time. Araujo and McElwain [9] modified the Jones et al. model to consider the case where the direction of growth is dependent on stress. By doing this, they were able to prevent unbounded stresses from arising.

\(^{16}\)For example, a tumour may grow preferentially towards a source of oxygen and nutrients.
Given that the principal directions of growth are parallel to the principal directions of stress, we find that the eigenvectors of the growth tensor, $G$, will be identical to the eigenvectors of the Cauchy stress tensor, $\sigma$. Since the Cauchy stress tensor is symmetric, we find that it can be expressed in the form

$$\sigma = P D_\sigma P^T,$$

where $P$ is the orthogonal tensor of eigenvectors and $D_\sigma$ is the diagonal tensor of eigenvalues. For notational convenience, we will express the eigenvalues of $\sigma$ as $\lambda_\sigma i$ and the corresponding eigenvectors as $p_i; i = 1, 2, 3$.

Since the eigenvectors of $G$ are equal to the eigenvectors of $\sigma$, we find that

$$G = P D_G P^T,$$

where $D_G$ is a diagonal tensor containing the eigenvalues of $G$. As before, we denote these eigenvalues as $\lambda_G i, i = 1, 2, 3$.

In order to fully determine $G$, we need to construct a relationship between the eigenvalues of $G$ and the eigenvalues of $\sigma$. In their model of stress-dependent growth, Rodriguez et al. [159] proposed that the principal stresses (i.e. the eigenvalues of $\sigma$) have biologically preferred values. Growth was then biased according to the difference between the current value of the principal stress and the preferred value of the principal stress.

Following the example of Rodriguez et al., let us define the preferred stress tensor, $\sigma^*$, to represent the biologically preferred stress state at any point. In general, it would seem physically reasonable to assume that the biologically preferred stress is isotropic. In this case, we would find that

$$\sigma^* = s^* I,$$
where $s^*$ is a scalar. As we will see, a preferred stress tensor of this form leads to significant simplifications later in our analysis.

However, for now let us assume that the preferred stress tensor is not isotropic. Many biological tissues (e.g. bone, arterial walls etc.) are anisotropically stressed in physiological situations [96]. Thus, it is not unreasonable to consider an anisotropic preferred stress. However, in order to continue our analysis, we must assume that the eigenvectors of $\sigma^*$ are parallel to the eigenvectors of $\sigma$. If this is not the case, we would probably encounter a situation where the principal directions of growth are not the same as the principal directions of stress.

Thus, we will assume that

$$\sigma^* = P D^* P^T,$$

where $D^* = \text{diag} [\lambda_{\sigma 1}^*, \lambda_{\sigma 2}^*, \lambda_{\sigma 3}^*]$ is the diagonal tensor of eigenvalues. If $\sigma^*$ is isotropic, we would find that $\lambda_{\sigma 1}^* = \lambda_{\sigma 2}^* = \lambda_{\sigma 3}^*$.

Now, if $\lambda_{\sigma i} < \lambda_{\sigma i}^*$ we find that the tissue is compressed in the direction of $p_i$ relative to its biologically preferred state. Conversely, if $\lambda_{\sigma i} > \lambda_{\sigma i}^*$ we find that the tissue is extended in the direction of $p_i$ relative to its biologically preferred state. Thus, we would expect growth to be greatest in directions where $\lambda_{\sigma i} > \lambda_{\sigma i}^*$ and smallest in directions where $\lambda_{\sigma i} < \lambda_{\sigma i}^*$.

We recall that $g(x, t) = \text{tr} (G) = \lambda_{G1} + \lambda_{G2} + \lambda_{G3}$. Thus, we seek an expression for $\lambda_{Gi}$ with the following two properties:

- The mean $\lambda_{Gi}$ is $\frac{1}{3} g(x, t)$.
- The values of $\lambda_{Gi}$ are biased so that an above average value of $\lambda_{\sigma i} - \lambda_{\sigma i}^*$ implies an above average value of $\lambda_{Gi}$.

There are many ways that we could satisfy these conditions. However, the sim-
ples approach is to take $\lambda_{Gi}$ to be given by the equation

$$\lambda_{Gi} = \frac{g(x, t)}{3} + \kappa \left( \lambda_{\sigma_i} - \lambda_{\sigma^*_i} - \frac{1}{3} \sum_{j=1}^{3} (\lambda_{\sigma_j} - \lambda_{\sigma^*_j}) \right).$$ (6.12)

Equation (6.12) includes a parameter, $\kappa$, that represents the sensitivity of growth to stress. We will call $\kappa$ the stress sensitivity and we note that it has units of inverse stress multiplied by inverse time. As with $g$, we note that $\kappa$ may vary through space and time. Hypothetically, cells might respond to some chemical factor or other signal by becoming more or less responsive to mechanical signals. Generally, however, we will assume that $\kappa$ is constant.

If $\kappa = 0$, we find that the rate of growth is completely independent of the stress experienced (i.e., we recover the equation for isotropic growth). As $\kappa$ becomes large and positive, the growth becomes increasingly sensitive to stress. A negative value of $\kappa$ corresponds to the unphysical (and generally unstable) situation where growth is preferred in the directions of relative compression. We will explore the meaning of the parameter $\kappa$ in greater detail later in this section.

Now, equation (6.12) implies that

$$D_G = \frac{g(x, t)}{3} I + \kappa \left( (D_\sigma - D^{*}) + \frac{1}{3} \text{tr} (D_\sigma - D^{*}) I \right).$$

Thus, we can use the fact that $G = PD_G P^T$ to show that

$$G = \frac{g(x, t)}{3} I + \kappa \left( (\sigma - \sigma^*) - \frac{\text{tr}(\sigma - \sigma^*)}{3} I \right).$$ (6.13)

If the biologically preferred stress is isotropic, we note that

$$\sigma^* = \frac{\text{tr}(\sigma^*)}{3} I.$$
In this case, equation (6.13) becomes
\[
G = \frac{g(x, t)}{3} I + \kappa \left( \sigma - \frac{\text{tr} (\sigma)}{3} I \right).
\]  
(6.14)

In all of the work that follows, we will use equation (6.14) in preference to equation (6.13). However, it is interesting to note that we can use the theory described above to deal with the case where the biologically preferred stress is anisotropic.

Equation (6.14) gives us an expression for \( G \) that we can use in a complete model of growth. Thus, consider what happens when we substitute (6.14) into equation (5.64). This yields
\[
\frac{D e}{D t} + \text{tr} (e) \text{ sym} (L) + e \text{ skew} (L) - \text{skw} (L) e
= \text{sym} (L) - \left( \frac{g(x, t)}{3} I + \kappa \left( \sigma - \frac{\text{tr} (\sigma)}{3} I \right) \right),
\]  
(6.15)

where \( e \) is the strain and \( L \) is the velocity gradient. Given an appropriate expression for \( g(x, t) \), a constitutive law relating stress to strain and a force balance equation, we can use equation (6.15) to construct a well-posed model of growth.

It should be noted that equation (6.15) is only valid for small strains. Intuitively, we expect that small strains will generally arise when growth is highly responsive to stress (i.e. where \( \kappa \) is large). However, it is interesting to note that this is not necessarily the case. If a growing body is completely constrained (i.e. no displacement is permitted on the boundaries), we will find that significant strains and stresses will develop, even if \( \kappa \) is large. In cases where significant strains develop, it becomes necessary to use equation (5.44) or equation (5.74) instead.

Equation (6.15) is very general and, as such, it is difficult to interpret. In order to understand it better, it is constructive for us to consider a number of simplifying cases. In Section 6.3.2, for example, we make the assumption of spherical symmetry to construct a general model of a growing tumour spheroid. In this case,
we find that (6.15) can be simplified to yield two scalar equations: one for the evolution of radial strain and another for the evolution of circumferential strain.

Another interesting problem to consider is the case where \( \kappa \) is taken to be infinitely large. In this case, we find that equation (6.15) infers that

\[
\sigma - \frac{\text{tr}(\sigma)}{3} I = 0.
\]

This will be satisfied if and only if \( \sigma \) takes the form \( \sigma = s I \), where \( s \) is some constant. That is, \( \kappa \to \infty \) corresponds to the situation where our growing solid can only tolerate an isotropic stress (like a hydrostatic pressure). If there is any shearing, the solid will adapt immediately by growing in some directions and shrinking in others. Thus, we find that our material effectively behaves as a growing inviscid fluid.

Furthermore, it would appear to be possible to use equation (6.15) to construct a model of a growing Maxwell material.\(^{18}\) By choosing \( \kappa = \frac{1}{\eta} \) and \( g(x, t) = \frac{1}{\eta} \text{tr}(\sigma) + g^*(x, t) \), we are able to obtain equations that are very similar to what we would expect for a growing Maxwell material. However, there are aspects of this approach that require further thought. In particular, we need to be careful with the constitutive law that we use to relate stress with the elastic strain. As it stands, the full analysis of a three-dimensional Maxwell model is beyond the scope of the current work.

\(^{18}\)For an example of a mathematical model that incorporates a growing Maxwell material, see MacArthur and Please [114].
6.3.2 A model of spherically-symmetric growth

Consider the system that we obtain when we combine equation (6.14) with the model of growth given in equations (5.65) to (5.68):

\[ G = \frac{g(x, t)}{3} I + \kappa \left( \sigma - \frac{\text{tr}(\sigma)}{3} I \right), \]

\[ \frac{D e}{D t} + \text{tr}(e) \text{sym}(L) + e \text{skw}(L) - \text{skw}(L) e = \text{sym}(L) - G, \]

\[ \sigma = \lambda \text{tr}(e) I + 2 \mu e, \]

\[ \text{div } \sigma = 0, \]

\[ L = \frac{\partial v}{\partial x}. \]

If \( g(x, t) \) is specified and appropriate boundary conditions are given, we can use this system to model the stress dependent growth of a linear elastic, isotropic solid. Thus, one possible application of these equations is the modelling of tumour spheroid growth.

Tumour spheroids are small colonies of tumour cells that are limited in size by their lack of vasculature. Despite some arguments about their usefulness, tumour spheroids have been used extensively in experiments as \textit{in vitro} equivalents of solid tumours [160]. Furthermore, the growth of tumour spheroids has received a significant amount of attention from mathematical modellers. For more details on the current state of mathematical research in this area, see the recent reviews by Araujo and McElwain [8] and Roose \textit{et al.} [160].

Of particular interest to us, several models have been developed to describe the internal stresses experienced by a growing tumour. Interestingly, these models approach spheroid growth in very different ways. For example, Jones \textit{et al.} [97] and MacArthur and Please [114] (amongst others) use a modified Maxwell model in which growth is incorporated into the representation of the velocity gradient. In contrast, Ambrosi and Mollica [5] use energy arguments based on their
consideration of the evolving zero stress state. More recently, Kim et al. [100] have proposed a hybrid model in which both continuum and discrete mechanical elements are used to model different parts of the growing spheroid.

Despite our use of an evolving zero stress state, the approach that we take is closer to the approach used by Jones et al. and MacArthur and Please than the approach used by Ambrosi and Mollica [5]. As we can see clearly from equation (5.64), our treatment of growth effectively involves modifying the velocity gradient in the equation for the evolution of elastic strain. Furthermore, we note from earlier that it is possible to obtain a Maxwell model by making appropriate assumptions about $g$ and $\kappa$. Thus, very little modification of the equations above would be needed in order to obtain a model that is equivalent to the models of Jones et al. [97] or MacArthur and Please [114].

However, there are two important differences between our work and the earlier work of Jones et al. and MacArthur and Please. Firstly, we have used the infinitesimal strain definition that we developed in Section 4.4. This means that we can be confident that our model is valid at large displacement gradients, as long as the strains associated with elastic stresses are small. In contrast, most previous mechanical models have relied on the assumptions of classical linear elasticity. As such, they may encounter difficulties when attempting to model the large deformations involved in growth.

Furthermore, our model incorporates stress-dependence in the growth term. In their original model, Jones et al. [97] assumed that growth was isotropic. This is equivalent to assuming that $\kappa = 0$ in the model that we give above. Unfortunately, Jones et al. found that the assumption of isotropic growth led to the development of infinite stresses within the growing tumour. This led Araujo and McElwain [9] to modify the Jones et al. model to include a term that represents the dependence of the growth rate on stress.

In their model, Araujo and McElwain [9] introduce stress dependent terms $\eta_r$ and $\zeta_r$ that reflect the stress-dependent bias in proliferation and apoptosis re-
pectively. These functions are defined so that it is impossible for the tumour to be growing in one direction and shrinking in another direction at a single point. In contrast, we propose a linear relationship between $G$ and $\sigma$ and we use a single parameter, $\kappa$, to represent the stress-dependence of growth. For sufficiently large values of $\kappa$, this means that our model may predict simultaneous radial growth and circumferential shrinkage (or vice versa) at a single point. Experimental work is needed to determine whether or not this is physically reasonable. Possibly, we need to adopt a dependence of $\kappa$ on $\sigma$ that is analogous to the approach taken by Araujo and McElwain.

Now, it is possible to use the assumption of spherical symmetry to significantly simplify the tensor equations given above. Following previous researchers [8, 26, 71, 114], we find that the stress tensor has only two independent components: the radial stress, $\sigma_r$, and the circumferential stress, $\sigma_\theta$. In spherical coordinates, we can write the stress tensor in the form

$$\sigma = \text{diag} (\sigma_r, \sigma_\theta, \sigma_\theta).$$

Using equation (6.14), this implies that $G$ can also be separated into a radial component and a circumferential component:

$$G = \text{diag} (g_r, g_\theta, g_\theta),$$

where

$$g_r = \frac{g(r, t)}{3} + \frac{2}{3} \kappa (\sigma_r - \sigma_\theta),$$

and

$$g_\theta = \frac{g(r, t)}{3} + \frac{2}{3} \kappa (\sigma_\theta - \sigma_r).$$

As above, $g(r, t)$ represents the net rate of volumetric growth and $\kappa$ is the stress sensitivity of growth. We assume that $g(r, t)$ is a known function and that $\kappa$ is a constant.

Similarly to the stress tensor and growth tensor, we find that the strain tensor
can be separated into a radial strain and a circumferential strain. That is,

\[ e = \text{diag}(e_r, e_\theta, e_\theta). \]

Lastly, we note from classical continuum mechanics (see, for example, Spencer [176]) that the assumption of spherical symmetry yields a velocity gradient of the form

\[ L = \text{diag}\left(\frac{\partial v}{\partial r}, \frac{v}{r}, \frac{v}{r}\right), \]

where \( v(r, t) \) is the radial velocity.

Given these results, we find that the assumption of spherical symmetry allows us to rewrite equations (5.65), (5.66), (5.67), (5.68) and (6.14) as the following system of scalar equations:

\[
\begin{align*}
g_r &= \frac{g(r, t)}{3} + \frac{2\kappa}{3} (\sigma_r - \sigma_\theta), \\
g_\theta &= \frac{g(r, t)}{3} + \frac{\kappa}{3} (\sigma_\theta - \sigma_r), \\
\frac{De_r}{Dt} + (e_r + 2e_\theta) \frac{\partial v}{\partial r} &= \frac{\partial v}{\partial r} - g_r, \\
\frac{De_\theta}{Dt} + (e_r + 2e_\theta) \frac{v}{r} &= \frac{v}{r} - g_\theta, \\
\sigma_r &= \lambda(e_r + 2e_\theta) + 2\mu e_r, \\
\sigma_\theta &= \lambda(e_r + 2e_\theta) + 2\mu e_\theta, \\
\frac{\partial \sigma_r}{\partial r} + \frac{2}{r} (\sigma_r - \sigma_\theta) &= 0.
\end{align*}
\]

In order to construct a well-posed problem, we need initial conditions for \( e_r \) and \( e_\theta \) as well as boundary conditions for \( v \) and \( \sigma_r \). These are relatively simple to obtain. For example, we note that symmetry requires that \( v(0, t) \equiv 0 \). Any other result would correspond to a tear or a superposition at the centre of the tumour.
Now, consider the boundary condition on $\sigma_r$. If we are considering a spheroid that is submerged in a fluid medium, it is appropriate to assume that there is a constant hydrostatic stress acting on the surface of the spheroid. That is, we find that

$$\sigma_r(R(t), t) \equiv -p,$$

where $R(t)$ is the position of the outer boundary and $p$ is the pressure in the surrounding medium. In order to simplify the analysis that follows, we make the assumption that this pressure is negligible compared to the internal stresses experienced by the spheroid.\(^{19}\) Thus, we will instead use the condition

$$\sigma_r(R(t), t) \equiv 0. \quad (6.23)$$

We also need to construct an equation to specify $R(t)$. We note that the velocity of the boundary is equal to the velocity of a particle on the boundary. Thus,

$$\frac{dR}{dt} = v(R(t), t). \quad (6.24)$$

This equation in turn requires an initial condition. Thus, we specify that $R(0) = R_0$, where $R_0$ is the initial radius of the spheroid.

Lastly, we need initial conditions on $e_r$ and $e_\theta$. One possibility is to assume that the tumour is at a zero stress state when $t = 0$. Given the definition of strain, this would imply that

$$e_r(r, 0) \equiv e_\theta(r, 0) \equiv 0, \quad 0 \leq r \leq R_0.$$

In order to proceed further, it is useful for us to rearrange the system of equations given above. For example, consider what happens when we double equation (6.19)\(^{19}\) Note that we are already making assumptions that imply that the internal stresses of the tumour are small. Hence, our use of condition (6.23) implies that the external pressure is extremely small. This may not be a valid assumption but we will use it because of the simplifications that it entails. It may be necessary to perform an asymptotic expansion in orders of the nondimensionalised external pressure to investigate the effects of the boundary condition.
and add it to equation (6.18). This yields
\[
\frac{D}{Dt} (e_r + 2e_\theta) + (e_r + 2e_\theta) \left( \frac{\partial v}{\partial r} + \frac{2v}{r} \right) = \frac{\partial v}{\partial r} + \frac{2v}{r} - (g_r + 2g_\theta).
\]

From equations (6.16) and (6.17), it is clear that \( g_r + 2g_\theta = g(r, t) \). Thus, if we let \( e_T = e_r + 2e_\theta \), we find that
\[
\frac{De_T}{Dt} + e_T \left( \frac{\partial v}{\partial r} + \frac{2v}{r} \right) = \frac{\partial v}{\partial r} + \frac{2v}{r} - g(r, t).
\]

The variable \( e_T \) is equal to the trace of the strain tensor and it is effectively a measure of the total strain. For mathematical convenience, we will consider the evolution of \( e_\theta \) and \( e_T \) instead of the evolution of \( e_\theta \) and \( e_r \). That is, we will use the substitution
\[
e_r = e_T - 2e_\theta,
\]

to replace every occurrence of \( e_r \) with an equivalent expression involving \( e_T \) and \( e_\theta \).

Similarly, it is convenient for us to follow the example of Araujo and McElwain [9] and define \( \beta = \sigma_r - \sigma_\theta \). That is, \( \beta \) represents the difference between the radial stress and the circumferential stress. As with the strain, we will make the substitution
\[
\sigma_\theta = \sigma_r - \beta,
\]

and thus eliminate \( \sigma_\theta \) from our system.\(^{20}\) For example, we note that subtracting equation (6.21) from equation (6.20) yields
\[
\sigma_r - \sigma_\theta = 2\mu (e_r - e_\theta),
\]

\(^{20}\) It may appear strange that we have eliminated the radial strain but not the radial stress. There is no physical motivation for these substitutions; our only reason for eliminating \( e_r \) and \( \sigma_\theta \) is to simplify later parts of our mathematical analysis. If there is a particular interest in the radial or circumferential behaviour of the spheroid, it is possible to make different substitutions.
and hence,

\[ \beta = 2\mu (e_T - 3e_\theta). \]

Using these results, and substituting the expression for \( g_\theta \) from (6.17) into equation (6.19), we obtain the following system:

\[
\begin{align*}
\frac{D e_T}{D t} + e_T \left( \frac{\partial v}{\partial r} + \frac{2v}{r} \right) &= \frac{\partial v}{\partial r} + \frac{2v}{r} - g(r, t), \\
\frac{D e_\theta}{D t} + e_T \frac{v}{r} &= \frac{v}{r} - \frac{g(r, t)}{3} + \frac{\kappa}{3} \beta, \\
\sigma_r &= (\lambda + 2\mu) e_T - 4\mu e_\theta, \\
\beta &= 2\mu (e_T - 3e_\theta), \\
\frac{\partial \sigma_r}{\partial r} + \frac{2}{r} \beta &= 0.
\end{align*}
\]

We now wish to nondimensionalise this system. In order to do this, we need to find a length scale, a time scale and a stress scale that are characteristic of the problem being considered. Equations (6.27) and (6.28) suggest that it is appropriate to scale stress according to the shear modulus, \( \mu \). However, the model given above does not have obvious scalings for space and time.\(^{21}\) Thus, we will use \( L \) to denote the length scale and \( T \) to denote the time scale, but we will leave \( L \) and \( T \) unspecified.

Noting that \( e_T \) and \( e_\theta \) are already dimensionless, we thus propose the following

\(^{21}\)In a full tumour spheroid model, we could scale length with respect to the final radius of the spheroid and time with respect to the maximum rate of proliferation. However, in order to find these values, we would need to know the form that \( g(r, t) \) takes. Since \( g(r, t) \) is unspecified in the current model, it seems appropriate for us to make the spatial and temporal scalings as general as possible.
nondimensionalisations:

\[ r = r^* L, \quad t = t^* T, \quad g_r = g_r^* \frac{1}{T}, \quad g_\theta = g_\theta^* \frac{1}{T}, \]
\[ g = g^* \frac{1}{T}, \quad v = v^* \frac{L}{T}, \quad \sigma_r = \sigma_r^* \mu, \quad \beta = \beta^* \mu. \]

Using these scalings and omitting stars from dimensionless variables, we obtain the following dimensionless system:

\[ \frac{De_T}{Dt} + e_T \left( \frac{\partial v}{\partial r} + \frac{2v}{r} \right) = \frac{\partial v}{\partial r} + \frac{2v}{r} - g(r, t), \quad (6.30) \]
\[ \frac{De_\theta}{Dt} + e_T \frac{v}{r} = \frac{v}{r} - \frac{g(r, t)}{3} + \frac{k}{3} \beta, \quad (6.31) \]
\[ \sigma_r = (l + 2)e_T - 4e_\theta, \quad (6.32) \]
\[ \beta = 2(e_T - 3e_\theta), \quad (6.33) \]
\[ \frac{\partial \sigma_r}{\partial r} + \frac{2}{r} \beta = 0; \quad (6.34) \]

where \( k = \kappa \mu T \) and \( l = \frac{\lambda}{\mu} \). The parameter \( k \) is effectively a dimensionless measure of the sensitivity of growth to stress, while \( l \) is related to the elastic properties of the spheroid.

In fact, we recall from classical solid mechanics (see, for example, Malvern [118]) that the Lamé coefficients can be related to Young’s modulus, \( E \) and Poisson’s ratio, \( \nu \), as follows:

\[ \lambda = \frac{E \nu}{(1 + \nu)(1 - 2\nu)}, \quad \mu = \frac{E}{2(1 + \nu)}. \]

Thus,

\[ l = \frac{\lambda}{\mu} = \frac{2\nu}{1 - 2\nu}, \]

and we find that \( l \) is dependent only on Poisson’s ratio.

Interestingly, we note that this definition of \( l \) could potentially lead to problems
Applications of zero stress state theory and morphoelasticity

with the asymptotic analysis that follows. Since we expect the tumour spheroid to be almost incompressible, we find that $\nu$ should be close to $\frac{1}{2}$. This would imply that $l$ is large and not of order unity (as we will assume). Although this does not affect the validity of our results,\textsuperscript{22} it is important to note that we must always be careful when defining dimensionless parameters.

Now, we are particularly interested in the case where the rate of growth is very sensitive to the stress. In this situation, we expect to find that the stresses and strains are all small, allowing us to justify our use of linear elasticity. More importantly, small strains would imply that it is appropriate to use equation (5.64) to describe the evolution of strain. If the strains were to become large, we would find that we would need to rethink our approach to strain evolution as well as our constitutive assumptions for the stress-strain relationship.\textsuperscript{23}

Thus, it is appropriate for us to assume that $k$, the dimensionless stress sensitivity, is large. Formally, we will investigate the case where $k \gg 1$ but $l$ and $g(r, t)$ are both of order unity. Hence, we propose asymptotic expansions for our dependent variables in inverse powers of $k$. That is,

$$e_T(r, t; \kappa^*) \sim e_{T0}(r, t) + k^{-1} e_{T1}(r, t) + \ldots,$$

$$e_\theta(r, t; \kappa^*) \sim e_{\theta0}(r, t) + k^{-1} e_{\theta1}(r, t) + \ldots,$$

$$\ldots$$

where $e_{T0}$, $e_{\theta0}$, $e_{T1}$, $e_{\theta1}$ etc. are all of order unity.

Collecting the highest order terms from equation (6.31), we readily find that

$$\beta_0(r, t) \equiv 0.$$  

\textsuperscript{22}If anything, a large value of $l$ strengthens our analysis. When $l$ is large, we find that $e_T$ is small, which is also a consequence of $k$ being large.

\textsuperscript{23}Furthermore, we recall that it is physically realistic for cells to preferentially divide in the direction of least compressive stress. As shown by Jones et al. \cite{97}, isotropic growth can sometimes lead to an unrealistic prediction of infinite stress.
Now, the leading order terms of equation (6.34) yield the result that

$$\frac{\partial \sigma_{r0}}{\partial r} + \frac{2}{r} \beta_0 = 0.$$  

Hence, it follows that

$$\frac{\partial \sigma_{r0}}{\partial r} = 0,$$

and we can use the boundary condition from (6.23) to show that

$$\sigma_{r0}(r, t) \equiv 0.$$

Given that $\beta_0$ and $\sigma_{r0}$ are both identically equal to zero, we can use the leading order equations obtained from (6.32) and (6.33) to show that

$$e_{r0}(r, t) \equiv e_{T0}(r, t) \equiv 0.$$  

Thus, we find that the stresses and strains are all zero to leading order. This is consistent with our expectation that a large value of $k$ should lead to small stresses and strains. Effectively, we have shown that the stresses and strains are all of asymptotic order $k^{-1}$ as $k$ tends to infinity.\textsuperscript{24}

Now, consider the equation obtained from the leading order terms of equation (6.30). Given that $e_{T0} \equiv 0$, this yields the following equation for $v_0(r, t)$:

$$\frac{\partial v_0}{\partial r} + \frac{2 v_0}{r} = g(r, t).$$

Rearranging, we find that

$$\frac{1}{r^2} \frac{\partial}{\partial r} (r^2 v_0) = g(r, t). \quad (6.35)$$

\textsuperscript{24}This has consequences for our original nondimensionalisation. Clearly, the scalings that we used originally are not actually the natural scales for the variables. Thus, it would be appropriate to rescale the dimensionless stresses and strains with respect to $k^{-1}$ to obtain a new dimensionless system. However, since the analysis of this new system is analogous to the analysis that we have described, we will not rescale.
For a spherically-symmetric problem, we recall that \( \frac{1}{r^2} \frac{\partial}{\partial r} (r^2 v) \) is equal to the divergence of velocity. Thus, equation (6.35) has a simple physical interpretation. To leading order, we find that the divergence of velocity is equal to the rate of growth.

This is consistent with our earlier comments about small strains and stresses. As we discussed in Section 5.3.1, the assumption of stress free growth leads to

\[
\text{div } \mathbf{v} = g(\mathbf{x}, t).
\]

Thus, our leading order solution for velocity is identical to the complete solution that we would have obtained if we had made an initial assumption of zero stresses.

Furthermore, we recall that many models of tumour growth use the divergence of velocity or an equivalent as a measure of the rate of growth (see, for example, the influential models of Greenspan [73] and Casciari et al. [22]). If \( k \) is indeed large, our work suggests that these models can be thought of as leading-order approximations to the true mechanical behaviour of the growing spheroid. In contrast, if \( k \) is small, it would be more appropriate to use the isotropic growth approach proposed by Jones et al. [97] and MacArthur and Please [114].

Given that \( v_0(0, t) \equiv 0 \), we can solve equation (6.35) by integrating. This yields the result that

\[
v_0(r, t) = \frac{1}{r^2} \int_0^r \bar{r}^2 g(\bar{r}, t) \, d\bar{r}.
\]  

(6.36)

Thus, we have obtained leading order solutions for \( v, e_T, e_\theta, \sigma_r \) and \( \beta \). However, only the solution for \( v \) is nontrivial. For the rest, we find that our original scaling was incorrect. In order to obtain nonzero terms, we need to look at the first order correction.

We will not investigate the first order correction in detail. However, we are able to make some interesting observations about the variable \( \beta \), the difference between the radial stress and the circumferential stress. Collecting the order unity terms
of equation (6.31), we find that

\[
\frac{1}{3} \beta_1 = \frac{g(r, t)}{3} - \frac{v_0}{r}.
\]

Using equation (6.36), this yields the result that

\[
\beta_1(r, t) = g(r, t) - 3 \int_0^r \bar{r}^2 g(\bar{r}, t) \, d\bar{r}
\]

\[
= g(r, t) - 3 \left( \left[ \frac{\bar{r}^3}{3} g(\bar{r}, t) \right]_0^r - \frac{1}{3} \int_0^r \bar{r}^3 \frac{\partial g(\bar{r}, t)}{\partial \bar{r}} \, d\bar{r} \right)
\]

\[
= g(r, t) - g(r, t) + \frac{1}{r^3} \int_0^r \bar{r}^3 \frac{\partial g(\bar{r}, t)}{\partial \bar{r}} \, d\bar{r}
\]

\[
= \frac{1}{r^3} \int_0^r \bar{r}^3 \frac{\partial g(\bar{r}, t)}{\partial \bar{r}} \, d\bar{r}. \tag{6.37}
\]

This is a very powerful and useful result. Equation (6.37) demonstrates that the gradient of the growth rate can be used to obtain an approximate solution for the difference between the radial stress and the circumferential stress. If the growth rate is uniform throughout the spheroid, we would find that \( \frac{\partial g}{\partial r} = 0 \) and hence \( \beta_1(r, t) = 0 \). That is, a constant rate of growth leads to an isotropic stress field.\textsuperscript{25} However, if different parts of the spheroid are growing at different rates, we find that internal stresses develop.

It is well known that the cells within a tumour spheroid proliferate fastest when they are near the surface of the spheroid [160]. Cells that are too far away from the surface stop proliferating and die because they are deprived of oxygen and essential nutrients. The dead cells form a collapsing necrotic core that has the effect of dragging the outermost cells inwards.

Thus, most mathematical models of spheroid growth use a monotone increasing function of \( r \) to describe the local rate of growth. Within the necrotic core, the

\textsuperscript{25}Given the stress-free boundary conditions that we are using, we would even find that the stress is universally zero in this case.
growth rate is negative, reflecting the shrinkage that occurs as cell bodies collapse. In contrast, the growth rate is positive in the outermost regions where the cells continue to proliferate.

If we assume that \( g(r, t) \) is a monotone increasing function of \( r \), we find that the integrand in equation (6.37) is always positive. Thus, \( \beta \) is always positive, indicating that the radial stress is always larger (i.e. more tensile) than the circumferential stress. In fact, it is possible to show that \( \sigma_r \) is always positive and \( \sigma_\theta \) is always negative. Hence, our model predicts that a floating tumour spheroid will experience radial tension and circumferential compression throughout.\(^{26}\)

The observation that radial stress is greater than circumferential stress is consistent with the predictions of earlier models [9, 97, 114]. However, it remains to be seen whether these predictions are borne out by experimental results. We note that all of these mechanical models (including ours) assume that cells are particles that are passively advected into the centre of the spheroid. In actuality, experimental results indicate that the cells in tumour spheroids exert tractional forces that need to be taken into account [67].

Furthermore, it is reasonable to expect that cells actively migrate towards the surface of the spheroid in search of oxygen and nutrients. For example, Dorie and coworkers showed that inert polystyrene spheres move towards the interior of a spheroid at a different rate from radiolabelled cells [48, 49]. By taking into account the active chemotactic movement of cells, Pettet et al. [151] and Landman and Please [106] were able to develop a mathematical model of spheroid growth that replicates these results. However, it would appear that none of the models in the existing mathematical literature consider the effects of chemotactic movement on mechanical stress.

Also, we have not considered the possibility that the stress profiles we obtain are unstable. Ben Amar and Goriely [18, 69] showed that the residual stresses

\(^{26}\)This is dependent on the gradient of the growth rate. As we noted earlier, if \( g(r, t) \) is independent of \( r \) then no stresses develop.
associated with growth can lead to the buckling of hollow shells. Although we
would not expect to see similar effects for solid spheres, further investigation is
needed in order to confirm that this is the case.27

In conclusion, we have demonstrated that our theory of morphoelasticity can be
used to develop a model of tumour spheroid growth. By using asymptotic meth-
ods, we have been able to show that our model predicts that tumour spheroids
exist in a global mechanical state of radial tension and circumferential compres-
sion. Experimental investigation is necessary in order to determine whether or
not this is the case. Also, numerical simulations are needed in order to properly
compare our model with other models in the existing literature.

6.4 A mechanochemical model of dermal wound
healing

6.4.1 Modelling approach

In this section, we unite the themes of this thesis by proposing a novel mechano-
chemical model of dermal wound healing that incorporates morphoelasticity. As
described in Section 2.3, the present models of dermal wound healing fail to
combine mechanical effects and biochemical effects in a realistic way. Moreover,
most of the current models assume that the dermal ECM behaves as a viscoelastic
solid; in Section 3.4 we demonstrated that a morphoelastic model would be more
appropriate.

27We note that Goriely and Ben Amar [69] used their cumulative description of growth
to construct models of growing (and shrinking) hollow, spherical shells in an incompressible
medium. Because the internal and external pressures on the shells were not in equilibrium,
Goriely and Ben Amar obtained results where the circumferential stress is not globally larger
than the radial stress. It would be interesting to see whether these results can be replicated
using our model.
Unfortunately, we found that the present theories of morphoelasticity are too complex to incorporate into a model of dermal wound healing. This necessitated the development of the new theory described in Chapters 4 and 5. We now apply this theory to develop a description of the mechanics of the dermis.

Although our theory of morphoelasticity is well-suited to two-dimensional and three-dimensional modelling, we use a one-dimensional Cartesian coordinate system for the model presented here. This allows us to use the one-dimensional strain evolution equations given in Section 5.4.1. Furthermore, a one-dimensional framework enables us to avoid the complexities of directional anisotropy and it simplifies the task of comparing our work with previous work.

In many ways, the model that we propose is similar to the one-dimensional model of dermal wound healing proposed by Olsen and coworkers [136, 140]. Like Olsen et al., we consider two species of fibroblast (‘normal’ fibroblasts and myofibroblasts) and we incorporate a freely-diffusing growth factor that affects the behaviour of these cells. Also, our model includes a mathematical description of ECM synthesis and degradation that is analogous to Olsen et al.’s work.

Unlike Olsen et al., we use an elastic constitutive law to describe the mechanical behaviour of the dermis instead of a viscoelastic constitutive law. Moreover, we use a morphoelastic model to describe the evolution of effective strain. This aspect of our model is similar to that of Cook [33]; like Cook, we consider the plastic effects of ECM turnover. However, there are some subtle but important differences between our approach and Cook’s approach. These are discussed later in Section 6.4.2.

The dependence of myofibroblast differentiation on stress

None of the previous models of dermal wound healing have accounted for the fact that the differentiation of fibroblasts into myofibroblasts is dependent on mechanical stress. As described in Section 2.2, it is known that fibroblasts develop
stress fibres when they are cultured under mechanical tension and/or on a stiff substrate [45, 75, 76, 186]. These fibroblasts, which are called protomyofibroblasts, are able to exert more stress on the surrounding ECM and they synthesise collagen fibres more quickly than ‘normal’ fibroblasts. With the further introduction of growth factors such as TGFβ, these protomyofibroblasts differentiate into myofibroblasts, which are characterised by the presence of α-smooth muscle actin [75, 186].

Mathematically, we find that making the rate of myofibroblast differentiation dependent on stress prevents the mechanical aspects of the model from becoming decoupled from the rest of the system. In the Olsen and Tranquillo-Murray models, fibroblasts are passively advected with the moving ECM but are otherwise unaffected by the mechanical properties of the dermis. However, we find that advection is generally small when compared to the other terms in the cell density equation. Thus, the mechanical parts of the Olsen and Tranquillo-Murray models can be removed without significantly changing the model predictions. We wish to avoid this decoupling by ‘closing the loop’ and incorporating feedback from the mechanical model to the cell-behaviour model.

However, modelling the effect of mechanical stress on fibroblast-myofibroblast differentiation introduces a number of difficulties. For example, we note that the stress experienced by the dermis is expressed as the sum of an elastic component and a traction component. It is unclear as to whether the rate of fibroblast-myofibroblast differentiation is dependent on the elastic stress, the traction stress or the total stress.

---

28Our model does not include protomyofibroblasts as a separate species from myofibroblasts. Thus, we assume that fibroblast to myofibroblast differentiation is a single stage process that depends on mechanical tension and the concentration of growth factors.

29We note that the smallness of advective terms is a direct result of the smallness of velocity. Moreover, we expect the velocity of the dermis to be relatively small on the time scale of wound healing. Although it may be possible to obtain more advection by changing the parameter values, the resulting model would exhibit unrealistically large velocities.

30In the Tranquillo-Murray and Olsen models, the total stress is expressed as the sum of an elastic component, a viscous component and a traction component. However, we use an elastic model rather than a viscoelastic model.
In order to determine which of these is most appropriate, let us consider the initial condition of the elastic-only model described in Section 3.4.2. From equation (3.74), we find that the initial displacement is given by

\[ u(x, 0) = \begin{cases} \frac{x}{\sqrt{s}} e^{-\sqrt{s}} \sinh (x \sqrt{s}), & 0 \leq x \leq 1; \\ \frac{\tau}{\sqrt{s}} \sinh (\sqrt{s}) e^{-\sqrt{s} x}, & x \geq 1; \end{cases} \]  

where \( \tau = \frac{\tau_0}{1+\gamma} \) is representative of the traction stress exerted by cells in the healthy skin and \( s \) describes the strength of the tethering forces.

In this dimensionless model, the elastic stress is equal to the elastic strain. Thus,

\[ \sigma_e = \frac{\partial u}{\partial x}. \]

Differentiating equation (6.38) yields

\[ \sigma_e = \begin{cases} \tau e^{-\sqrt{s}} \cosh (x \sqrt{s}), & 0 \leq x < 1; \\ -\tau e^{-x \sqrt{s}} \sinh (\sqrt{s}), & x > 1. \end{cases} \]

Also, we note that the dimensionless traction stress is initially 0 within the wound and \( \tau \) outside:

\[ \psi = \begin{cases} 0, & 0 \leq x < 1; \\ \tau, & x > 1. \end{cases} \]

Using these results, we can plot the initial profiles of elastic stress, traction stress and total stress. These are shown in Figure 6.2.

Interestingly, we find that the total stress is a monotone increasing function of the distance from the wound centre, as is the cell traction stress. Since positive stress corresponds to tension and negative stress corresponds to compression, these measures would suggest that the skin is under greatest tensile stress outside the wound. If we make the reasonable assumption that increased tension
leads to increased myofibroblast differentiation, we would find that myofibroblast differentiation is encouraged outside the wound rather than inside the wound.

In contrast, we note that the elastic stress is positive inside the wound and negative outside it. Thus, we propose that it is more realistic to make the rate of myofibroblast differentiation dependent on the elastic stress than the total stress. This is consistent with a number of experimental observations. Most notably, Hinz et al. [90] demonstrated that myofibroblast differentiation slows and/or stops when the tension is released from splinted wounds. Thus, per-

\[\tau = s = 1.\]
mitting a wound to contract (i.e. allowing $\sigma_e$ to decrease and possibly become negative) decreases myofibroblast differentiation, even if the cell traction stress remains large.

Model variables

As described above, we wish to develop a mechanochemical model of wound healing that takes into account the plasticity of the dermis and the importance of mechanical and biochemical signals. We distinguish between fibroblasts and myofibroblasts but we will not consider any other cell types; although macrophages and endothelial cells are important to granulation tissue formation, they are not included in this model. As is conventional, we adopt the continuum hypothesis and thus represent fibroblasts and myofibroblasts using density variables.

In addition to these cell densities, our model includes several variables that describe the mechanical behaviour of the ECM (displacement, velocity, strain etc.). Also, we represent the effects of inflammation using a generic diffusible growth factor. The concentration of this growth factor is assumed to reflect the concentration of TGF-β and other growth factors that are important in wound healing (e.g. platelet derived growth factor, macrophage derived growth factor etc.). A full list of model variables is given in the table below:
It should be noted that not all of the dependent variables above require differential equations. For example, the traction stress will be given explicitly as a function of the fibroblast density, the myofibroblast density and the ECM density. However, it is convenient for us to track the development of the traction stress as a separate species from the species given above.

Outline of expected behaviour

Before we construct equations for our model, it is important for us to consider the behaviour that we want our model to exhibit. Furthermore, we need to think carefully about plausible mechanisms for this behaviour. The simplified description of wound healing given below outlines the essential features of dermal wound healing that we wish to reflect in our model. This narrative will be used to guide the development of our model through the rest of this section.

- When a wound occurs, cells are cleared from the wound space and the
dermis is replaced with a provisional fibrin matrix.

- Inflammation at the site of the wound means that growth factors are produced; these factors encourage fibroblasts to enter the wound space by chemotaxis.

- The fact that there is initially a greater density of fibroblasts outside the wound than inside the wound means that the traction stresses are unbalanced. As a result, the wound stretches open, leading to a positive elastic stress within the wound space.

- In response to a combination of mechanical stress and the presence of growth factors, some fibroblasts differentiate into myofibroblasts. These myofibroblasts synthesise and degrade ECM at a different rate from fibroblasts and they are also less motile.

- Fibroblasts and myofibroblasts both exert traction stresses on the ECM, but myofibroblasts exert more stress than fibroblasts.

- These traction stresses cause the wound to contract, thus relieving the initial elastic stress. As the elastic stress decreases, myofibroblast differentiation slows and stops.

- At all times, there is a balance between the elastic stress, the traction stress and a restoring force that represents subdermal tethering. Inertial effects and viscous effects are treated as being negligible on the time scale of wound healing.

- In a healthy wound, the density of myofibroblasts will decrease over time, ultimately leading to a contracted scar that has the same cell density as the surrounding skin. However, some wounds may lead to pathological scars where myofibroblasts persist in the scar tissue and/or the scar grows beyond the original boundaries of the wound. This may depend on the level of tension in the healthy skin and other model parameters.
6.4.2 Development of model equations

Following the outline given above, we now construct a system of partial differential equations to define our model. Despite the fact that we expect advection to be small, it should be noted that we include advective fluxes in the equations given below. Most probably, these terms can be removed without significantly affecting the model predictions.

Fibroblast density

Expressed as a word equation, the evolution of fibroblast density will take the form:

\[ \frac{\partial}{\partial t} \text{(fibroblast density)} + \frac{\partial}{\partial x} \text{(flux)} = \text{proliferation} - \text{death} - \text{conversion to myofibroblasts}. \]

Given this general form, we need to develop suitable expressions for the flux, proliferation, death and conversion terms.

For the flux term, we propose that fibroblasts undergo random motion, chemotaxis and advection with the underlying ECM. Using classical expressions for these terms (e.g. Fickian diffusion to represent random motion), we find that the total flux takes the form

\[ \text{flux} = -\mu_n \frac{\partial n}{\partial x} + \chi_n n \frac{\partial c}{\partial x} + n v, \]

where \( \mu_n \) represents the rate of random movement and \( \chi_n \) represents the sensitivity of the fibroblasts to chemotaxis.

Under mechanical stress, fibroblasts show upregulated expression of genes that

---

\(^{32}\)Since there is little experimental evidence to support the view that myofibroblasts can differentiate back into fibroblasts, we assume that conversion only occurs in one direction [186].
are associated with proliferation [99]. Given that this is the case, an appropriate expression for the net rate of proliferation and death might take the form

\[ \text{proliferation} - \text{death} = (a_n + a_{n\sigma} \sigma^+ - \theta_{nn} n) n, \]

where \( a_n \) represents the base rate of proliferation, \( a_{n\sigma} \) represents the dependence of proliferation on mechanical tension and \( \theta_{nn} \) represents the rate of death due to overcrowding. If \( a_{n\sigma} \) were zero, this would yield a conventional logistic growth term.

The term \( \sigma^+ \) refers to the positive part of elastic stress; that is,

\[ \sigma^+ = \begin{cases} 0, & \sigma \leq 0; \\ \sigma, & \sigma > 0. \end{cases} \]

By using \( \sigma^+ \) instead of \( \sigma \), we ensure that \( a_n \) is the minimum rate of fibroblast proliferation. Mechanical tension increases the rate of proliferation but mechanical compression has no effect at all. Further investigation is needed in order to determine whether this is realistic.

Lastly, we assume that the net rate of fibroblast-myofibroblast conversion is directly proportional to the positive part of elastic stress and to the growth factor concentration. That is,

\[ \text{conversion to myofibroblasts} = \gamma \sigma^+ c n, \]

where \( \gamma \) represents the rate of conversion and \( \sigma^+ \) is defined as above.

Combining these terms, we obtain the following evolution equation for the fibroblast density:

\[
\frac{\partial n}{\partial t} = \mu_n \frac{\partial^2 n}{\partial x^2} - \frac{\partial}{\partial x} \left( \chi_n n \frac{\partial c}{\partial x} \right) - \frac{\partial}{\partial x} \left( n v \right) + (a_n + a_{n\sigma} \sigma^+ - \theta_{nn} n) n - \gamma \sigma^+ c n
\]
Myofibroblast density

Similar to the fibroblast evolution equation, we propose that the myofibroblast evolution takes the following form:

\[
\frac{\partial}{\partial t} \text{(myofibroblast density)} + \frac{\partial}{\partial x} \text{(flux)} = \text{proliferation} - \text{death} + \text{conversion from fibroblasts}.
\]

Since myofibroblasts are less motile than fibroblasts, we will assume that the only flux of myofibroblasts is due to advection. That is, myofibroblasts move passively with the underlying ECM but are otherwise stationary. This yields a flux term of the form

\[
\text{flux} = m v.
\]

The conditions that are needed to stimulate the proliferation of myofibroblasts are not fully understood. However, recent studies suggest that certain growth factors can increase the rate of myofibroblast proliferation [65]. Furthermore, it is reasonable to expect that, like fibroblast proliferation, myofibroblast proliferation is enhanced in conditions of mechanical tension.

Although mechanical and chemical signals can increase myofibroblast proliferation, we note that myofibroblasts are rare or absent in the healthy skin [45, 75]. Thus, we assume that the net rate of myofibroblast proliferation is zero or negative in the absence of these signals.

Combining these observations, we propose that the net rate of myofibroblast proliferation and death takes the form

\[
\text{proliferation} - \text{death} = (a_{m\sigma} \sigma + a_{mc} c - \theta_m - \theta_{mm} m) m,
\]

where \(a_{m\sigma}\) is the rate of myofibroblast proliferation in response to mechanical
tension, \( a_{mc} \) is the rate of myofibroblast proliferation in response to the growth factor, \( \theta_m \) is the rate of myofibroblast death in the absence of chemical and mechanical signals and \( \theta_{mm} \) is the rate of death from overcrowding.\(^{33}\)

Using the fibroblast-myofibroblast conversion terms described above, this yields the following evolution equation:

\[
\frac{\partial m}{\partial t} = - \frac{\partial}{\partial x} (m v) + \gamma \sigma^+ c n + (a_{m\sigma} \sigma^+ + a_{mc} c - \theta_m - \theta_{mm} m) m.
\]

**ECM density**

As with the fibroblast densities, the evolution of the ECM density can be represented using a continuity equation:

\[
\frac{\partial}{\partial t} \text{(ECM density)} + \frac{\partial}{\partial x} \text{(flux)} = \text{production} - \text{degradation}.
\]

The only flux of ECM will be due to passive advection, which yields

\[
\text{flux} = \rho v.
\]

Also, we assume that the rates of ECM synthesis and ECM degradation are both proportional to the cell densities. However, fibroblasts and myofibroblasts may synthesise and degrade the ECM at different rates; thus, we find that

\[
\text{production} - \text{degradation} = \left( \eta_n n + \eta_m m \right) - \left( \pi_n n \rho + \pi_m m \rho \right),
\]

where \( \eta_n \) and \( \eta_m \) are rates of production and \( \pi_n \) and \( \pi_m \) are rates of degradation.\(^{34}\)

\(^{33}\)Death from overcrowding may be unnecessary in this model. However, we include this term in order to ensure that \( m \) is always bounded.

\(^{34}\)It is interesting to note that Olsen [136, 140] proposed a model in which the rate of ECM synthesis increases with increasing growth factor concentration. Similarly, there is experimental evidence that indicates that protomyofibroblasts (i.e. fibroblasts that have partially differenti-
Hence, we model the evolution of ECM density using an equation of the form

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x} (\rho v) + (\eta_n n + \eta_m m) - (\pi_n n \rho + \pi_m m \rho).$$

**Growth factor concentration**

The growth factor concentration equation is identical to the ECM density in its general form:

$$\frac{\partial}{\partial t} \text{ (growth factor concentration)} + \frac{\partial}{\partial x} \text{ (flux)} = \text{production} - \text{degradation}.$$  

However, the specific expressions used to represent these terms are very different.

Unlike the ECM, the growth factor can diffuse freely; this leads to a Fickian contribution to the flux term. We assume that the growth factor is not bound to the fibres of the ECM and it is not necessary to include advection in the growth factor equation. However, we note that many important growth factors do bind to the ECM, so this assumption may need to be revisited [40]. This yields an overall flux of the form

$$\text{flux} = -\mu_c \frac{\partial c}{\partial x},$$

where $\mu_c$ is the rate of random motility.

The production and degradation of the growth factor require careful attention. In order to simplify the model, we have not included macrophages or any other cells that might synthesise the growth factor. Thus, we will assume that the growth factor is generated directly in the wound space. For the purposes of growth factor production, the ‘wound’ is defined to be the region of the skin where the density of the ECM is below some critical value, $\rho_{\text{crit}}$. If the rate of growth factor production

_ated due to mechanical tension_ synthesise collagen at a considerably faster rate than ‘normal’ fibroblasts. Thus, it would be reasonable for us to make $\eta_n$ (and possibly $\eta_m$) dependent on other factors. For simplicity, however, we will focus on a simple model where collagen synthesis is dependent only on the cell density.
is constant within the wounded region, we find that

production = \( a_c H(\rho_{\text{crit}} - \rho) \),

where \( a_c \) is the rate of production and \( H \) represents the Heaviside function.

It should be noted that the Olsen et al. model uses a very different formulation for the rate of growth factor production. In their model, it is assumed that fibroblasts and myofibroblasts generate the growth factor as well as respond to it. Indeed, this may be necessary in our model in order to prevent the growth factor from being eliminated too rapidly. Once the ECM density reaches \( \rho_{\text{crit}} \), the growth factor will quickly be removed by diffusion and degradation. Further investigation is necessary in order to determine whether the expression above needs to be altered.

Also, the Olsen et al. model assumes that fibroblasts and myofibroblasts consume the growth factor when they respond to it. In contrast, we assume that the growth factor is only removed through spontaneous degradation. That is,

\[ \text{degradation} = \theta_c c, \]

where \( \theta_c \) is the rate of elimination.

Combining these terms, we develop the following evolution equation for the growth factor concentration:

\[
\frac{\partial c}{\partial t} = \mu_c \frac{\partial^2 c}{\partial x^2} + a_c H(\rho_{\text{crit}} - \rho) - \theta_c c.
\]

Displacement, velocity and force balance

In general, it seems sensible for us to use the true velocity rather than the approximate velocity. For example, we recall that the evolution of strain in equation
(5.55) depends significantly on the spatial gradient of the true velocity. Moreover, using the true velocity enables us to correctly define advective fluxes.

The true velocity is the material derivative of displacement. That is,

$$\frac{\partial u}{\partial t} + v \frac{\partial u}{\partial x} = v. \quad (6.39)$$

As with the elastic-only model described in Section 3.4.2, we cannot easily use this equation to describe the evolution of the displacement profile.\(^{35}\) Thus, we may find it more useful to rearrange (6.39) to obtain an explicit equation for velocity:

$$v = \frac{\frac{\partial u}{\partial t}}{1 - \frac{\partial u}{\partial x}}. \quad (6.40)$$

As before, we have two options for finding \(u\) and \(v\): we can use equation (6.40) in conjunction with a standard force balance equation or we can use (6.39) in conjunction with the time derivative of the force balance equation. In this second case, it is useful for us to rearrange equation (6.39) so that it is more like a classical continuity equation. One way of achieving this is to introduce \(w\), the displacement gradient. We recall from Section 5.4.2 that

$$\frac{D}{Dt} \left( \frac{\partial u}{\partial x} \right) = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial x} \frac{\partial v}{\partial x}.$$ 

Thus, if \(w = \frac{\partial u}{\partial x}\) we find that

$$\frac{\partial w}{\partial t} + \frac{\partial}{\partial x} (w v) = \frac{\partial v}{\partial x}.$$ 

Once \(w\) is known, it is possible to integrate with respect to \(x\) to obtain \(u\).

---

\(^{35}\)This is a result of the fact that an elastic-only force balance equation contains no explicit dependence on velocity. If we were to use (6.39) to define the evolution of the displacement profile, we would need to adapt the force balance equation to obtain an equation that can be solved for velocity.
As mentioned above, the force balance equation is also important for calculating displacement and velocity. We recall that there is no viscous contribution to the stress tensor in our formulation. However, our force balance equation is otherwise identical to the analogous equations used in earlier models. This yields

$$\frac{\partial}{\partial x} (\sigma + \psi) = b \rho u,$$

where $b$ represents the strength of the subdermal tethering.\(^{36}\)

Since $\sigma$ and $\psi$ are explicitly expressed functions of the other model variables, it is most appropriate to use this equation to solve for $u$ or $v$. If using the force balance equation to solve for $v$, we must differentiate it with respect to $t$ and use the fact that $\frac{\partial u}{\partial t} = v (1 - w)$. This will yield an ordinary differential equation in space that can be solved at each time step to obtain $v$. Interestingly, it is possible to eliminate $u$ altogether by rearranging the force balance equation before differentiating.

**Stress and strain**

We assume that the dermis behaves as a linear elastic material. That is, the elastic stress will be directly proportional to the effective strain at all times. However, we also assume that the elastic modulus of the dermis will change according to the density of the ECM. This is not unusual in biomechanical models; Ramtani \([155, 156]\) developed models of collagen lattice contraction in which the elastic modulus is dependent on the collagen density.

For simplicity, we assume that the dependence of elastic modulus on density is linear. Thus,

$$\sigma = \frac{E_0 \rho}{\rho_0} e,$$

where $E_0$ is the elastic modulus of the healthy skin and $\rho_0$ is the density of the

\(^{36}\)The parameter $b$ is equivalent to $s$ in Tranquillo and Murray’s notation \([189]\).
Applications of zero stress state theory and morphoelasticity

healthy skin.

Similarly, we assume that the traction stress increases with increasing cell density and with increasing ECM density. Although Olsen et al. proposed more complex expressions, we will assume that traction stress is directly proportional to both ECM density and cell density. Considering fibroblasts and myofibroblasts separately, this yields

\[ \psi = \beta_n n \rho + \beta_m m \rho, \]

where \( \beta_n \) represents the traction strength of fibroblasts and \( \beta_m \) represents the traction strength of myofibroblasts.

The last equation to consider is the evolution of strain. From our one-dimensional linear theory of morphoelasticity, we recall that

\[ \frac{\partial e}{\partial t} + \frac{\partial}{\partial x} (e v) = \frac{\partial v}{\partial x} - g(x, t), \]

where \( g(x, t) \) represents the rate of growth (i.e. the rate at which the zero stress state grows larger).

In Cook’s model of wound healing [33], the zero stress state changes in response to matrix turnover; as stressed fibres are replaced with unstressed fibres, the effective strain decreases towards zero.\(^{37}\) If fibres are degraded at random and all new fibres are synthesised at an unstressed state, the rate of relaxation is given by the rate of ECM synthesis relative to the total ECM density. In our notation, this yields a growth term of the form\(^{38}\)

\[ g = \frac{(\eta_n n + \eta_m m) e}{\rho}. \]

\(^{37}\)Cook also considers the possibility of classical plasticity (i.e. the zero stress state changes when some yield threshold is exceeded). We do not include classical yielding in our model of wound healing.

\(^{38}\)Note that there is no need for an arbitrary constant of proportionality in this expression. See Cook [33] for further details.
The problem with this expression is that it implies that the strain universally tends towards zero. However, it is known that the skin exists in a continual state of tension. Moreover, skin tension persists after death, leading to the well-recorded phenomenon of Langer’s lines. Thus, we would like our model to predict the existence of a nonzero steady-state strain distribution.

Cook resolved this problem by assuming that a proportion of the fibres are synthesised with the same stress distribution as the surrounding fibres. This proportion is dependent on the density; once a critical density is exceeded, all new fibres are ‘prestressed’. Thus, the rate of remodelling plasticity goes to zero once the critical density has been reached.

However, this seems unrealistic. We would expect that fibres are synthesised in a stress-free state and they later develop stress as a result of cells moving the fibres and ‘tying them together’ in a new configuration. Effectively, this means that the desired distance between material points in the ECM decreases because of the action of cells. In order to reflect this, our model includes terms that describe the direct action of fibroblasts and myofibroblasts on the zero stress state. Algebraically, this yields

\[ g = \frac{(\eta_n n + \eta_m m) e}{\rho} - \zeta_n \rho n - \zeta_m \rho m, \]

where \( \zeta_n \) and \( \zeta_m \) represent the respective abilities of fibroblasts and myofibroblasts to change the zero stress state.\(^{39}\)

It should be noted that changing the zero stress state is different from applying a traction stress. While \( \beta_n \) and \( \beta_m \) represent the capacity of cells to ‘pull’ on the ECM, \( \zeta_n \) and \( \zeta_m \) represent the capacity of cells to deform the ECM and then rebind the collagen network so that these local changes are made permanent. Effectively, traction depends on the ‘brute strength’ of the fibroblasts, while changing the zero stress state requires a combination of strength and the ability

\(^{39}\)Note that the action of cells leads to contraction, not expansion. This means that the corresponding terms in the growth expression must be negative.
Applications of zero stress state theory and morphoelasticity

to refashion the local ECM.

It is not clear as to the correct formulation that should be used to represent the direct effect of cells on the ECM. We propose that the rate of permanent contraction due to cellular effects is proportional to the cell density and to the ECM density. Further work is needed in order to establish whether a different expression would be more appropriate.

Combining these effects yields the following equation for strain evolution:

$$\frac{\partial e}{\partial t} = -\frac{\partial}{\partial x}(e v) + \frac{\partial v}{\partial x} - \frac{(\eta_n n + \eta_m m) e}{\rho} + \zeta_n \rho n + \zeta_m \rho m.$$  

6.4.3 Initial conditions and boundary conditions

We have constructed nine equations for the nine unknowns of our system.\(^40\) These are given below.

$$\frac{\partial n}{\partial t} = \mu_n \frac{\partial^2 n}{\partial x^2} - \frac{\partial}{\partial x} \left( \chi_n n \frac{\partial c}{\partial x} \right) - \frac{\partial}{\partial x} (n v) + \left( a_n + a_{n\sigma} \sigma^+ - \theta_n \right) n - \gamma \sigma^+ c n;$$

(6.41)

$$\frac{\partial m}{\partial t} = -\frac{\partial}{\partial x} (m v) + \gamma \sigma^+ c n + \left( a_{m\sigma} \sigma^+ + a_{mc} c - \theta_m - \theta_{mm} \right) m;$$

(6.42)

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x} (\rho v) + \left( \eta_n n + \eta_m m \right) - \left( \pi_n \rho + \pi_m \rho \right);$$

(6.43)

$$\frac{\partial c}{\partial t} = \mu_c \frac{\partial^2 c}{\partial x^2} + a_c H(\rho_{crit} - \rho) - \theta_c c;$$

(6.44)

$$v = \frac{\frac{\partial n}{\partial t}}{1 - \frac{\partial n}{\partial x}}.$$
\[ u = \frac{1}{b \rho} \frac{\partial}{\partial x} (\sigma + \psi); \]  
(6.46)

\[ \sigma = \frac{E_0 \rho}{\rho_0} e; \]  
(6.47)

\[ \psi = \beta_n n \rho + \beta_m m \rho; \]  
(6.48)

\[ \frac{\partial e}{\partial t} = -\frac{\partial}{\partial x} (e v) + \frac{\partial v}{\partial x} - \frac{(\eta_n n + \eta_m m) e}{\rho} + \zeta_n \rho n + \zeta_m \rho m. \]  
(6.49)

Given this system of equations, we now need to specify appropriate initial and boundary conditions before the model can be solved. Like earlier one-dimensional models, we consider a semi-infinite domain where \( x = 0 \) corresponds to the centre of the wound, \( x = L \) corresponds to the initial wound boundary and \( x > L \) corresponds to healthy skin.

By symmetry at the wound centre, we find that the displacement, the velocity and all fluxes must be zero at \( x = 0 \). This yields the following boundary conditions:

\[ u(0, t) = 0, \quad v(0, t) = 0, \quad n_x(0, t) = 0, \quad c_x(0, t) = 0. \]

As \( x \) tends to infinity, we require that all model variables approach bounded limits. Specifically, we require that

\[ n_x(\infty, t) \to 0, \quad c_x(\infty, t) \to 0 \quad \text{and} \quad u_x(\infty, t) \to 0. \]

If these three conditions are satisfied, it is possible to use the methods described in Section 3.2.2 to demonstrate that all spatial derivatives tend to zero as \( x \) tends to infinity.\(^{41}\)

---

\(^{41}\)Technically, this will only be true if the \( x \to \infty \) limit of the initial condition corresponds to a stable, spatially homogeneous equilibrium of the model. Since \( x \to \infty \) corresponds to the healthy skin, it is reasonable to expect that this will be the case.
Having obtained a complete set of boundary conditions, we now need to construct appropriate initial conditions. As with the elastic-only Tranquillo-Murray model described in Section 3.4.2, we expect to find that the initial conditions for strain and displacement are spatially nonhomogeneous. However, we now have the added complexity that the initial unwounded state may not correspond to a zero stress state. That is, the effective strain may be nonzero even if the apparent strain \( (i.e. \frac{\partial u}{\partial x}) \) is zero.

As a first step towards determining the wounded initial condition, let us consider the homogeneous steady state that corresponds to healthy skin. In healthy skin, the fibroblast density and the ECM density must be nonzero. However, the growth factor concentration and myofibroblast density should be zero, as should the velocity and the displacement. Thus, we seek a ‘healthy’ steady state of the form

\[
[n, m, \rho, c, v, u, \sigma, \psi, e] = [n_0, 0, \rho_0, 0, 0, 0, \sigma_0, \psi_0, e_0],
\]

where \( n_0 \) and \( \rho_0 \) are both nonzero.

Setting all derivatives in equation (6.41) to zero, we find that the equilibrium fibroblast density must satisfy

\[
\left( a_n + a_{n\sigma} (\sigma_0^+) - \theta_{nn} n_0 \right) n_0 = 0.
\]

Since \( n_0 \) is nonzero, we find that

\[
 n_0 = \frac{a_n + a_{n\sigma} (\sigma_0^+)}{\theta_{nn}}, \tag{6.50}
\]

where \( (\sigma_0^+) \) is the positive part of the equilibrium elastic stress.

Similarly, equation (6.43) yields the result that

\[
\eta_n n_0 - \pi_n n_0 \rho_0 = 0.
\]

Since \( n_0 \) is nonzero, this implies that \( \rho_0 = \frac{\eta_n}{\pi_n} \).
From equation (6.49), we find that

\[ -\frac{\eta_n n_0 e_0}{\rho_0} + \zeta_n \rho_0 n_0 = 0. \]

Using the fact that \( \rho_0 = \frac{\eta_n}{\pi_n} \), this yields

\[ e_0 = \frac{\zeta_n \eta_n}{\pi_n^2}. \]

We note that the ECM density of healthy skin appears as a parameter in equation (6.47). Since this is the physical interpretation of \( \rho_0 \) as given above, we find that equation (6.47) yields

\[ \sigma_0 = E_0 e_0 = \frac{E_0 \zeta_n \eta_n}{\pi_n^2}. \]

This is positive, which implies that \((\sigma_0)^+\) in equation (6.50) is nonzero. Substituting the above expression into (6.50), we find that

\[ n_0 = \frac{a_n \pi_n^2 + a_{n\sigma} E_0 \zeta_n \eta_n}{\theta_{nn} \pi_n^2}. \] (6.51)

Lastly, equation (6.48) gives the result that

\[ \psi_0 = \beta_n \rho_0 n_0 = \frac{\beta_n \eta_n \left( a_n \pi_n^2 + a_{n\sigma} E_0 \zeta_n \eta_n \right)}{\theta_{nn} \pi_n^3}. \]

In summary, the spatially homogeneous steady state associated with healthy skin is given by

\[ n_0 = \frac{a_n \pi_n^2 + a_{n\sigma} E_0 \zeta_n \eta_n}{\theta_{nn} \pi_n^2}, \quad m_0 = 0, \quad \rho_0 = \frac{\eta_n}{\pi_n}, \]

\[ e_0 = 0, \quad v_0 = 0, \quad u_0 = 0, \]

\[ \sigma_0 = \frac{E_0 \zeta_n \eta_n}{\pi_n^2}, \quad \psi_0 = \frac{\beta_n \eta_n \left( a_n \pi_n^2 + a_{n\sigma} E_0 \zeta_n \eta_n \right)}{\theta_{nn} \pi_n^3}, \quad e_0 = \frac{\zeta_n \eta_n}{\pi_n^2}. \]
Interestingly, we found that this is the only steady state that has the properties that we associate with the healthy skin (e.g. no myofibroblasts and no inflammatory mediator). If the steady state above does not exist (or if it is not stable), we conclude that either the model parameters are incorrect or the entire model is invalid.

For example, we require that $\frac{\eta_n}{\pi_n} > \rho_{\text{crit}}$. If this were not the case, we would find that the growth factor is being generated in the healthy skin. Thus, the ‘healthy steady state’ given above would not satisfy equation (6.44). Furthermore, we note that the stability of the healthy steady state must be taken into consideration when determining whether a given set of parameters is valid. However, we have not performed a detailed study of stability at this stage.

Having determined the spatially homogeneous steady state that corresponds to healthy skin, we now consider the effect of introducing a wound. At the beginning of the proliferative stage of wound healing, we expect that a provisional fibrin matrix has formed within the wound space, but the ECM density will be smaller in the wound than in the surrounding skin. Also, we assume that there will be no fibroblasts or myofibroblasts within the wound space, but there will be a burst of inflammatory mediator that will encourage the ingrowth of cells.

There are a number of ways to describe this situation mathematically. One possibility is to use initial conditions for $n$, $m$, $\rho$ and $c$ that take the form

$$
n(x, 0) = \begin{cases} 
0, & x < L, \\
n_0, & x > L; 
\end{cases} \quad m(x, 0) = 0; 
$$

$$
\rho(x, 0) = \begin{cases} 
\rho_{\text{wound}}, & x < L, \\
\rho_0, & x > L; 
\end{cases} \quad c(x, 0) = \begin{cases} 
c_{\text{wound}}, & x < L, \\
0, & x > L; 
\end{cases}
$$

where $n_0$ and $\rho_0$ are defined as above.\(^{42}\)

\(^{42}\)Note that $\rho_{\text{wound}}$ must be less than $\rho_{\text{crit}}$ in order for the growth factor to be produced in the wound at $t = 0$. 
Now, we wish to specify the initial stress and strain distributions.\textsuperscript{43} In order to do this, we assume that the wound is not accompanied by any immediate plastic changes. That is, the Lagrangian zero stress deformation gradient of the initial wounded state is identical to the Lagrangian zero stress deformation gradient of the healthy skin. Although the displacement profile changes when a wound occurs, the relationship between the initial configuration and the zero stress configuration does not.

Following the definition of strain given in Section 4.4.2, we find that

\[ e = I - Z F^{-1}, \]  

where \( Z \) is the principal Lagrangian zero stress deformation gradient.

In one-dimensional Cartesian coordinates, all zero stress deformation gradients are principal. Furthermore, we note that

\[ F^{-1} = \frac{\partial X}{\partial x} = 1 - \frac{\partial u}{\partial x}. \]

Thus, equation (6.52) yields

\[ e = 1 - Z \left( 1 - \frac{\partial u}{\partial x} \right), \]

or equivalently,

\[ Z = \frac{1 - e}{1 - \frac{\partial u}{\partial x}}. \]

In the healthy steady state, \( e = e_0 \) and \( u = 0 \). Thus,

\[ Z = 1 - e_0. \]

\textsuperscript{43}As in Section 3.4.2, the lack of viscosity means that wound retraction is immediate, leading to an initially nonzero displacement profile. Interestingly, we note that this retraction should also affect the cell density and ECM density profiles; for example, the contracted region immediately outside the wound will exhibit a greater density of ECM than the rest of the healthy skin. However, we assume that such effects are negligible and we do not attempt to represent them mathematically.
If the Lagrangian zero stress deformation gradient is unchanged after wounding, this means that
\[
\frac{1 - e(x, 0)}{1 - u_x(x, 0)} = 1 - e_0,
\]
and thus,
\[
e(x, 0) = e_0 + (1 - e_0) u_x(x, 0).
\] (6.53)

Substituting the initial conditions that we have already obtained into equations (6.47) and (6.48), we find that
\[
\sigma(x, 0) = \begin{cases} 
E_0 \frac{\rho_{\text{wound}}}{\rho_0} \left( e_0 + (1 - e_0) u_x(x, 0) \right), & x < L, \\
E_0 (e_0 + (1 - e_0) u_x(x, 0)), & x > L;
\end{cases}
\] (6.54)
\[
\psi(x, 0) = \begin{cases} 
0, & x < L, \\
\beta_n n_0 \rho_0, & x > L.
\end{cases}
\] (6.55)

We now need to find an appropriate formulation for \( u(x, 0) \) that satisfies equation (6.46). Since this will involve solving an ordinary differential equation, we introduce the notation \( U(x) = u(x, 0) \). Considering only the region \( x < L \), we find that equation (6.46) yields
\[
U = \frac{1}{b \rho_{\text{wound}} E_0 \rho_{\text{wound}} \rho_0} \left( 1 - e_0 \right) \frac{d^2U}{dx^2},
\]
or equivalently,
\[
\frac{d^2U}{dx^2} - \lambda^2 U = 0,
\]
where
\[
\lambda = \sqrt{\frac{b \rho_0}{E_0 (1 - e_0)}}.
\]

Interestingly, we obtain the same result when we consider only the region \( x > L \).
In this case, equation (6.46) yields

\[ U = \frac{1}{b \rho_0} E_0 (1 - e_0) \frac{d^2 U}{dx^2}, \]

which also becomes

\[ \frac{d^2 U}{dx^2} - \lambda^2 U = 0. \]

As described earlier, we are given boundary conditions on \( u \) at \( x = 0 \) and \( x \to \infty \). In terms of \( U \), we find that

\[ U(0) = 0, \quad U_x(\infty) = 0. \]

Solving the above differential equations subject to these conditions, we find that

\[ U(x) = \begin{cases} A \sinh(\lambda x), & x < L, \\ B e^{-\lambda x}, & x > L; \end{cases} \tag{6.56} \]

where \( A \) and \( B \) are constants to be determined.

Now, we note that equation (6.53) depends on \( \frac{\partial u}{\partial x} \). In order for \( e \) to be bounded, it follows \( U \) must be continuous. That is,

\[ \lim_{x \to L^-} U(x) = \lim_{x \to L^+} U(x). \]

From equation (6.56), this means that

\[ A \sinh(\lambda L) = B e^{-\lambda L}. \tag{6.57} \]

Similarly, we note that equation (6.46) involves the spatial derivative of \( \sigma + \psi \). Thus, we require \( \sigma + \psi \) to be continuous in order for \( u \) to be bounded. From equations (6.54) and (6.55), this implies that

\[ \lim_{x \to L^-} E_0 \frac{\rho_{\text{wound}}}{\rho_0} (e_0 + (1 - e_0) u_x(x, 0)) = \lim_{x \to L^+} E_0 (e_0 + (1 - e_0) u_x(x, 0)) + \beta_n n_0 \rho_0. \]
In terms of $U$, this means that

$$
\lim_{x \to L^-} \left( \frac{\rho_{\text{wound}}}{\rho_0} U'(x) + \frac{\rho_{\text{wound}}}{\rho_0} e_0 \right) \frac{1}{1 - e_0} = \lim_{x \to L^+} \left( U'(x) + \frac{e_0}{1 - e_0} + \frac{1}{E_0 \left( 1 - e_0 \right)} \beta_n n_0 \rho_0 \right).
$$

Thus,

$$
\frac{\rho_{\text{wound}} \lambda \cosh(\lambda L)}{\rho_0} A e_0 \frac{1}{1 - e_0} = -B \lambda e^{-\lambda L} + \frac{e_0}{1 - e_0} + \frac{1}{E_0 \left( 1 - e_0 \right)} \beta_n n_0 \rho_0.
$$

Solving (6.57) and (6.58) simultaneously, we find that

$$
A = \frac{e_0}{1 - e_0} \left( 1 - \frac{\rho_{\text{wound}}}{\rho_0} \right) + \frac{\beta_n n_0 \rho_0}{E_0 \left( 1 - e_0 \right)} \lambda \left( \sinh(\lambda L) + \frac{\rho_{\text{wound}}}{\rho_0} \cosh(\lambda L) \right),
$$

$$
B = \frac{\left( e_0 - \frac{\rho_{\text{wound}}}{\rho_0} \right) + \frac{\beta_n n_0 \rho_0}{E_0 \left( 1 - e_0 \right)} \lambda \left( 1 + \frac{\rho_{\text{wound}}}{\rho_0} \coth(\lambda L) \right) e^{\lambda L}}{\lambda \left( 1 + \frac{\rho_{\text{wound}}}{\rho_0} \coth(\lambda L) \right)}.
$$

Using equations (6.56) and (6.53), these expressions can be used to obtain explicit representations of the initial conditions for $u$ and $e$.

By a similar method, it is possible to obtain an ordinary differential equation for $v(x, 0)$.\(^{44}\) However, this equation is highly nonlinear and probably cannot be solved analytically. If we were to use the force balance equation to specify $v$ in our numeric approach, it would be necessary to solve the resulting ordinary differential equation numerically at every time step, including $t = 0$.

### 6.4.4 Nondimensionalisation

To summarise, our dimensional model takes the following form:

\(^{44}\)Note that this requires the use of the time derivative of the force balance equation
Equations:

\[ \frac{\partial n}{\partial t} = \mu_n \frac{\partial^2 n}{\partial x^2} - \frac{\partial}{\partial x} \left( \chi_n n \frac{\partial c}{\partial x} \right) - \frac{\partial}{\partial x} (n v) + \left( a_n + a_{ns} \sigma^+ - \theta_{nm} n \right) n - \gamma \sigma^+ c n; \]

\[ \frac{\partial m}{\partial t} = -\frac{\partial}{\partial x} (m v) + \gamma \sigma^+ c n + \left( a_{ms} \sigma^+ + a_{mc} c - \theta_m - \theta_{mm} m \right) m; \]

\[ \frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x} (\rho v) + (\eta_n n + \eta_m m) - (\pi_n n \rho + \pi_m m \rho); \]

\[ \frac{\partial c}{\partial t} = \mu_c \frac{\partial^2 c}{\partial x^2} + a_c H(\rho_{crit} - \rho) - \theta_c c; \]

\[ v = \frac{\frac{\partial u}{\partial t}}{1 - \frac{\partial u}{\partial x}}; \]

\[ u = \frac{1}{b \rho} \frac{\partial}{\partial x} (\sigma + \psi); \]

\[ \sigma = \frac{E_0 \rho}{\rho_0} e; \]

\[ \psi = \beta_n n \rho + \beta_m m \rho; \]

\[ \frac{\partial e}{\partial t} = -\frac{\partial}{\partial x} (e v) + \frac{\partial v}{\partial x} - \frac{(\eta_n n + \eta_m m) e}{\rho} + \zeta_n \rho n + \zeta_m \rho m. \]

Boundary conditions:

\[ n_x(0, t) = 0, \quad n_x(\infty, t) = 0; \]

\[ c_x(0, t) = 0, \quad c_x(\infty, t) = 0; \]

\[ u(0, t) = 0, \quad u_x(\infty, t) = 0. \]

Initial conditions:

\[ n(x, 0) = \begin{cases} 
0, & x < L, \\
n_0, & x > L; 
\end{cases} \]

\[ m(x, 0) = 0; \]
\[
\rho(x, 0) = \begin{cases} 
\rho_{\text{wound}}, & x < L, \\
\rho_0, & x > L;
\end{cases}
\]

\[
e(x, 0) = \begin{cases} 
\epsilon_{\text{wound}}, & x < L, \\
0, & x > L;
\end{cases}
\]

\[
u(x, 0) = \begin{cases} 
A e^\lambda (1 - e_0) \sinh(\lambda x), & x < L, \\
B e^\lambda (1 - e_0) e^{-\lambda x}, & x > L;
\end{cases}
\]

\[
\epsilon(x, 0) = \begin{cases} 
e_0 + A e \sinh(\lambda x), & x < L, \\
e_0 - B e^{-\lambda x}, & x > L.
\end{cases}
\]

The initial conditions for \(\sigma\) and \(\psi\) can be calculated from the conditions given above.

Importantly, we note that not all of the parameters given in the equations above are independent. In particular, we recall that

\[
n_0 = \frac{a_n \pi_n^2}{\theta_{nn} \pi_n^2} + a_n \sigma \frac{E_0 \zeta_n \eta_n}{\theta_{nn} \pi_n^2},
\]

\[
\rho_0 = \frac{\eta_0}{\pi_n},
\]

\[
e_0 = \frac{\epsilon_0}{\pi_n^2},
\]

\[
\lambda = \sqrt{\frac{b \rho_0}{E_0 (1 - e_0)}},
\]

\[
A e = \frac{e_0 (1 - \rho_{\text{wound}}) + \beta_n \eta_0 \rho_0}{\sinh(\lambda L) + \rho_{\text{wound}} \rho_0 \cosh(\lambda L)},
\]

\[
B e = \frac{\left(e_0 (1 - \rho_{\text{wound}}) + \beta_n \eta_0 \rho_0\right) e^{\lambda L}}{1 + \frac{\rho_{\text{wound}} \rho_0}{\eta_0} \coth(\lambda L)}.
\]

We now wish to construct an appropriate nondimensionalisation for this system. That is, we wish to find scalings \(\hat{n}, \hat{m}, \hat{\rho}, \text{ etc.}\); so that \(\hat{n}\) is a characteristic fibroblast density, \(\hat{m}\) is a characteristic myofibroblast density and so on.

To begin with, we note that fibroblasts convert directly into myofibroblasts; thus it is convenient for us to use the same nondimensionalisation for both species. Scaling the cell densities with respect to the healthy equilibrium fibroblast density as an appropriate scale, this yields the result that

\[
\hat{n} = \hat{m} = n_0 = \frac{a_n \pi_n^2 + a_n \sigma \frac{E_0 \zeta_n \eta_n}{\theta_{nn} \pi_n^2}}{\theta_{nn} \pi_n^2}.
\]
Similarly, we can use the healthy ECM density as a scaling for $\rho$:

$$\hat{\rho} = \rho_0 = \frac{\eta_n}{\pi_n}.$$  

For the growth factor concentration, we choose to scale with respect to the equilibrium concentration that would arise if $\rho < \rho_{\text{crit}}$ at all times.\(^{45}\) Hence,

$$\hat{c} = \frac{a_c}{\theta_c}$$

Now, we note that displacement has units of length and strain is already dimensionless. Although it is possible to use a nondimensionalisation that reflects our expectation that strain and displacement will both be small, we find that this is unnecessary at this stage.\(^{46}\) Thus, we scale $u$ according to the characteristic length scale of the model (i.e. the initial length of the wound) and we leave $e$ unchanged.\(^{47}\) That is,

$$\hat{u} = \hat{x} = L; \quad \hat{e} = 1.$$  

The variables $\sigma$ and $\psi$ both have units of stress, so it is appropriate for us to use the same scaling for both of these. Since $e$ will not generally be of order unity, we find that the scaling obtained from the $\psi$ equation will be more characteristic of stress than the scaling obtained from the $\sigma$ equation. Thus,

$$\hat{\sigma} = \hat{\psi} = \beta_n \hat{n} \hat{\rho} = \frac{\beta_n \eta_n \left( a_n \pi_n^{-2} + a_n \sigma E_0 \zeta_n \eta_n \right)}{\theta_{nn} \pi_n^3}.$$
Lastly, we require a scaling for time. There are many time-dependent processes in the model that could be used to nondimensionalise time. Following the example of previous modellers, we use the fibroblast motility. Thus,

\[
\hat{t} = \frac{\hat{x}^2}{\mu_n} = \frac{L^2}{\mu_n},
\]

Since velocity is distance divided by time, this yields the result that

\[
\hat{v} = \frac{\mu_n}{L}.
\]

Having determined these scalings, we now introduce dimensionless variables \( n^* \), \( m^* \), etc.; these are defined so that

\[
n = \hat{n} n^*, \quad m = \hat{m} m^*, \quad \text{etc.}
\]

We note that our original model involved 27 independent parameters. Also, we note that the nondimensionalisation that we propose involves 6 independent scalings. Hence, we expect our nondimensionalised model to involve 21 independent dimensionless parameters. An appropriate set of dimensionless parameters is listed below:\(^{48}\)

\[
\chi_n^* = \frac{\chi_n a_c}{\rho_n \theta_c}, \quad a_n^* = \frac{a_n L^2}{\mu_n}, \quad a_{n\sigma}^* = \frac{a_{n\sigma} \eta_n \eta_n L^2}{\pi_n \mu_n}, \quad \gamma^* = \frac{\gamma \beta_n \eta_n \eta_n a_c L^2}{\pi_n \theta_c \mu_n},
\]

\[
a_{n\sigma}^* = \frac{a_{n\sigma} \beta_n \eta_n \eta_n L^2}{\pi_n \mu_n}, \quad a_{m\sigma}^* = \frac{a_{m\sigma} \eta_n \eta_n L^2}{\theta_c \mu_n}, \quad \theta_m^* = \frac{\theta_m L^2}{\mu_n}, \quad \theta_m^* = \frac{\theta_m \eta_n \eta_n L^2}{\mu_n},
\]

\[
k_m^* = \frac{\pi_n \eta_n L^2}{\mu_n}, \quad \eta_m^* = \frac{\eta_m}{\eta_n}, \quad \pi_m^* = \frac{\pi_m}{\pi_n}, \quad b^* = \frac{b L^2}{\beta_n \eta_n},
\]

\[
E_c^* = \frac{E_0 \pi_n}{\beta_m \eta_n \eta_n}, \quad \beta_m^* = \frac{\beta_m}{\beta_n}, \quad k_c^* = \frac{\zeta_n \eta_n \eta_n L^2}{\pi_n \mu_n}, \quad \zeta_m^* = \frac{\zeta_m}{\zeta_n},
\]

\[
\mu_c^* = \frac{\mu_c}{\mu_n}, \quad a_c^* = \frac{a_c L^2}{\mu_n}, \quad \beta_{\text{crit}}^* = \frac{\beta_{\text{crit}} \pi_n}{\eta_n}, \quad \rho_{\text{crit}}^* = \frac{\rho_{\text{crit}} \pi_n}{\eta_n}, \quad \rho_w^* = \frac{\rho_{\text{wound}} \pi_n}{\eta_n},
\]

\[
c_{\text{wound}}^* = \frac{c_{\text{wound}} \theta_c}{a_c}.
\]

\(^{48}\)Note that \( n_0 = \frac{a_n \pi_n \eta_n + a_{n\sigma} E_0 \zeta_n \eta_n}{\pi_n \eta_n} \) is dependent on other parameters.
In addition to these independent parameters, it is convenient for us to define some further dimensionless parameters that are dependent on those listed above. These are given below:

\[ \lambda^* = L \sqrt{\frac{b \rho_0}{E_0 (1 - e_0)}} = \sqrt{\frac{b^*}{E^* (1 - \frac{k^*}{E^*_p})}}, \]

\[ A^* = \frac{k^*_p (1 - \rho^*_w) + 1/E^*}{\sinh(\lambda^*) + \rho^*_w \cosh(\lambda^*)}, \]

\[ B^* = \frac{k^*_p (1 - \rho^*_w) + 1/E^* e^{\lambda^*}}{1 + \rho^*_w \coth(\lambda^*)}. \]

Using dimensionless variables and parameters but omitting stars, we obtain the model given below.

Equations:

\[ \frac{\partial n}{\partial t} = \frac{\partial}{\partial x} \left( \frac{\partial n}{\partial x} - \chi_n \frac{\partial c}{\partial x} - n v \right) + a_n (1 - n) n + a_{n\sigma} (\sigma^+ - E \frac{k^*_p}{E^*_p} n) n - \gamma \sigma^+ c n; \] (6.59)

\[ \frac{\partial m}{\partial t} = \frac{\partial}{\partial x} (-m v) + \gamma \sigma^+ c n + (a_{m\sigma} \sigma^+ + a_{mc} c - \theta_m - \theta_{mm} m) m; \] (6.60)

\[ \frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x} (-\rho v) + k_p (n + \eta_m m) - k_p (n + \pi_m m) \rho; \] (6.61)

\[ \frac{\partial c}{\partial t} = \frac{\partial}{\partial x} (\mu_c \frac{\partial c}{\partial x}) + a_c H(\rho_{crit} - \rho) - a_c c; \] (6.62)

\[ v = \frac{\partial u}{\partial t}; \] (6.63)

\[ u = \frac{\partial}{\partial x} \left( \frac{\sigma + \psi}{b \rho} \right); \] (6.64)

\[ \sigma = E \rho e; \] (6.65)

\[ \psi = \rho (n + \beta_m m); \] (6.66)

\[ \frac{\partial e}{\partial t} = \frac{\partial}{\partial x} ((1 - e) v) - \frac{k_p e (n + \eta_m m)}{\rho} + k_e \rho (n + \zeta_m m). \] (6.67)
Boundary conditions:

\[ n_x(0, t) = 0, \quad n_x(\infty, t) = 0; \]
\[ c_x(0, t) = 0, \quad c_x(\infty, t) = 0; \]
\[ u(0, t) = 0, \quad u_x(\infty, t) = 0. \]

Initial conditions:

\[ n(x, 0) = \begin{cases} 0, & x < 1, \\ 1, & x > 1; \end{cases} \quad m(x, 0) = 0; \]
\[ \rho(x, 0) = \begin{cases} \rho_w, & x < 1, \\ 1, & x > 1; \end{cases} \quad c(x, 0) = \begin{cases} c_w, & x < 1, \\ 0, & x > 1; \end{cases} \]
\[ u(x, 0) = \begin{cases} \frac{A}{\lambda \left(1 - \frac{k_c}{k_p}\right)} \sinh(\lambda x), & x < 1, \\ \frac{B}{\lambda \left(1 - \frac{k_c}{k_p}\right)} e^{-\lambda x}, & x > 1; \end{cases} \quad e(x, 0) = \begin{cases} \frac{k_c}{k_p} + A \sinh(\lambda x), & x < 1, \\ \frac{k_c}{k_p} - B e^{-\lambda x}, & x > 1. \end{cases} \]

6.4.5 Preliminary observations and further work

We have not yet conducted a thorough analysis of the model stated above, nor have we performed numeric simulations. In this section, we present a plan for further investigation of our model.
Caricature models

The model that we have proposed is very complex and there is a high level of interdependence between the model species. This limits the amount of practical information that can be obtained using classical techniques of model analysis. For example, steady state analysis of the full model is theoretically possible but it does not yield many results that are meaningful and interpretable. In order to gain a better understanding of the general behaviour of a morphoelastic model of dermal wound healing, it is useful to construct and analyse simplified caricature models.

Moreover, we note that we will potentially encounter problems when trying to perform numeric simulations of our model. Given the unusual nature of the system that we wish to solve, it is difficult to determine which numeric approach will yield the most accurate results. Thus, it is desirable for us to construct analogous ‘toy’ problems with well-characterised solutions that can be used to test our numeric methods.

One possible caricature model to consider would be the morphoelastic equivalent of the base Tranquillo-Murray model. Such a model would not include myofibroblasts or growth factors and the effect of mechanical stress on fibroblasts would be ignored. This would allow us to investigate the different effects of traction stress (as represented by $E$) and ‘permanent contraction’ (as measured by $k_\zeta$). Furthermore, the steady states of such a model should be more easily characterised than the steady states of the full model.

If further simplifications are made (e.g. removing advection and fibroblast proliferation), it may be possible to obtain an exact or approximate solution to the caricature model. This solution could be compared with numeric simulations in

\[^{49}\text{To illustrate this, consider the problem of finding a spatially homogeneous steady state where } n \text{ and } c \text{ are both nonzero. Even if we make the simplifying assumption that } \theta_{mm} = 0, \text{ it can be shown that finding this steady state requires us to solve a quartic equation (not given here). Analysing the stability of known steady states is even more difficult, particularly when spatially nonhomogeneous perturbations are considered.}\]
order to validate our numeric methods. Furthermore, we could compare the analytic behaviour of a morphoelastic model with the general viscoelastic behaviour described in Section 3.3 and Appendix A. This would enable us to formally characterise the similarities and differences between our approach to mechanochemical modelling and previous approaches.

**Numeric simulations**

As described above, numeric simulation of the full model presents some challenges. In Section 6.4.2, we outlined two possible approaches to solving the model numerically. One approach involves numerically differentiating \( u \) to obtain \( v \), while the second approach involves eliminating \( u \) and constructing a differential equation in space that can be solved for \( v \). We hypothesise that the first approach will lead to greater errors; this could be confirmed by testing both approaches on a well-characterised ‘toy’ problem.

If both approaches yield acceptable results for the ‘toy’ problem, it would seem sensible to use both approaches to perform simulations of the full model. Thus, the two numeric approaches can be used to validate each other; if both approaches yield similar results, it is likely that they are both close to the actual solution.

**Possible alterations to the model**

Once numeric simulations have been obtained and validated, it may be necessary to make alterations to the model. There are a number of assumptions that we have made that might cause our model to be inaccurate. Most notably, we have neglected to include any cellular agents that contribute to the inflammatory response; the growth factor is simply generated in regions where the ECM density is low. Although this is appears to be an appropriate simplification, it may be necessary to extend our model to include macrophages and endothelial
6.5 Further applications of morphoelasticity

In this chapter, we have presented three very different applications of morphoelasticity theory: viscoelastic fluid flow, tissue growth and wound contraction. It is clear from this work that morphoelasticity can be applied to a wide variety of different problems; any physical phenomenon that involves plastic flow or evolving residual stresses can be described using a morphoelastic model. To conclude this chapter, we discuss some more of the many possible applications of morphoelasticity.

In Section 6.2, we described a morphoelastic equivalent of Maxwell’s model of a one-dimensional viscoelastic fluid. It would be interesting to use the tensorial strain-evolution equations described in Section 5.4 to develop a three-dimensional equivalent of this theory. We would expect the three-dimensional morphoelastic model to be similar to the upper convected model of viscoelastic flow developed by Oldroyd (see, for example, Bird et al. [20]). However, our experience with the one-dimensional model means that we also expect that there will be some interesting differences when the displacement gradient is large. Experimental work is needed in order to determine which model is more appropriate in these cases.

Another interesting problem that could be investigated using a morphoelastic model is the contraction of collagen lattices by fibroblasts. In Section 2.3.3, we discussed the fact that a variety of experimental researchers have used fibroblast populated collagen lattices (FPCLs) to investigate the mechanical behaviour of fibroblasts. While some mathematical models of lattice contraction have been developed, they struggle to explain the permanent contraction that remains after
fibroblasts are killed. This phenomenon could easily be explained and investigated using a morphoelastic model.

In addition to problems of biological growth and contraction, there are a variety of industrial problems that involve evolving residual stresses. For example, wood undergoes changes to its fundamental mechanical structure during the process of oven drying [147]. In some cases, these changes lead to undesirable outcomes like board warping. Our theory of morphoelasticity could be used to develop improved models of the mechanical processes that lead to warping. Such models could be used to develop new methods for making the drying process more effective.

Another industrial application that could be considered is classical metal plasticity. As described by Xiao et al. [196], a variety of mathematical models can be used to describe elastoplastic deformation at finite strains. It would be interesting to develop a morphoelastic model of metal plasticity and compare it with other models that have been developed.

As discussed, our theory of morphoelasticity can be used in a variety of different applications. However, there is a fundamental problem that needs to be addressed before morphoelastic modelling is truly practical. Having constructed a general model, there is now a need for accurate and efficient numeric methods that can be used to solve the systems of equations that we have proposed. We note that it is unusual to define elastic strain using an evolution equation. As a result, conventional numeric methods for solving problems in finite elasticity are unlikely to be useful. Moreover, the strain evolution equations that we have developed are distinctly different from classical conservation equations. This means that conventional numeric methods for solving advection problems will probably also struggle.50

In conclusion, morphoelasticity is a useful theoretical framework for describing the evolution of residual strains. However, further work is needed in order to

50 As described above, problems also arise because of the fact that the force balance equation is independent of velocity. See Section 3.4.2 for a discussion of one way of resolving this issue in a one-dimensional model.
confirm that the predictions obtained from morphoelastic models are consistent with physical observations. Furthermore, developing appropriate numeric algorithms to solve morphoelastic problems requires careful thought, both for the three-dimensional case and the one-dimensional case.
Chapter 7

Conclusions and further work

Over the past fifty years, our understanding of human biology on the microscale has moved forward in leaps and bounds. With the advent of electron microscopy, modern spectroscopy, genomics and proteomics, we have learnt more and more about cells and the chemical signals that they use for communication. However, the revolutionary progress achieved in cell biology and biochemistry is yet to be matched by similar progress in biomechanics. Despite a growing awareness of its importance, the mechanical interaction of cells with their surrounding environment is still poorly understood.

One reason for this is that it is much more difficult to measure the mechanical stresses experienced by a cell than it is to study a cell’s chemical environment. Although a variety of methods for detecting biochemicals have been developed, it is currently impossible to measure mechanical stresses directly. Instead, a tissue’s stress distribution must be inferred from observations of displacement.

The process of using displacement data to obtain information about internal stresses depends on effective mathematical modelling. As a result, there is a present need for practical mathematical models that can be used to describe the mechanical behaviour of biological tissues. Improved models of tissue mechanics
could potentially lead to more effective treatments for severe scarring and other disorders. Furthermore, mechanical modelling could be used to investigate such diverse biological phenomena as tumour growth, aneurysm formation and embryo development. Without mathematical models, however, biomechanics will never become a quantitative and predictive science.

In this thesis, we have developed a powerful and flexible theory that can be used to describe the mechanical behaviour of biological solids. Building on current models of biological ‘plasticity’ and morphoelasticity, we have constructed evolution equations for the effective strain that can be incorporated into general mathematical models. These equations have the potential to be used in a wide range of biomechanical applications and in other fields.

This work was prompted by our investigation of the mechanochemical models of dermal wound healing in the present literature. As described in Chapter 3, viscoelastic models of dermal wound healing yield physically implausible oscillations in the displacement. This indicates that different assumptions are necessary. Although the skin may behave as a viscoelastic solid on a short time scale, this does not account for the gradual ‘flow’ observed during wound healing. This can be better described using some form of plasticity.

However, classical plasticity theory is poorly suited to modelling the continual growth and contraction that occurs in biological tissues. Instead, we required a theory of ‘morphoelasticity’ that would allow us to construct mechanical models in which the evolution of the zero stress state is followed over time. A number of mathematical frameworks for morphoelasticity can be found in the present literature but some of these depend on cavalier assumptions about the process of morphoelastic change and others are too complex to be practical and interpretable. Thus, we found it appropriate to develop a new theory that builds on the work to date.

Chapters 4 and 5 form the core of this thesis; they contain the mathematical development of our theory of morphoelasticity. In Chapter 4, we reviewed the
use of a multiplicative decomposition of the deformation gradient to describe the zero stress state of a body. This led us to formally define the effective strain in terms of the zero stress deformation gradient. In Chapter 5, we then considered how changes to the zero stress state should be modelled. Although the ambiguous nature of the zero stress state made this difficult, we were able to introduce the concept of the ‘growth tensor’ as a measure of the rate at which the zero stress state changes. Moreover, we found that we could use this growth tensor to develop evolution equations for the effective strain in both one dimension and three dimensions.

Lastly, in Chapter 6 we developed a number of applications of this general theory. In doing this, we demonstrated that morphoelasticity is a flexible theory that yields results that are consistent with our physical expectations. However, much further work is needed in order to explore the potential of morphoelastic modelling. Most importantly, there is a need for appropriate numeric methods that can be used to obtain simulations of the models presented. Although we present three different applications of morphoelasticity in Chapter 6, we are not yet at the stage of producing numeric simulations for any of them.

From a theoretical perspective, one promising avenue of further investigation is the area of multiphase modelling. Biological tissues are not homogeneous solids. Instead, they are porous materials that contain a mixture of cells, fibres and water. It would be interesting to attempt to incorporate morphoelasticity into a simple multiphase model (e.g. a model of collagen lattice contraction). However, it is unclear as to how the zero stress state should be defined in such a situation; possibly, each solid phase would need its own separate zero stress state. Further investigation of this area is needed in order to explore the possibilities.

In conclusion, mathematical modelling is a necessary part of biomechanics and the theory of morphoelasticity presented in this thesis provides a useful framework for constructing biomechanical models. By developing an interpretable and concise representation of biological plasticity, we have increased the opportunity for mathematical techniques to be used to investigate biological problems. Future
work in this area can only serve to strengthen Cohen’s confident assertion:

\[
\text{Mathematics is biology’s next microscope, only better;}
\]
\[
\text{biology is mathematics’ next physics, only better.}
\]

Cohen (2004) [32], p2017
Appendices

A  Asymptotic analysis of a general viscoelastic wound healing model

In Section 3.3.2, we showed that the caricature Tranquillo-Murray model developed in Section 3.3.1 exhibits analytic oscillations in displacement at large values of $x$. We will now show that this behaviour can be generalised to all models of dermal wound healing with a similar construction, including the full Tranquillo-Murray model and the extensions proposed by Olsen and coworkers.

Thus, let us consider a general model of dermal wound healing of the form

$$\frac{\partial a_1}{\partial t} + \frac{\partial}{\partial x} (a_1 \frac{\partial u}{\partial t}) = A_1(a_1, a_2, \ldots),$$

$$\frac{\partial a_2}{\partial t} + \frac{\partial}{\partial x} (a_2 \frac{\partial u}{\partial t}) = A_2(a_1, a_2, \ldots),$$

$$\ldots$$

$$\frac{\partial^3 u}{\partial x^2 \partial t} + E \frac{\partial^2 u}{\partial x^2} + \tau \frac{\partial}{\partial x} (\psi(a_1, a_2, \ldots)) = s B(a_1, a_2, \ldots) u; \quad (A.1)$$

where $u$ represents displacement and each $a_i$ represents the concentration of some species relevant to wound healing. For example, $a_1$ could represent fibroblast density, $a_2$ could represent ECM density, $a_3$ could represent the concentration of a
growth factor and so on. The model also includes unspecified functions (described below) and two dimensionless parameters, $\tau$ and $s$. As before, $\tau$ represents the relative magnitude of the cellular tractional stresses and $s$ represents the magnitude of the tethering force. Later in our analysis, we will assume both of these to be small.

The functions $A_i$ describe how the concentrations represented by $a_i$ vary over time due to random motion, growth, conversion between species etc. Importantly, we expect all of the $a_i$ equations to be either reaction-advection equations or reaction-advection-diffusion equations. Thus, the functions $A_i$ may depend on the spatial derivatives of $a_1$, $a_2$ etc., but not on their time derivatives. The function $\psi$ represents the traction stresses exerted by the cells, while $B$ represents the dependence of the tethering forces on the ECM density and related species. Both $\psi$ and $B$ are taken to depend only on $a_1$, $a_2$ etc., not on temporal or spatial derivatives of these variables.

We wish to solve these equations subject to relevant initial and boundary conditions. The conditions on displacement are the same as those used previously:

$$u(x, 0) = 0, \quad u(0, t) = 0, \quad \lim_{x \to \infty} \frac{\partial u}{\partial x} = 0. \quad (A.2)$$

Furthermore, the symmetry of our linear wound requires that the diffusive flux of all randomly moving species $a_i$ is zero at the wound centre.\footnote{Note that not all of the species represented by the variables $a_i$ necessarily undergo random motion. For example, the ECM density in the Tranquillo-Murray model does not diffuse and therefore does not require a boundary condition at $x = 0$.} Thus,

$$\frac{\partial a_i}{\partial x}(0, t) = 0$$

in all cases where $A_i$ involves $\frac{\partial^2 a_i}{\partial x^2}$. Similarly, we require that all of the dependent variables of the model tend to some value as $x$ approaches infinity. That is,

$$\lim_{x \to \infty} \frac{\partial a_i}{\partial x} = 0$$
for all \( a_i \).

Now consider the initial conditions for the variables \( a_i \). Following the example of the Tranquillo-Murray model, we assume that these will all take the form

\[
a_i(x, 0) = \begin{cases} 
\alpha_i^{\text{wound}}, & 0 < x < 1, \\
\alpha_i^{\text{healthy}}, & x > 1.
\end{cases}
\]

That is, the concentration of each species is initially constant inside the wound and constant outside the wound, but these two concentrations may be different. Since the healthy dermis should not undergo major changes unless a significant disruption takes place, we also require that the \( x > 1 \) initial condition corresponds to a stable, spatially homogeneous steady state of the system.\(^2\)

Now, we wish to apply the same analytic techniques as described in the previous sections. Firstly, let us assume that \( \tau \) is small. Considering the leading-order problem associated with (A.1) and applying the initial and boundary conditions given in (A.2) we find that \( u_0(x, t) \equiv 0 \). Thus, the leading-order \( a_i \) equations become a conventional reaction-diffusion system with no dependence on displacement. If we solve this system, we obtain leading-order solutions for \( a_i \) which we will denote \( a_i^{(0)} \).

Next, we wish to find the first nonzero term in the solution for \( u(x, t) \). Thus, let \( \psi_0(x, t) \) and \( B_0(x, t) \) be defined so that

\[
\psi_0(x, t) = \psi(a_1^{(0)}, a_2^{(0)}, \ldots),
\]

and

\[
B_0(x, t) = B(a_1^{(0)}, a_2^{(0)}, \ldots).
\]

\(^2\)As discussed by Tranquillo and Murray [189], the requirement of stability is not trivial. For certain parameter values, the steady state that we would use to represent the healthy dermis is unstable to small perturbations. Since this is biologically unreasonable, we will assume that the parameter values have been chosen to ensure that the \( x > 1 \) initial condition corresponds to a stable steady state.
Given these definitions, we find that the first order problem obtained from (A.1) becomes
\[
\frac{\partial^3 u_1}{\partial x^2 \partial t} + E \frac{\partial^2 u_1}{\partial x^2} - s B_0(x, t) u_1 = -\frac{\partial \psi_0}{\partial x}.
\] (A.3)

A general analysis of this equation is not easy. Instead, we will focus on considering the case where \( t \) is sufficiently small for viscous effects to dominate elastic effects. In this case, it is possible to show that the solution to (A.3) exhibits oscillations similar to those described in Section 3.3.2.

This assumption of small times is very important. From the analysis described here, we can only be confident that oscillations will occur on the short time scale where dermal viscosity is more important than dermal elasticity. Although it is possible for the oscillations to continue at longer times, we have not been able to show that this is the case. Thus, our work suggests that the time scale of the displacement oscillations is linked to the time scale of viscous effects.

Before we begin our formal analysis, it is important to make some observations about the short time behaviour of \( B_0 \) and \( \psi_0 \). Firstly, we note that there should be no significant change at short times to the \( a_i \) concentrations far from the wound. This is a natural consequence of the fact that the \( x > 1 \) initial condition should correspond to a stable steady state of the model. Thus, \( B_0 \) and \( \psi_0 \) should rapidly tend to constant values as \( x \) approaches infinity. In fact, without loss of generality we will assume that \( B_0 \to 1 \) as \( x \to \infty \) and redefine \( s \) appropriately.

Secondly, we note that \( \psi_0 \) should undergo a sudden and significant change in the vicinity of \( x = 1 \). This is a result of the fact that \( \psi_0 \) represents the traction stress applied by the cells, which is initially much larger outside the wound than inside it. Effectively, it is this jump in the value of \( \psi_0 \) that drives the movement of the ECM.

---

3Interestingly, we note that the ‘healthy’ steady state may not be the only stable steady state of the model. It is possible that the wound will disrupt the system so that the long time equilibrium is different from the initial \( x > 1 \) equilibrium. For example, it is possible for there to be a ‘keloidal’ steady state, corresponding to a scar that spreads outside its original boundaries. However, in the case where \( t \) is small and \( x \) is large, this will not be important.
Let us now consider the case where \( t \sim O(\epsilon) \) and \( 0 < \epsilon \ll 1 \). We assume that \( \epsilon \) has been chosen so that \( E \epsilon \) and \( s \epsilon \) are both small. For a given \( \epsilon \), we can define a rescaled time variable, \( T \), so that \( T \sim O(1) \) and \( t = \epsilon T \). Substituting for \( t \), equation (A.3) becomes

\[
\frac{\partial^3 u_1}{\partial x^2 \partial T} + E \epsilon \frac{\partial^2 u_1}{\partial x^2} - s \epsilon B_0(x, \epsilon T) u_1 = -\epsilon \frac{\partial \psi_0}{\partial x}.
\] (A.4)

Equation (A.4) is very similar to equation (3.39) and we expect to be able to use matched asymptotics to find an approximate solution for \( u_1 \). In order to do this, we first need to determine the scalings of \( x \) that correspond to the distinguished limits of (A.4).

Now, since \( E \epsilon \) is small, we find that \( \frac{\partial^3 u_1}{\partial x^2 \partial T} \) is much larger than \( E \epsilon \frac{\partial^2 u_1}{\partial x^2} \), regardless of any spatial rescaling. That is, during the short time scale that we are considering, viscous forces completely dominate elastic forces. Similarly, when \( x \sim O(1) \), we find that \( \frac{\partial^3 u_1}{\partial x^2 \partial T} \) is much larger than \( -s \epsilon B_0(x, \epsilon T) u_1 \). Thus, viscous forces also dominate tethering forces when \( t \) is small and \( x \) is not large.

In contrast, we recall that \( \psi_0 \) changes abruptly at \( x = 1 \). It follows from this that \( \frac{\partial \psi_0}{\partial x} \) is very large around \( x = 1 \) and we will find that the dominant balance of (A.4) for \( x \sim O(1) \) takes the form

\[
\frac{\partial^3 u_{1, \text{inner}}}{\partial x^2 \partial T} = -\epsilon \frac{\partial \psi_0}{\partial x},
\] (A.5)

where \( u_{1, \text{inner}} \) represents the leading order term in the asymptotic expansion of the inner solution to (A.4).

However, as \( x \) grows large, we find that \( \frac{\partial \psi_0}{\partial x} \to 0 \) and \( B_0 \to 1 \). Thus, we can construct a new dominant balance where \( x \sim O((s \epsilon)^{-\frac{1}{2}}) \). Letting \( X = x \sqrt{s \epsilon} \) represent the rescaled spatial variable, we find that the dominant balance of (A.4) for large \( x \) takes the form

\[
\frac{\partial^3 u_{1, \text{outer}}}{\partial X^2 \partial T} - B_0(X (s \epsilon)^{-\frac{1}{2}}, \epsilon T) u_{1, \text{outer}} = 0,
\]
where \( u_{1}^{\text{outer}} \) represents the leading order term in the outer expansion. Furthermore, since \( B_0 \to 1 \) when \( x \) is large and \( t \) is small, this is effectively equivalent to

\[
\frac{\partial^3 u_{1}^{\text{outer}}}{\partial X^2 \partial T} - u_{1}^{\text{outer}} = 0. \tag{A.6}
\]

As we did in Section 3.3.2, we will use Prandtl’s matching condition to obtain valid solutions for \( u_{1}^{\text{inner}} \) and \( u_{1}^{\text{outer}} \). Fortunately, finding a solution of (A.5) that does not diverge as \( x \to \infty \) is a simple matter of integration. That is,

\[
u_{1}^{\text{inner}}(x, T) = \epsilon \int_0^T \int_0^x \int_x^\infty \frac{\partial \psi_0(x^{**}, \epsilon T^*)}{\partial x^{**}} dx^{**} dx^* dT^*.
\]

Having found \( u_{1}^{\text{inner}} \), we can use this to specify a boundary condition on \( u_{1}^{\text{outer}} \). By Prandtl’s matching condition, we require that

\[
\lim_{X \to 0} u_{1}^{\text{outer}}(X, T) = \lim_{x \to \infty} u_{1}^{\text{inner}}(x, T).
\]

Hence, let

\[
F(T) = \lim_{x \to \infty} u_{1}^{\text{inner}}(x, T).
\]

Since we have a solution for \( u_{1}^{\text{inner}} \) in terms of \( \psi_0 \), we note that \( F(T) \) is known. Furthermore, it follows from the definition of \( u_{1}^{\text{inner}} \) that \( F(0) = 0 \) and that \( F(T) \) is continuous and smooth.

Now, we wish to solve for \( u_{1}^{\text{outer}} \). For notational convenience in the rest of this section, let \( Y(X, T) = u_{1}^{\text{outer}}(X, T) \). Thus, we intend to solve the differential equation,

\[
\frac{\partial^3 Y}{\partial X^2 \partial T} - Y(X, T) = 0, \tag{A.7}
\]
subject to the conditions,

\begin{align}
  Y(X, 0) &= 0, \quad (A.8) \\
  Y(0, T) &= F(T), \quad (A.9) \\
  Y(\infty, T) &= 0. \quad (A.10)
\end{align}

This is almost identical to the outer problem obtained from our caricature model (see (3.58) to (3.61)). The only difference is that we now have a more general left-hand boundary condition. Unfortunately, the change of boundary condition means that the Laplace transform approach no longer easily interpretable results; instead we would find our solution for \( Y(X, T) \) expressed as a convolution. Thus, we will need to use a different approach to solve the system stated above.

Now, we recall that the solution to (3.58) took the form

\[ Y_0(X, t) = f_1(t) f_2(X \sqrt{t}), \]

where \( f_1(t) \) was a polynomial and \( f_2(X \sqrt{t}) \) was a Wright function. Since we expect the solution to (A.7) to be similar, let us rewrite (A.7) so that the dependent variables are \( T \) and \( X \sqrt{T} \). Once we have a partial differential equation in this form, we will approach it by considering a separation of variables.

Hence, let \( Z = X \sqrt{T} \). In this case, we find that

\[ \frac{\partial Y}{\partial X} = \sqrt{T} \frac{\partial Y}{\partial Z}, \]

and hence

\[ \frac{\partial^2 Y}{\partial X^2} = T \frac{\partial^2 Y}{\partial Z^2}. \]

Remembering that \( Z \) is a function of \( T \), we can then differentiate with respect to
Appendix A – Analysis of a viscoelastic wound healing model

\[ \frac{\partial^3 Y}{\partial X^2 \partial T} = \frac{\partial^2 Y}{\partial Z^2} + T \frac{\partial^3 Y}{\partial Z^2 \partial T} + \frac{T X \partial^3 Y}{2\sqrt{T} \partial Z^3}, \]

or, equivalently,

\[ \frac{\partial^3 Y}{\partial X^2 \partial T} = \frac{\partial^2 Y}{\partial Z^2} + T \frac{\partial^3 Y}{\partial Z^2 \partial T} + \frac{Z \partial^3 Y}{2 \partial Z^2}. \]

Thus, our rearranged differential equation takes the form

\[ \frac{Z}{2} \frac{\partial^3 Y}{\partial Z^3} + T \frac{\partial^3 Y}{\partial Z^2 \partial T} + \frac{\partial^2 Y}{\partial Z^2} - Y(Z, T) = 0. \] (A.11)

Now, consider what happens to the boundary conditions. When \( T = 0 \), we find that \( Z = T = 0 \); when \( X = 0 \), we find that \( Z = 0 \); and when \( X \to \infty \), we find that \( Z \to \infty \). Thus, our boundary conditions on \( Y(Z, T) \) become

\[ Y(0, 0) = 0, \] (A.12)
\[ Y(0, T) = F(T), \] (A.13)
\[ Y(\infty, T) = 0. \] (A.14)

Given that equation (A.11) has triple derivatives in \( Z \) as well as single derivatives in \( T \), this would appear to be underdetermined. Since condition (A.13) implies (A.12), we only have two independent conditions when we would expect four.

However, we must still have enough information to construct a well-defined problem. We expect that it is possible to derive the two remaining boundary conditions that we need from the boundary conditions that we already have.

For example, consider the case where \( T = 0 \). Then, (A.11) becomes

\[ \frac{Z}{2} \frac{\partial^3 Y}{\partial Z^3}(0, 0) + \frac{\partial^2 Y}{\partial Z^2}(0, 0) - Y(0, 0) = 0, \] (A.15)

which is an ordinary differential equation in \( Z \).
We have two relevant boundary conditions on this equation:

\[ Y(0, 0) = 0, \]

and

\[ Y(\infty, 0) = 0. \]

Furthermore, we note that substituting \( Z = 0 \) into (A.15) yields

\[ \frac{\partial^2 Y}{\partial Z^2}(0, 0) - Y(0, 0) = 0, \]

and thus,

\[ \frac{\partial^2 Y}{\partial Z^2}(0, 0) = 0. \quad (A.16) \]

This gives us the third boundary condition that we need to solve the third-order equation.

Trivially, we see that \( Y(Z, 0) = 0 \) is a solution to (A.15). Furthermore, since (A.15) is a linear ordinary differential equation, we can be confident that this solution is unique. Thus, it follows that

\[ Y(Z, 0) = 0, \quad (A.17) \]

is an appropriate initial condition for equation (A.15).

Similarly, consider what happens when \( Z = 0 \) is substituted into equation (A.11). In this case we find that

\[ T \frac{\partial^3 Y}{\partial Z^2 \partial T}(0, T) + \frac{\partial^2 Y}{\partial Z^2}(0, T) - Y(0, T) = 0. \]

Since \( Y(0, T) = F(T) \) from (A.13), it follows that

\[ T \frac{\partial^3 Y}{\partial Z^2 \partial T}(0, T) + \frac{\partial^2 Y}{\partial Z^2}(0, T) = F(T). \]
This is an ordinary differential equation where the independent variable is $T$ and the dependent variable is $\frac{\partial^2 Y}{\partial Z^2}(0, T)$. Furthermore, we recall from (A.16) that $\frac{\partial^2 Y}{\partial Z^2}(0, 0) = 0$, giving us the initial condition that we need. Using the integrating factor method, we thus conclude that

\[
\frac{\partial^2 Y}{\partial Z^2}(0, T) = 1 \frac{1}{T} \int_0^T F(S) dS. \tag{A.18}
\]

This is the remaining condition that we need for a well-defined system. As we expected, we have three boundary conditions and one initial condition.

Now, we propose to solve (A.11) by separation of variables. Thus, let

\[Y(Z, T) = M(Z) \Theta(T),\]

where $M(Z)$ and $\Theta(T)$ are to be determined. Substituting into (A.11), this yields

\[
\frac{Z}{2} M'''(Z) \Theta(T) + T \Theta'(T) M''(Z) + M''(Z) \Theta(T) - M(Z) \Theta(T) = 0,
\]

and thus we find that

\[
\frac{-\left(\frac{Z}{2} M'''(Z) + M''(Z) - M(Z)\right)}{M''(Z)} = T \frac{\Theta'(T)}{\Theta(T)} = \lambda, \tag{A.19}
\]

where $\lambda$ is independent of $T$ and $Z$ but can otherwise be freely chosen.

Since (A.11) is a linear equation, the superposition principle means that the solution for $Y(Z, T)$ may take the form

\[Y(Z, T) = \sum_{\lambda} A_{\lambda} M_{\lambda}(Z) \Theta_{\lambda}(T),\]

where the summation is performed over relevant values of $\lambda$, as specified by the boundary and/or initial conditions.

Equation (A.19) allows us to construct separate equations for $M_{\lambda}(Z)$ and $\Theta_{\lambda}(T)$.
as follows:

\[ Z M''(Z) + 2(1 + \lambda) M'(Z) - 2M(Z) = 0, \]  
\[ \Theta'_\lambda(T) - \frac{\lambda}{T} \Theta_\lambda(T) = 0. \]  
\[ \text{(A.20)} \]
\[ \text{(A.21)} \]

Equation (A.21) clearly has a general solution of \( \Theta_\lambda(T) = k_\lambda T^\lambda \). Since \( A_\lambda \) is arbitrary, we can take \( k_\lambda = 1 \) without loss of generality. Substituting this into the summation definition of \( Y(Z, T) \), we find that

\[ Y(Z, T) = \sum_\lambda A_\lambda M_\lambda(Z) T^\lambda. \]  
\[ \text{(A.22)} \]

Now, consider the boundary and initial conditions that \( Y(Z, T) \) needs to satisfy. Substituting the definition of \( Y \) from (A.22) into (A.17), we find that

\[ 0 = Y(Z, 0) = \sum_\lambda A_\lambda M_\lambda(Z) 0^\lambda. \]

It follows from this that \( \lambda \) cannot be negative or zero.

Substituting into (A.14), we find that

\[ 0 = Y(\infty, T) = \sum_\lambda A_\lambda M_\lambda(\infty) T^\lambda, \]

which may be guaranteed if

\[ M_\lambda(\infty) = 0, \]

for all \( \lambda \).

Next, consider the left-hand boundary conditions. Substituting into (A.13), we find that

\[ F(T) = Y(0, T) = \sum_\lambda A_\lambda M_\lambda(0) T^\lambda. \]
Without loss of generality, let $M_{\lambda}(0) = 1$. Thus, we find that

$$F(T) = \sum_{\lambda} A_{\lambda} T^\lambda. \quad (A.23)$$

Since $F(T)$ is smooth, it can be represented using a Maclaurin series. Noting that $F(0) = 0$, this means that

$$F(T) = \sum_{\lambda=1}^{\infty} \frac{F^{(\lambda)}(0) T^\lambda}{\lambda!}. \quad (A.24)$$

By comparison with (A.23), this suggests that it may be appropriate to pick $\lambda$ to be the positive integers. In this case, we would find that

$$Y(Z, T) = \sum_{\lambda=1}^{\infty} A_{\lambda} M_{\lambda}(Z) T^\lambda, \quad (A.25)$$

and

$$A_{\lambda} = \frac{F^{(\lambda)}(0)}{\lambda!}. \quad (A.26)$$

Before we can be confident that this is valid, we need to ensure that we can satisfy our remaining boundary condition. From (A.18), this takes the form

$$\frac{1}{T} \int_0^T F(S) \, dS = \frac{\partial^2 Y}{\partial Z^2}(0, T) = \sum_{\lambda} A_{\lambda} M_{\lambda}''(0) T^\lambda.$$

Using the Maclaurin series representation of $F(T)$, it follows that

$$\frac{1}{T} \int_0^T F(S) \, dS = \sum_{i=1}^{\infty} \frac{F^{(i)}(0) T^i}{(i+1)!}. \quad (A.27)$$

Thus, if we assume that $Y(Z, T)$ takes the form given in (A.24), we find that

$$\sum_{\lambda=1}^{\infty} \frac{F^{(\lambda)}(0) T^\lambda}{(\lambda + 1)!} = \sum_{\lambda=1}^{\infty} \frac{F^{(\lambda)}(0) T^\lambda}{\lambda!} M_{\lambda}'(0).$$
This will be satisfied if

\[ M''(0) = \frac{1}{\lambda + 1}. \]

Interestingly, this is identical to the result that we obtain when we substitute \( Z = 0 \) and \( M(0) = 1 \) into equation (A.20). Thus, it would appear that considering only positive integer values of \( \lambda \) is consistent with all of the restrictions that we need to satisfy.

Now, we are left with the problem of solving (A.20) to find \( M_\lambda(Z) \). That is, we wish to solve the ordinary differential equation

\[ Z M''''_\lambda(Z) + 2 (1 + \lambda) M''_\lambda(Z) - 2 M_\lambda(Z) = 0, \]

subject to the boundary conditions

\[ M_\lambda(0) = 1, \quad M''_\lambda(0) = \frac{1}{1 + \lambda}, \quad \lim_{Z \to \infty} M_\lambda(Z) = 0. \]

There is no obvious way of solving this differential equation. However, based on the results obtained in Section 3.3.2, we propose that \( M_\lambda(Z) \) will take the form of a Wright function. That is, we assume that

\[ M_\lambda(Z) = \alpha \sum_{k=0}^{\infty} \frac{(-Z)^k}{k! \Gamma(\rho k + \beta)}, \]

where \( \alpha, \rho \) and \( \beta \) are constants to be determined that are dependent on \( \lambda \). Substituting this into (A.20), we find that

\[
\alpha Z \sum_{k=0}^{\infty} \frac{k(k-1)(k-2)(-Z)^k}{k! \Gamma(\rho k + \beta)} + 2 (1 + \lambda) \alpha \sum_{k=0}^{\infty} \frac{k(k-1)(-Z)^k}{k! \Gamma(\rho k + \beta)}
- 2 \alpha \sum_{k=0}^{\infty} \frac{(-Z)^k}{k! \Gamma(\rho k + \beta)} = 0,
\]
or, equivalently,

\[ \sum_{k=0}^{\infty} \left[ \frac{(k-1)(k-2)+2(1+\lambda)k(k-1)}{k! \Gamma(\rho k + \beta)} (-Z)^k \right] - 2 \sum_{k=0}^{\infty} \frac{(-Z)^k}{k! \Gamma(\rho k + \beta)} = 0. \]

Noting that \( k(k-1) = 0 \) when \( k = 0 \) or \( k = 1 \), we can combine these two sums as follows:

\[ \sum_{k=0}^{\infty} \left[ \frac{(k+2)(k+1)k+2(1+\lambda)(k+2)(k+1)}{(k+2)! \Gamma(\rho k + \beta)} - \frac{2}{k! \Gamma(\rho k + \beta)} \right] (-Z)^k = 0. \]

In order for this to be true, we require that

\[ \frac{k+2(1+\lambda)}{k! \Gamma(\rho k + \beta + 2 \rho)} - \frac{2}{k! \Gamma(\rho k + \beta)} = 0, \]

for all values of \( k \). By a simple rearrangement, this is equivalent to

\[ \frac{\Gamma(\rho k + \beta + 2 \rho)}{\Gamma(\rho k + \beta)} = \frac{k}{2} + 1 + \lambda. \]

(A.26)

Now, it is well-known that

\[ \frac{\Gamma(a+1)}{a} = a. \]

Thus, it is obvious that equation (A.26) will be satisfied if \( \rho = \frac{1}{2} \) and \( \beta = 1 + \lambda \). That is,

\[ M_\lambda(Z) = \alpha \sum_{k=0}^{\infty} \frac{(-Z)^k}{k! \Gamma(\frac{k}{2} + 1 + \lambda)} = \alpha \phi(\frac{1}{2}, 1 + \lambda; -Z), \]

(A.27)

is a valid solution to (A.20), where \( \phi(\rho, \beta; z) \) represents the Wright function, defined in (3.64).

Having obtained this solution, we need to ensure that it satisfies the boundary
conditions. Gorenflo et al. [68] give the result that, for $\rho > 0$ and $z \in \mathbb{R}$,

$$
\lim_{z \to -\infty} \phi(\rho, \beta, z) = 0.
$$

Thus, it follows that

$$
\lim_{Z \to \infty} M_\lambda(Z) = 0,
$$

meaning that the right-hand boundary condition is satisfied.

Furthermore, when $Z = 0$, we find that

$$
M_\lambda(0) = \alpha \sum_{k=0}^{\infty} \frac{(0)^k}{k! \Gamma\left(\frac{k}{2} + 1 + \lambda\right)} = \frac{\alpha}{\Gamma(1 + \lambda)}.
$$

Thus, in order to satisfy $M_\lambda(0) = 1$, we require that $\alpha = \Gamma(1 + \lambda) = \lambda!$.

Also, given that $\alpha = \lambda!$, we find that

$$
M'_\lambda(0) = \frac{\alpha}{\Gamma(2 + \lambda)} = \frac{\lambda!}{(1 + \lambda)!} = \frac{1}{1 + \lambda},
$$

as expected.

Hence, we conclude that

$$
M_\lambda(Z) = \lambda! \phi\left(\frac{1}{2}, 1 + \lambda; -Z\right)
$$

satisfies the differential equation, (A.20), and all relevant boundary conditions.

Substituting back into the definition of $Y(Z, T)$, we find that

$$
Y(Z, T) = \sum_{\lambda=1}^{\infty} F^{(\lambda)}(0) \phi\left(\frac{1}{2}, 1 + \lambda; -Z\right) T^\lambda,
$$

is a solution to equation (A.11). Returning to the original, unrescaled coordinates,
this means that

\[ u_1(x, t) \sim \sum_{\lambda=1}^{\infty} F_{\lambda} \phi \left( \frac{1}{2}, 1 + \lambda; -x\sqrt{st} \right) t^{\lambda}, \quad (A.28) \]

where the constants \( F_{\lambda} \) can be determined from the Maclaurin expansion of \( F \left( \frac{1}{t} \right) \).

Equation (A.28) is only valid when \( t \) is small and \( x\sqrt{t} \) is large.

Significantly, we note that equation (A.28) is expressed in terms of Wright functions where \( \rho > 0 \) and \( \beta \in \mathbb{R} \). Thus, the oscillating asymptotic expansion given in equation (3.65) is relevant and we can conclude that all viscoelastic models of wound healing following the Tranquillo-Murray model will exhibit some oscillations in displacement.

Importantly, we note that the oscillations arose because of the three-way interaction of the viscous, cell traction and tethering terms of the force balance equation. Cellular traction stresses and subdermal tethering of the ECM are essential features of a realistic mechanical model of dermal wound healing; as described in Section 3.2.3, models where \( \tau \) or \( s \) are set to zero are unable to describe a healing wound. However, it is perfectly plausible to construct a model of wound healing using an elastic constitutive law for the ECM instead of a viscoelastic law. Thus, the most promising method for eliminating displacement oscillations is to remove the viscous term from the force balance equation.
B Existence and uniqueness of the zero stress state

B.1 Existence of the zero stress Cauchy-Green tensor

Let \( C \) be the right Cauchy-Green tensor of a deformed body, \( \mathcal{B} \), following the definitions given in Section 4.2.4. Now, let \( \sigma \) be a measure of stress, and let the function \( w \) represent the elastic constitutive law that applies at position \( X_0 \) and time \( t \), such that \( \sigma(X_0, t) = w(C) \). If there exists a \( C \) such that \( w(C) = 0 \) (that is, if there is a choice of \( C(X_0, t) \) such that \( \sigma(X_0, t) = 0 \)), then \( C \) is called a zero stress Cauchy-Green tensor and is denoted \( \Theta(X_0, t) \).

In this appendix, we describe sufficient conditions on \( w(C) \) for \( \Theta \) to exist and be unique. It is clear that not all functions \( w \) will satisfy this; for example, if \( \sigma \) is a discontinuous function of \( C \), it is possible to go from a state of contractile stress to a state of tensile stress without ever passing through the zero stress state. However, although this is mathematically possible, we can safely assume in real applications that \( w \) will be continuous and smooth.

Even with the requirement that \( w \) is smooth, we cannot guarantee the existence of a zero stress state. For example, consider a solid with the property that the eigenvalues of \( \sigma \) at \( X = X^* \) are positive for all deformations (i.e., a nonzero force at \( X^* \) is required in order to prevent the solid from collapsing in on itself). In this case, there will be no solutions to \( w(C) = 0 \) at \( X = X^* \) and thus \( \Theta \) cannot be defined at that point.

Before we proceed, it should be noted that we will use a Lagrangian coordinate system throughout this section, since it is necessary to use the right Cauchy-Green tensor to define the zero stress state. Thus, we need to use the Piola-Kirchoff stress tensors instead of the more natural Cauchy stress tensor. In particular, \( \sigma \) will be taken to represent the second (symmetric) Piola-Kirchoff tensor through-
out this appendix. We prefer the second Piola-Kirchoff stress over other possible measures for two reasons: firstly, because the second Piola-Kirchoff tensor is symmetric (and thus has the same number of independent components as $C$); and secondly, because there is a close association between the components of $C$ and the components of $\sigma$. Thus, increasing or decreasing the strain in a given direction (i.e. changing a particular component of $C$), will tend to change the corresponding component of $\sigma$ in a similar manner.

Now, in order to establish our sufficient conditions for the existence of $\Theta$, we first note that $C$ and $\sigma$ both have six independent components. Thus, $w$ can be considered to be a mapping from $\mathbb{R}^6$ to $\mathbb{R}^6$ and solving $w(\Theta) = 0$ can be thought of as solving six simultaneous equations in six variables. Hence, one way of proving the existence of $\Theta$ would be to use the higher dimensional analogue of Bolzano’s theorem (equivalent to the Intermediate Value Theorem) described below.

**Theorem:** Let $f_1, f_2, \ldots, f_n$ be continuous real-valued functions of $x$ where $x = (x_1, x_2, \ldots, x_n)$ is in the domain $R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$. Moreover, let the functions $f_i$ have the property that $f_i < 0$ when $x_i = a_i$ and $x \in R$; and similarly $f_i > 0$ when $x_i = b_i$ and $x \in R$. We will show that there exists $\xi \in R$ such that $f_i(\xi) = 0$ for all $i = 1, \ldots, n$.

**Proof:**

Firstly consider the case where $n = 2$. For convenience, we use the notation $x$ and $y$ for the two independent variables instead of $x_1$ and $x_2$, and we use $a$, $b$, $c$ and $d$ in place of $a_1$, $b_1$, $a_2$ and $b_2$ defined above. Thus, $R = [a, b] \times [c, d]$.

Using this notation, we have two continuous functions, $f_1(x, y)$ and $f_2(x, y)$ defined on the domain $R = [a, b] \times [c, d]$ with the properties that

\[
\begin{align*}
f_1(a, y) &< 0 \quad \forall y \in [c, d], \\
f_1(b, y) &> 0 \quad \forall y \in [c, d], \\
f_2(x, c) &< 0 \quad \forall x \in [a, b], \\
f_2(x, d) &> 0 \quad \forall x \in [a, b].
\end{align*}
\]
We wish to prove that there exists at least one point \((x^*, y^*)\) such that
\[ f_1(x^*, y^*) = f_2(x^*, y^*) = 0. \]

To show that this is the case, firstly let \(x(\zeta)\) and \(y(\zeta)\) be two continuous functions of \(\zeta\) such that \(x(0) = a, x(1) = b, x(\zeta) \in [a, b] \ \forall \zeta \in [0, 1],\) and \(y(\zeta) \in [c, d] \ \forall \zeta \in [0, 1].\) That is, as \(\zeta\) varies from 0 to 1 it parametrically describes a curve \(C\) within the rectangle \(R,\) that extends from the left-hand boundary \(x = a\) to the right-hand boundary \(x = b.\) (see Figure B.1 for an illustration).

![Diagram](image_url)

Figure B.1: As shown above, \(C\) can be any curve that runs from the left-hand edge (where \(f_1\) is negative) to the right-hand edge (where \(f_1\) is positive). As a result of the intermediate value theorem, there must be a point on this curve where \(f_1 = 0.\) Moreover, this must be true for all possible choices of \(C.\)

Now, consider the function of one variable defined by \(g(\zeta) = f_1(x(\zeta), y(\zeta)).\)
As \(f_1(x, t), x(\zeta)\) and \(y(\zeta)\) are all continuous, we find that \(g(\zeta)\) is continuous.
Furthermore, we note that \(g(\zeta)\) is negative when \(\zeta = 0\) and positive when \(\zeta = 1.\)
Thus, by Bolzano’s theorem (or, equivalently, by the intermediate value theorem), we can find \(\zeta^*\) such that \(0 < \zeta^* < 1\) and \(g(\zeta^*) = 0.\) That is, there exists at least one point on the curve \(C\) such that \(f_1(x, y) = 0.\)

Importantly, this holds true for all choices of \(x(\zeta)\) and \(y(\zeta)\) with the required properties stated above. Thus, for any curve within our rectangle that extends
from the left-hand boundary to the right-hand boundary, there will exist a point on that curve where \( f_1 \) is zero. We now wish to show that this implies the existence of a continuous one-dimensional curve \( G \subset R \), such that \( G \) extends from the upper boundary \( y = c \) to the lower boundary \( y = d \) and \( f_1(x, y) = 0 \) if \( (x, y) \in G \). Effectively, \( G \) forms a ‘wall’ where \( f_1 = 0 \) so that no curve can connect \( x = a \) to \( x = b \) without either crossing the ‘wall’ or leaving the rectangle.

In order to show that \( G \) exists, we can apply the converse of Jordan’s theorem (see White [195] for a description and a proof). Firstly, let \( \Omega_R \) be the set of all points on the boundary of \( R \). Now, we take an arbitrary \( y^* \in [c, d] \) and consider two points, \( A = (a + \epsilon, y^*) \) and \( B = (b - \epsilon, y^*) \), where \( \epsilon \) is very small. Next, consider a curve \( Q \) in \( \mathbb{R}^2 \) connecting \( A \) and \( B \). If we follow \( Q \) starting from \( A \), we note that in order to get to \( B \), \( Q \) must either cross \( \Omega_R \) at some point where \( f_1 < 0 \) or it must pass through a point where \( f_1 = 0 \). Hence, by the converse of Jordan’s theorem, it follows that there exists a closed curve consisting of continuous regions of \( \Omega_R \) where \( f_1 < 0 \) and curves where \( f_1 = 0 \). Importantly, one of the curves where \( f_1 = 0 \) must connect \( y = c \) to \( y = d \) (i.e. it must connect the upper and lower parts of \( \Omega_R \)). Hence, \( G \) exists.

Now, we define the functions \( \hat{x}(\lambda) \) and \( \hat{y}(\lambda) \) such that varying \( \lambda \) from 0 to 1 is a parametric description of \( G \). Since \( G \) extends from \( y = c \) to \( y = d \), it is possible to choose \( \hat{y}(\lambda) \) such that \( \hat{y}(0) = c \) and \( \hat{y}(1) = d \). Furthermore, \( G \) is a continuous curve and therefore \( \hat{x} \) and \( \hat{y} \) are both continuous functions. Thus, we can use the intermediate value theorem as described earlier to show that there exists at least one point on \( G \) where \( f_2 = 0 \). Furthermore, we recall that \( f_1 = 0 \) at every point on \( G \) and \( G \subset R \). Thus, there exists \( (x^*, y^*) \in R \) such that \( f_1(x^*, y^*) = 0 \) and \( f_2(x^*, y^*) = 0 \), proving the required result in two dimensions.

In more than two dimensions, the approach is analogous, but somewhat more

---

1Specifically, we should choose \( \epsilon \) so that \( f_1(x, y^*) < 0 \) for all \( x \in [a, a + \epsilon] \) and similarly \( f_1(x, y^*) > 0 \) for all \( x \in [b - \epsilon, b] \). This ensures that \( G \) cannot separate \( (a + \epsilon, y^*) \) or \( (b - \epsilon, y^*) \) from \( x = a \) or \( x = b \) respectively. By using the intermediate value theorem and the well-ordered principle, it is trivial to show that \( \epsilon \) must exist with the required properties.
Figure B.2: The curve $G$ is defined so that $f_1 = 0$ at all points on $G$. Since all curves that connect $x = a$ with $x = b$ must pass through at least one point where $f_1 = 0$, we find that it is always possible to construct $G$ so that $G$ forms a 'wall' that connects $y = c$ with $y = d$. By parameterising this curve and using the intermediate value theorem, we are able to show that at least one point exists where $f_1 = 0$ and $f_2 = 0$ simultaneously.

complicated. Firstly, we use the intermediate value theorem to show that any parameterised curve in $R$ that connects $x_1 = a_1$ to $x_1 = b_1$ must pass through a point where $f_1 = 0$. We then apply the converse of multidimensional Jordan’s theorem to obtain a continuous $(n - 1)$-dimensional manifold, $G_1$, where $f_1(x) = 0 \ \forall x \in G_1$ and $G_1$ separates $R$ into two regions. We note that the boundary of $G_1$ will consist of a set of points where $x_2 = a_2$, a set of points where $x_2 = b_2$, a set of points where $x_3 = a_3$, etc. Thus, we can use the intermediate value theorem to show that any parameterised curve in $G_1$ connecting $x_2 = a_2$ to $x_2 = b_2$ must pass through a point where $f_2 = 0$. Applying the converse of Jordan’s theorem again, we can obtain a continuous $(n - 2)$-dimensional manifold, $G_2$, where $f_1(x) = 0$ and $f_2(x) = 0$ simultaneously. We repeat the process until we obtain a zero-dimensional manifold (i.e. a point) where $f_1$ through to $f_n$ are all 0.

It may be possible to generalise this theorem to the case where $R$ is not a hyperprism, but it is not intuitively obvious how to formulate the conditions on the bounding manifold so as to ensure the existence of a solution. However, the region $R$ as defined above has a natural interpretation in the context of solid mechanics.
As we mentioned earlier, it is reasonable to assume that increasing (decreasing) a particular component of $C$ would tend to increase (decrease) the corresponding component of $\sigma$. That is, if all components of the stress tensor are able to take on both positive and negative values (avoiding the situation described at the beginning), we would expect that $\sigma_{IJ}$ would be positive for sufficiently large values of $C_{IJ}$ and negative for sufficiently small values of $C_{IJ}$.

More rigorously, let us assume that $\sigma$ is a continuous function of $C$ and there exist values $C^+_{IJ}$ and $C^-_{IJ}$ such that if $C_{IJ} \in [C^+_{IJ}, C^-_{IJ}]$ holds for all $I$ and $J$, and $C_{AB}$ is equal to $C^+_{AB}$ (equal to $C^-_{AB}$), it follows that $\sigma_{IJ}$ is positive (negative).

In this case, we can apply the theorem described above to show that there will exist at least one $C$ for which $\sigma = 0$. Thus, as long as stress is a ‘relatively monotonous’ function of strain (i.e. corresponding stress and strain components tend to increase/decrease together) that satisfies the conditions described above, at least one zero stress state will exist.

However, while this gives us sufficient conditions to guarantee the existence of $\Theta$, it is clear that they are not necessary. There may exist solids for which the above conditions cannot be satisfied, and in these cases it may still be possible to find $\Theta$. Despite the fact that the result above does not apply universally to all constitutive laws, the proof above gives us some confidence that the zero stress state can be found except for solids which exhibit highly unusual stress-strain relationships.

### B.2 Uniqueness of the zero stress Cauchy-Green tensor

It is considerably more difficult to construct sufficient conditions for establishing the uniqueness of the zero stress state. Given the existence of some known zero stress state, let us define the Lagrangian strain $E^L$ using equation (4.30). We are interested in whether there exist nonzero values of $E^L$ for which $\sigma = 0$.

More generally, as long as we can construct bases for $C$ and $\sigma$ such that we can specify ranges on the independent components of $C$ with the required properties, the proof still holds.
Intuitively, we would expect the zero stress state to be unique; that is, given two neighbouring particles in the initial configuration there is a unique ‘preferred’ distance between them which corresponds to a state of zero stress.

However, for a general constitutive law, there is no clear reason why there should not be nonzero strains corresponding to zero stress. For example, increasing one component of strain might increase the corresponding component of stress, but decrease others. If this is the case for two or more different components of strain, it might be possible to attain a nonzero strain corresponding to zero stress. Similarly, if a solid is perfectly plastic (i.e. unable to sustain any stress), it follows that all possible values of $E^L$ correspond to zero stress and thus there are infinitely many zero stress states.

Despite these problems, there are a few cases where we can be confident of the uniqueness of the zero stress state. Here, we describe one of these: the case of linear elasticity. For linear elasticity, let

$$\sigma_{IJ} = C_{IJKM} E^L_{KM},$$

where $C$ is a fourth-order tensor defining the elastic behaviour of the solid. Because of the symmetry of $E^L$ and $\sigma$, we note that this can be thought of as a system of six linear equations in six unknowns.\(^3\) That is, if we construct a six-by-six matrix $C^*$ with entries obtained from the appropriate entries of the tensor $C$, finding a nontrivial zero stress state is equivalent to finding a nontrivial solution to the homogeneous equation

$$C^* \mathbf{x} = \mathbf{0}.$$  

This will only be the case if $C^*$ has at least one zero eigenvalue. Furthermore, we note that a zero eigenvalue of $C^*$ implies the existence of some vector space of strains that correspond to zero stress. That is, in order to find a nontrivial zero stress state of a linearly elastic material, there must exist at least one straining

---

\(^3\)In practical solid mechanics, there are certain restraints on $C$ relating to the symmetry of the solid and the requirement of observer-independence (see Spencer [176] for details). However, these restrictions do not affect our analysis, so we will not discuss them further.
direction in which the solid is perfectly plastic.
Appendix C – Rodriguez et al.’s Lagrangian strain

C  Rodriguez et al.’s definition of Lagrangian strain

It is interesting to contrast our definitions of strain with the approach to strain taken by Rodriguez et al. [159]. As described earlier, Rodriguez et al. expressed the deformation gradient as the product of $F_e$, an elastic deformation gradient, and $F_g$, the growth deformation gradient:

$$F = F_e F_g.$$  

They then calculated the strain by using $F_e$ in the place of $F$ in equations (4.16) and (4.17).

Now, $F_e = F F_g^{-1}$, so the Eulerian strain definition used by Rodriguez et al. takes the form

$$E^E = \frac{1}{2} (I - F_e^{-T} F_e^{-1}) = \frac{1}{2} (I - F^{-T} F_g^T F_g F_g^{-1} F^{-1}),$$

which is identical to equation (4.31) with $Z$ replaced by $F_g$. However, when we substitute $F_e$ into the definition Lagrangian strain in (4.16), we find that

$$E^L = \frac{1}{2} (F_e^T F_e - I) = \frac{1}{2} (F_g^{-T} F_g F_g F_g^{-1} - I).$$

Replacing $F_g$ with $Z$, we obtain

$$E^{L*} = \frac{1}{2} (Z^{-T} F^T F Z^{-1} - I),$$  \hspace{1cm} (C.1)

which is clearly different to the Lagrangian strain that we obtained in equation (4.30). Worryingly, this ‘Lagrangian’ strain is not uniquely defined; we note that replacing $Z$ with $\hat{Z} = QZ$ (where $Q$ is proper orthogonal) will cause $E^{L*}$ to change, despite the fact that this transformation is physically meaningless.

In order to make sense of $E^{L*}$, let us consider the case where the zero stress coordinate system is definable. That is, we assume that we can construct functions.
Appendix C – Rodriguez et al.’s Lagrangian strain

Let \( a \) and \( b \) so that
\[
X = a(\chi, t), \quad x = b(\chi, t).
\]
We recall that \( a \) and \( b \) have the physical interpretation that the elastic strain (and stress) are zero in the case where \( b \) corresponds to a rigid motion. As before, we define the zero stress deformation gradient tensor by
\[
Z = \frac{\partial a^{-1}}{\partial X} = \frac{\partial \chi}{\partial X}.
\]

Let us now consider how we would construct a definition of strain with respect to the zero stress coordinate system in a form analogous to the Lagrangian and Eulerian strains given in equations (4.28) and (4.29):
\[
(ds)^2 - (dL)^2 = 2d\chi^T E^Z d\chi.
\]
From the definition of the zero stress coordinate system, we find that
\[
(dL)^2 = d\chi^T d\chi,
\]
while
\[
(ds)^2 = dx^T dx = d\chi^T \frac{\partial X^T}{\partial \chi} \frac{\partial x^T}{\partial \chi} \frac{\partial x}{\partial X} \frac{\partial X}{\partial \chi} d\chi.
\]
Combining these with the definition of the deformation gradient and the zero stress state, we find that
\[
(ds)^2 - (dL)^2 = d\chi^T (Z^{-T} F^T F Z^{-1} - I) d\chi.
\]
By comparison with equation (C.1), we find that \( E^{L^*} = E^Z \). That is, the ‘Lagrangian’ strain definition used by Rodriguez et al. is actually formulated with respect to the zero stress coordinate system. Although Rodriguez et al. comment briefly on this fact, they do not thoroughly explore its consequences.

For example, we note that this strain formulation is not uniquely defined. Since the zero stress coordinate system does not correspond to positions in real space,
it is possible to arbitrarily rotate and translate the zero stress coordinate system without changing the physical interpretation of the result. However, any rotation of the zero stress coordinate system will directly affect the strain tensor used by Rodriguez et al. Also, there may be difficulty in interpreting $E^L*$ in the case when residual stresses are present because the zero stress coordinate system can no longer be defined consistently through space.

Lastly, and most importantly, the strain formulation used by Rodriguez et al. creates considerable problems for the definition and interpretation of the stress tensor. If the Lagrangian strain tensor that we developed (i.e. $E^L$) is used, the related stress tensors will correspond to the conventional Piola-Kirchoff stress tensors. However, if $E^Z$ is used, it would be necessary to construct novel stress tensors defined with respect to the zero stress state coordinates. Although this is theoretically possible when the zero stress coordinate system exists, it is very unclear as to how to proceed if the zero stress coordinate system cannot be consistently defined. For these reasons, we prefer the definitions of strain given in equations (4.30) and (4.31) over those developed by Rodriguez et al.
D  Finite strain evolution and related topics

In this appendix, we present the derivation of an evolution equation for finite strain. In Section D.1, we present some useful preliminary results. Interestingly, we are able to show that equation (5.44), the evolution equation for the zero stress state, is valid regardless of whether or not the principal zero stress state is used.

In Section D.2, we use equation (5.44) to obtain an equation for the evolution of finite strain. Finally, in Section D.3, we demonstrate that the equation we obtain satisfies the requirement of observer independence. That is, our equation is unaffected by arbitrary rotations of the spatial coordinate axes.

D.1  Preliminary results

Equation (5.44) states that

\[ \det(Y) \ sym \left( \frac{DY}{Dt} Y^{-1} + Y L Y^{-1} \right) = sym \left( Y G Y^{-1} \right). \]

We note that this equation involves taking the symmetric part of some tensors. This is defined so that

\[ \text{sym} \left( A \right) = \frac{1}{2} \left( A + A^T \right). \]

The symmetric part of a tensor is relatively difficult to manipulate algebraically. For example, there is generally no relationship between \( \text{sym} \left( A \right) \text{sym} \left( B \right) \) and \( \text{sym} \left( A B \right) \).

However, there is one property of this operator that is particularly useful to us:

\[ B \ sym \left( A \right) B^T = sym \left( B A B^T \right). \quad (D.1) \]
To see that this is the case, we note that

\[
\text{sym} (B A B^T) = \frac{1}{2} \left( B A B^T + (B A B^T)^T \right)
\]

\[
= \frac{1}{2} \left( B A B^T + B A^T B^T \right)
\]

\[
= \frac{1}{2} B (A + A^T) B^T
\]

\[
= B \text{ sym} (A) B^T,
\]

as required.

We can use this result to show that that (5.44) is valid for any zero stress deformation gradient, not just the principal zero stress deformation gradient. Thus, let \( Y_P \) be the principal zero stress deformation gradient and let \( Y_* \) be any other valid zero stress deformation gradient. That is, let \( Y_* \) be defined so that

\[
Y_*(x, t) = Q(x, t) Y_P(x, t),
\]

(D.2)

where \( Q \) is a proper orthogonal tensor for all values of \( x \) and \( t \) and \( Q(x, t) \) is continuous and differentiable.

Now, consider

\[
\text{det} (Y_*) \text{ sym} \left( \frac{DY_*}{Dt} Y_*^{-1} + Y_* L Y_*^{-1} \right).
\]

We wish to show that this is equal to \( \text{sym} (Y_* G Y_*^{-1}) \).

Using equation (D.2), we find that

\[
\frac{DY_*}{Dt} Y_*^{-1} + Y_* L Y_*^{-1}
\]

\[
= \frac{DQ}{Dt} Y_P Y_P^{-1} Q^T + Q \frac{DY_P}{Dt} Y_P^{-1} Q^T + Q Y_P L Y_P^{-1} Q^T,
\]
and hence,
\[
\frac{DY_s}{Dt} Y_s^{-1} + Y_s L Y_s^{-1} = S + Q \left( \frac{DY_p}{Dt} Y_p^{-1} + Y_p L Y_p^{-1} \right) Q^T,
\]
where \( S = \frac{DQ}{Dt} Q \).

We recall from Section 5.3.3 that all tensors \( S \) of this form are skew symmetric.

Hence,
\[
\text{sym} \left( \frac{DY_s}{Dt} Y_s^{-1} + Y_s L Y_s^{-1} \right) = \text{sym} \left[ Q \left( \frac{DY_p}{Dt} Y_p^{-1} + Y_p L Y_p^{-1} \right) Q^T \right]
= Q \text{sym} \left( \frac{DY_p}{Dt} Y_p^{-1} + Y_p L Y_p^{-1} \right) Q^T,
\]
using (D.1).

Thus, we find that
\[
\det \left( Y_s \right) \text{sym} \left( \frac{DY_s}{Dt} Y_s^{-1} + Y_s L Y_s^{-1} \right)
= \det \left( Y_p \right) Q \text{sym} \left( \frac{DY_p}{Dt} Y_p^{-1} + Y_p L Y_p^{-1} \right) Q^T,
\]
since \( \det \left( Y_s \right) = \det \left( Q \right) \det \left( Y_p \right) = \det \left( Y_p \right) \).

This yields the result that
\[
\det \left( Y_s \right) \text{sym} \left( \frac{DY_s}{Dt} Y_s^{-1} + Y_s L Y_s^{-1} \right) = Q \text{sym} \left( Y_p G Y_p^{-1} \right) Q^T
= \text{sym} \left( Q Y_p G Y_p^{-1} Q^T \right)
= \text{sym} \left( Y_s G Y_s^{-1} \right),
\]
as required. This establishes the fact that equation (5.44) does not depend on using the principal zero stress deformation gradient.
D.2 Evolution of finite strain

As given in equation (4.33), the Eulerian finite strain is uniquely defined by

\[ E = \frac{1}{2} (I - Y^T Y), \]

where \( Y \) is any Eulerian zero stress deformation gradient.

Now, consider the material derivative of finite strain:

\[
\frac{DE}{Dt} = -\frac{1}{2} \left( \frac{DY^T}{Dt} Y + Y^T \frac{DY}{Dt} \right) = -\text{sym} \left( Y^T \frac{DY}{Dt} \right).
\]

Using (D.1) it follows that

\[
Y^{-T} \frac{DE}{Dt} Y^{-1} = -Y^{-T} \text{sym} \left( Y^T \frac{DY}{Dt} \right) Y^{-1} = -\text{sym} \left( \frac{DY}{Dt} Y^{-1} \right).
\]

From equation (5.44), we recall that

\[
\text{sym} \left( \frac{DY}{Dt} Y^{-1} \right) = -\text{sym} \left( Y L Y^{-1} - \frac{1}{\det Y} Y G Y^{-1} \right).
\]

Thus,

\[
\frac{DE}{Dt} = Y^T \text{sym} \left( Y L Y^{-1} - \frac{1}{\det Y} Y G Y^{-1} \right) Y = \text{sym} \left( Y^T Y L - \frac{1}{\det Y} Y^T Y G \right).
\]
From the definition of $E$, we note that

$$Y^T Y = I - 2E.$$ 

Furthermore, it is possible to show that

$$\det Y = \sqrt{\det (I - 2E)},$$

by using the fact that $\det (AB) = \det (A) \det (B)$.

This yields the following evolution equation for finite strain:

$$\frac{DE}{Dt} = \text{sym} \left( (I - 2E) L - \frac{1}{\sqrt{\det (I - 2E)}} (I - 2E) G \right).$$

### D.3 Observer independence

Finally, it is interesting to note that both of the strain evolution equations (i.e. equations (5.57) and (5.64)) satisfy the requirement of observer independence. In order to be observer-independent, it must be possible to apply a time-varying proper rigid transformation to the spatial and/or material coordinate systems without affecting the validity of the equations (see, for example, Gurtin [85]). In our case, there is no explicit dependence on the material coordinate system, so we only need to consider a rotation of the spatial coordinate system.

For a given position vector $x$ in the original coordinate system, we define the new position vector $x'$ to be given by

$$x' = c(t) + R(t)x,$$

where $c(t)$ is a differentiable vector function and $R(t)$ is a differentiable tensor function with the property that $R$ is always proper orthogonal.

---

1Note that this equation is labelled as (5.57) in the main text.
We recall that $L$ is defined by

$$ L = \frac{\partial}{\partial x} \left( \frac{Dx}{Dt} \right). $$

By analogy, let $L'$ be the velocity gradient in the new coordinate system, so that

$$ L' = \frac{\partial}{\partial x'} \left( \frac{Dx'}{Dt} \right). $$

Given this definition of $L'$, it can be shown that

$$ L' = R (L + S) R^T, $$

where $S = R^T \frac{DR}{Dt}$ is skew-symmetric.\(^2\)

Similarly, we find that

$$ E'(x', t) = RE(x, t) R^T, $$

and

$$ G'(x', t) = RG(x, t) R^T, $$

by using the standard rules for the rotation of a one-point tensor (see, for example, Malvern [118]).

\(^2\)To see this, we note that

$$ \frac{Dx'}{Dt} = \frac{dc}{dt} + \frac{DR}{Dt} x + R \frac{Dx}{Dt}. $$

Thus,

$$ L' = \frac{\partial}{\partial x} \left( \frac{Dx'}{Dt} \right) \frac{Dx}{Dt} $$

$$ = \frac{DR}{Dt} \frac{\partial x}{\partial x} R^T + R \frac{Dx}{Dt} \left( \frac{Dx}{Dt} \right) R^T $$

$$ = RR^T \frac{DR}{Dt} R^T + RL R^T, $$

yielding the required result.
Applying these results, we find that
\[
\text{sym}\left((I - 2E') L'\right) = \text{sym}\left((I - 2R E R^T)(L + S) R^T\right)
\]
\[
= \text{sym}\left(R(I - 2E)(L + S) R^T\right)
\]
\[
= R \text{ sym}\left((I - 2E)(L + S)\right) R^T
\]
\[
= R \text{ sym}\left((I - 2E)L - 2ES\right) R^T.
\]

Similarly,
\[
\text{sym}\left((I - 2E') G'\right) = \text{sym}\left((I - 2R E R^T) R G R^T\right)
\]
\[
= R \text{ sym}\left((I - 2E) G\right) R^T.
\]

Furthermore, we note that
\[
\det(I - 2E') = \det(I - 2R E R^T)
\]
\[
= \det(R(I - 2E) R^T)
\]
\[
= \det(I - 2E),
\]

since \(\det(AB) = \det(A) \det(B)\).

Combining these equations, we find that
\[
\text{sym}\left((I - 2E') L' - \frac{1}{\sqrt{\det(I - 2E')}} (I - 2E') G'\right)
\]
\[
= R \text{ sym}\left((I - 2E) L - \frac{1}{\sqrt{\det(I - 2E')}} (I - 2E) G - 2SE\right) R^T. \quad (D.3)
\]
Now,
\[
\frac{DE'}{Dt} = \frac{DR}{Dt} ER^T + R \frac{DE}{Dt} R^T + RE \frac{DR^T}{Dt}
\]
\[
= R \left( \frac{DE}{Dt} + R^T \frac{DR}{Dt} E + E \frac{DR^T}{Dt} R \right) R^T
\]
\[
= R \left( \frac{DE}{Dt} + S^T E + ES^T \right) R^T,
\]
where \( S = R^T \frac{DR}{Dt} \) as before. Since \( S \) is skew symmetric, this yields
\[
\frac{DE'}{Dt} = R \left( \frac{DE}{Dt} - 2 \ sym (S E) \right) R^T.
\]

Since \( E \) must satisfy equation (5.57), we find that
\[
\frac{DE'}{Dt} = R \ sym \left( (I - 2 E) L - \frac{1}{\sqrt{\det(I - 2 E)}} (I - 2 E) G - 2 S E \right) R^T.
\]

Combining this with equation (D.3), we find that
\[
\frac{DE'}{Dt} = \ sym \left( (I - 2 E') L' - \frac{1}{\sqrt{\det(I - 2 E')}} (I - 2 E') G' \right),
\]
as required.

From this, we conclude that equation (5.57) is observer-independent. Furthermore, we recall that the derivation of equation (5.57) used a general zero stress deformation gradient rather than the principal zero stress deformation gradient. Thus, equation (5.57) is valid for arbitrary rotations in the spatial coordinate system and arbitrary rotations in the local zero stress coordinate system.

Similarly, it is possible to show that (5.64), the infinitesimal strain evolution equation, is observer independent (not given here). However, we note that the
definition of infinitesimal strain depends directly on the principal zero stress deformation gradient. Hence, equation (5.64) is not valid for arbitrary rotations in the local zero stress coordinate system.
Bibliography


