Spiral Pinballs, Cardiac Tissue and Deforming Capacitors

by

Jacob Langham

A thesis submitted in partial fulfilment of the requirements for
the degree of

Doctor of Philosophy in Mathematics

University of Warwick, Mathematics Institute

April 2017
# Contents

List of figures
Acknowledgements
Declarations
Abstract

<table>
<thead>
<tr>
<th>1 Introduction</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Excitable media</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Spiral pinballs</td>
<td>5</td>
</tr>
<tr>
<td>1.3 Spiral waves in deforming media</td>
<td>7</td>
</tr>
<tr>
<td>1.4 Shape selection of dielectric elastomers</td>
<td>8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2 Spiral pinball reflections</th>
<th>11</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Introduction</td>
<td>11</td>
</tr>
<tr>
<td>2.2 Model and methods</td>
<td>13</td>
</tr>
<tr>
<td>2.3 Results</td>
<td>17</td>
</tr>
<tr>
<td>2.3.1 Small-core case</td>
<td>17</td>
</tr>
<tr>
<td>2.3.2 Large-core case</td>
<td>23</td>
</tr>
<tr>
<td>2.4 Discussion</td>
<td>28</td>
</tr>
<tr>
<td>2.4.1 Biktashev-Holden theory</td>
<td>28</td>
</tr>
<tr>
<td>2.4.2 Multiple reflections</td>
<td>29</td>
</tr>
<tr>
<td>2.4.3 Concluding remarks</td>
<td>32</td>
</tr>
<tr>
<td>2.A Appendix: Resonant forcing parameters</td>
<td>33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3 Asymptotic dynamics of spiral pinballs</th>
<th>35</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Introduction</td>
<td>35</td>
</tr>
<tr>
<td>3.2 Theory</td>
<td>38</td>
</tr>
<tr>
<td>3.3 Model and methods</td>
<td>42</td>
</tr>
<tr>
<td>3.4 Results</td>
<td>44</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Characteristic dynamics of an excitable system after a super-threshold perturbation. ........................................... 3
1.2 Example of a rigidly rotating spiral wave. ......................... 4
1.3 Example trajectory of a spiral pinball in a square box. .......... 6
1.4 Example of a spiral wave in a deforming excitable medium. 7
1.5 (a) Diagram showing the basic mode of operation for a dielectric elastomer. (b) Example of pattern formation in a dielectric elastomer with annular electrodes. 8

2.1 Illustration of resonant drift and reflection for spiral waves in excitable media. ..................................................... 12
2.2 Example results for a spiral pinball reflection from a no-flux boundary. ................................................................. 18
2.3 Example results for a spiral pinball reflection from a step inhomogeneity. ............................................................... 19
2.4 Illustration of the insensitivity of reflections throughout the small-core region of parameter space. ....................... 20
2.5 Effect of forcing amplitude on reflection of small-core spiral pinballs for the case of a Neumann boundary. ............... 22
2.6 Effect of forcing amplitude on reflection of small-core spiral pinballs for the case of a step boundary. ......................... 22
2.7 (a) Reflected angle as a function of forcing amplitude for small-core spiral pinballs approaching normal to a no-flux boundary. (b) Effect of phase on the annihilation of a spiral pinball at the boundary. (c) Annihilation at a very large forcing amplitude. 23
2.8 (a) Example of a large-core spiral wave. (b) Comparison between a small- and large-core spiral pinball reflection. ........ 24
2.9 Summary of results for large-core spiral pinballs. ............. 25
2.10 Catalogue of interesting trajectories for large-core spiral pinballs. ................................................................. 26
2.11 Sketch showing the geometry of multiple reflections in a portion of a square box. .................................................. 29
2.12 Examples of non-square paths for large core spiral pinballs. ................................................................. 31

3.1 Two examples of spiral pinballs reflecting from a step boundary: (a) small-core and (b) large-core. .......................... 37
3.2 $S_X$, $S_Y$ and $S_\Phi$ for a representative small-core spiral wave. ............................................................... 45
3.3 Theoretical trajectory of a small-core spiral pinball reflection with phase-dependent vector field plotted. ................................. 47
3.4 Two theoretical trajectories in the small-core regime at different incident angles. ................................................. 48
3.5 $S_X$, $S_Y$ and $S_\Phi$ curves, together with representative theoretical reflection trajectories for three different cases spanning a substantial extent of the small-core regime. .......................................................... 49
3.6 $S_X$, $S_Y$ and $S_\Phi$ for a representative large-core spiral wave. ................................................................. 51
3.7 Theoretical trajectory of a large-core spiral pinball reflection with phase-dependent vector field plotted. ......................... 52
3.8 Demonstration of the effect of incident angle for a large-core spiral pinball. ...................................................... 53
3.9 Phase dynamics for large-core spiral pinballs approaching the boundary with different incident angles. ....................... 53
3.10 Demonstration of the effect of forcing amplitude on large-core spiral pinballs. .................................................. 55
3.11 Reflected angle versus incident angle for large-core spiral pinballs at different forcing amplitudes. ......................... 55
3.12 Comparison between theory and simulation for a variety of parameter values and incident angles in the small-core regime. 57
3.13 Comparison between theory and simulation verifying the theoretical predictions for the role of incident angle and forcing amplitude in the large-core regime. .................................................. 58

4.1 Diagram showing the relationship between the reference and deformed configurations and the role of the motion and deformation maps. .......................................................... 71
4.2 Visualisation of the components of the Cauchy stress tensor. ................................................................. 77
4.3 Snapshot of a spiral wave in a deforming medium with Dirichlet boundary conditions. ................................ 96
4.4 Frames of a spiral wave moving in a deforming medium with traction boundary conditions. ........................................ 97
4.5 Frames of a spiral wave breaking up in a deforming medium with traction boundary conditions. ........................ 99
4.6 Snapshot of a spiral wave propagating in a three-dimensional slab with traction boundary conditions.

5.1 Diagrams explaining the operation of a dielectric elastomer.
5.2 Example result showing the effect of tangent forces.
5.3 Photograph showing the dielectric elastomer experiment from above.
5.4 The dielectric elastomer experiment with circular active region, before and after actuation.
5.5 Schematic of the dielectric elastomer model reference configuration.
5.6 (a) Example of a typical deformed elastomer configuration for a thin disc with circular active region. (b) Boundary conditions for the circular disc.
5.7 Comparison of experimentally measured deflections with simulation profiles for a succession of increasing voltages.
5.8 Effect of tangent force on the shape of model profiles.
5.9 Overhead view of an elastomer experiment with an annular active region.
5.10 Diagram of the model setup for a dielectric elastomer with annular active region, along with corresponding example solutions.
5.11 Deformed configurations for a circular disc elastomer with annular active regions of different widths.
5.12 Width of annular active region versus observed ripple wavelength for the experiment and model simulations.
5.13 Deformed configurations for a circular disc with annular active region, using different amounts of tangential force.
5.14 Diagram of the model setup for a dielectric elastomer with a rectangular active region, along with a corresponding model solution.
5.15 Deflated solutions for an infinite rectangular strip with applied tangential force.
5.16 Deflated solutions for an infinite rectangular strip with no applied tangential force.
5.17 Plots of the solution of electric potential for a rigid capacitor in two dimensions.
Acknowledgements

First and foremost, I thank my superlative PhD supervisor, Dwight Barkley: for always treating me as a valued colleague and not as a student (even when I had ideas that were clearly wrong!) and for your invaluable support and advice over the last five years. Without your help, I would not have got this far.

I would also like to thank my earlier supervisors, Kurt Debattista and Alan Chalmers at WMG, for introducing me to a research field that I was completely new to and for encouraging me always to work on whatever I found most appealing.

Next, I thank my scientific collaborators. Irina Biktasheva, for your kind help and expert advice on the material in Chapter 3 and for your work on the DXspiral software which was essential for that project. Andreas Dedner, for introducing me to DUNE and subsequently for helping me with my questions and conferring your deep knowledge of how it all works. Hadrien Bense, Benoît Roman and José Bico, for showing us the dielectric elastomer problem, for stimulating discussions on the physics and for your excellent experiments which underpin all the results in Chapter 5.

I thank the many people at Warwick who enriched my time here. Sam, for long walks, trips to the pub and for introducing me to many new ideas and experiences; Ed, for squash, running advice and your often fatalistic views on work and life; and everyone I know at the mathematics department, in particular Andrea, Costanza, Elena, José, Mim and Sara—thank you all for making this a fun and friendly place to work.

Finally, I thank both my partner Ellie and my family, for your encouragement, kindness and everything else you have done for me.
Declarations

The work in presented in this thesis was carried out by the author under the supervision of, and in collaboration with, Professor Dwight Barkley, save for the exceptions noted below.

Chapter 3 contains research conducted partly in collaboration with Dr. Irina Biktasheva at the University of Liverpool. The main contribution of the chapter is the results in Section 3.4, which are the work of the author.

Section 4.2 concerns expository theoretical material that is largely standard. I draw mainly upon the following references: [AF80, Gur81, Ogd84, Ant95, GS08]. While the presentation is my own, I do not claim that it is original.

In Chapter 5, all experimental data and photographs are due to our collaborators Hadrien Bense, Benoît Roman and José Bico at the laboratoire de Physique et Mécanique des Milieux Hétérogènes (PMMH), EPSCI, Paris.

Parts of this thesis have been published by the author:

- Chapter 2 was published in the journal Chaos in 2013 [LB13],

- Chapter 3 was published in the journal Physical Review E in 2014 [LBB14].

Aside from minor revisions intended to integrate the texts into the thesis as a whole, both papers are reproduced verbatim.

This thesis has not been submitted for a degree at any other university.
Abstract

‘Spiral pinballs’ are resonantly drifting spiral waves in excitable media that reflect from boundaries. Instead of reflecting at an angle equal to the one at which they approach the boundary—like a ray of light reflecting from a mirror—they reflect in a preferred direction. This invites comparison with a number of other complex systems that behave as nonspecular billiards, including bouncing droplets on a vibrated bath, swimming microorganisms and segments of chemical waves. In the first part of this thesis, we study the trajectories of spiral pinball reflections. A catalogue of interesting behaviours is discovered in both the small- and large-core rotation regimes and the long-term billiard dynamics is briefly considered. By using an asymptotic theory, we examine the reflection process in detail and thereby explain many of the observed phenomena.

The second part of this thesis concerns itself with modelling spiral wave activity in a deforming medium. Our motivation stems from cardiac tissue, in which spiral waves and mechanical deformation are reciprocally coupled. We describe a simple modelling approach for this system and discuss its implementation. Various different results are presented using this model.

Finally we consider a problem from the engineering world. Dielectric elastomers are flexible capacitors that undergo nonlinear elastic deformations in response to forces arising from electric surface charges. We propose a novel modelling approach that decomposes these forces into a compressive stress and a tangential shear. The tangential component corresponds to a fringing effect that is usually considered to be negligible. Via numerical simulations and comparison with experimental data we show that it nonetheless has an important role to play in selecting the deformed shapes that these systems adopt. In some cases, we are able to compute multiple equilibrium configurations and it is shown that doing so is necessary to obtain the most physically relevant states.
Chapter 1

Introduction

This thesis is about three different problems in nonlinear science. In each case, a simple mathematical model is used to capture the most important features of the problem. The systems considered are not easily studied analytically and we shall make extensive use of computational simulations to understand them.

1.1 Excitable media

The first two problems involve spiral waves in excitable media. A dynamical system is said to be excitable if a small perturbation from equilibrium, above a certain threshold, results in a comparatively large deviation before recovering to the initial state. This process is called excitation. Additional perturbations during the period of recovery are unable to excite the system. The prototypical example of excitation occurs in the dynamics of a neuron cell. In the resting state of a typical neuron there is a potential difference across its membrane of around $-60 \text{ mV}$ to $-70 \text{ mV}$. This equilibrium voltage is dictated by the concentration of various ions inside and outside the cell and is stable to small fluctuations. However, perturbations of sufficient magnitude result in rapid depolarisation of the neuron, followed by a slow recovery to rest. Such events are called action potentials. They provide a basis for the transmission of electrochemical signals in the brain and a variety of other contexts [GK02].
An excitable medium is a spatially extended dynamical system in which every point is excitable. These are commonly described by particular systems of reaction-diffusion partial differential equations (PDEs), which, in two spatial dimensions, typically admit rotating spiral-shaped wave solutions. In Chapter 3, we consider these systems in a very general form [Eq. (3.1)]. For the time being, we give the following two-variable example:

\[
\begin{align*}
\partial_t u &= \frac{1}{\epsilon} f(u,v) + D \nabla^2 u, \\
\partial_t v &= g(u,v).
\end{align*}
\]  

(1.1a)

(1.1b)

The state variables \( u \) and \( v \) are time-dependent spatial fields that are responsible for excitation and recovery respectively.

Neglecting the diffusion term on the end of Eq. (1.1a) for the moment, the functions \( f \) and \( g \) give the form of the local dynamics at each point. They are often called (reaction) kinetics terms because differential equations of this type may be used to model the reaction rates of chemical media. The parameter \( \epsilon \), which we shall take to be small and strictly positive, is included to indicate that the dynamics of \( f \) operate on a much faster time scale than the dynamics of \( g \). Figure 1.1 shows a solution of the ordinary differential equation (ODE) system \( \dot{u} = f/\epsilon, \dot{v} = g \), using the kinetics terms from the excitable model that we use in the later chapters [Eqs. (2.1a) and (2.1b)]. The details are not as important, at this stage, as the observed behaviour of the system. Both state variables are quiescent (at rest) when they are zero. This is the initial condition for the system at \( t = -1 \). At time \( t = 0 \) we introduce a perturbation to the \( u \) variable, shifting the state from 0 to 0.1, which exceeds the threshold for excitation in this case. Immediately, the fast \( u \)-dynamics rapidly excites the system. The excited state \( u = 1 \) is a zero of \( f \) and as \( u \) approaches this value, \( u \) decays to zero. Meanwhile, \( v \) increases on a slower time scale. The speed of the \( v \)-dynamics relative to the \( u \)-dynamics sets the length of time that the system may be excited for. As \( v \) peaks, it begins to inhibit \( u \), changing the sign of \( f \) and causing \( u \) to fall rapidly back to zero. Finally, \( v \) ‘recovers’ slowly back to rest. During this period, the excitation threshold is elevated.
Figure 1.1: Characteristic dynamics of an excitable system after a super-threshold perturbation occurring at $t = 0$.

and the system may not be excited by another small perturbation.

This picture of excitation and recovery captures the essence of a typical excitable system. Although the details of individual models and the exact interplay of their respective state variables may differ, the generic behaviour sketched above applies across a wide range of excitable systems.

Returning to the full PDE system in Eqs. (1.1a) and (1.1b), the Laplacian operator acting on $u$ with rate coefficient $D > 0$ causes excited regions to diffuse into nearby quiescent patches, thereby pushing them above the threshold for excitation. In most excitable systems of interest, this mechanism is sufficient to allow the propagation of undamped travelling waves, including the spiral waves that shall occupy our attention for two thirds of this thesis.

A picture of a spiral wave is shown in Figure 1.2. Regions of quiescence are plotted in dark blue and excited regions are plotted in orange. This particular example is a rigidly rotating spiral—it is a periodic solution to Eqs. (1.1a) and (1.1b) that rotates about a centre point. The tip of the wave\footnote{The exact definition of the spiral wave tip is a matter of convention and is not important to our discussion here. See Section 2.2 for the definition used in later results.} traces a small circle as the spiral rotates. The region inside this circle remains permanently unexcited. It is referred to as the spiral wave core.
Figure 1.2: Example of a spiral wave. Excited regions are shown in orange and quiescent regions are dark blue. This wave rotates clockwise about a point close to the centre of the square. The tip of the wave traces a small circle around this point, drawn in white.

Individual points far from the core resemble the plot in Fig. 1.1. Instead of an abrupt perturbation, smooth diffusion of $u$ from an approaching wave front brings the point above the excitation threshold. This is followed by rapid excitation, plateau and eventual return to quiescence.

Spiral waves were initially discovered in physical systems in an oscillating chemical reaction first studied by Belousov in the 1950s [Bel59] and several years later by Zhabotinsky [Zha64]. It was subsequently shown that a thin layer of the reagents spread onto a Petri dish can support the formation of various patterns, including spiral waves [ZZ70]. This marked the beginning of widespread interest in the reaction [Zha91]. The system became known as the Belousov-Zhabotinsky (BZ) reaction; to this day it represents one of the most important examples of an excitable medium, due to the ease with which experiments may be conducted and reproduced. Since the discovery of the BZ reaction, spiral waves were observed in many different contexts. Cases span a variety of scientific disciplines and cover a wide range of physical scales. We shall list but a few examples: electrical activity in chicken retinas [GB83], equations modelling the evolution of interstellar media [NI84, NI89], chemo-
tactic signalling in populations of slime mould [SW89] and rusting of steel plates [AS00]. Of particular importance is the case of cardiac tissue, in which spiral wave propagation was first experimentally demonstrated by Davidenko et al. [DPS+92]. This system will be discussed later and in Chapter 4.

1.2 Spiral pinballs

Not all spiral waves rotate rigidly, as in the example plotted in Fig. 1.2. Their cores may move as the spiral rotates. The rotation centres of so-called meandering spirals move spontaneously due to internal instabilities in the dynamics of rigidly rotating spirals [Bar95]. In doing so, their tips trace out a variety of interesting patterns [JSW89]. Alternatively, movement of the core may be induced by symmetry-breaking perturbations imposed on the medium [Bik07]. This phenomenon is known as drift. Such perturbations include, but are not limited to: medium inhomogeneities [DPS+92, MNUH92, MPMPV98], physical boundaries close to the spiral core [EP86, SB93, BH95], nearby spiral waves [EPS89] and external stimulation of the medium [ADM87, ADK92, SSM92].

The system that we refer to as a spiral pinball combines two of these kinds of perturbation. It is well known that external periodic forcing of the excitable medium close to the spiral rotation frequency can produce drift of the wave rotation centre along a straight line [ADM87, NvORE93, SBE95]. On approaching a boundary, such a resonantly drifting spiral changes its direction and in doing so ‘reflects’ [BH93, BH95]. This leads to a billiard-like or ‘pinball’ dynamics, demonstrated in Figure 1.3, which shows the trajectory of a spiral wave tip undergoing three subsequent reflections. While the plotted motion of the tip is cycloidal, the corresponding trajectory of the rotation centre is a straight line, apart from at the boundaries where it abruptly changes direction.

The reflections of spiral pinballs are typically nonspecular (incidence angle rarely equals reflection angle), inviting comparison with systems in various contexts that behave as nonspecular billiards. Studies of chemical
wave segments [STS08], swimming microorganisms [SL12, WLST15], optical solitons [PLTB11] and bouncing droplets on a vibrated bath [EFMC09, Shi13, PSFB16] have all demonstrated similar reflection properties to those of the spiral pinball system. Furthermore, with the exception of the microorganisms, each of these systems exhibits both wave and particle characteristics. This is particularly striking in the bouncing droplet experiment, which exhibits a number of fascinating phenomena associated with quantum mechanics on a macroscopic scale [CPFB05, CF06, EFMC09, FEB10, Bus10, PLM14]. A form of wave-particle duality is an established property of spiral waves [BB03] and the spiral pinball system in particular highlights this fact. Here, the ‘particle’ is the spiral core that ricochets around the medium, accompanied by the rotating wave train.

Motivated by the above connections, we study the reflections of spiral pinballs in detail, surveying the short- and long-term dynamics via extensive numerical simulations in Chapter 2 and using an asymptotic theory to investigate the underlying reflection mechanism in Chapter 3.
1.3 Spiral waves in deforming media

In Chapter 4 we consider spiral waves in a nonlinear elastic domain. An example is shown in Fig. 1.4 and should be contrasted with the simpler situation in Fig. 1.2. The motivation comes from cardiac tissue, where excitation waves cause local contraction of the medium. A spiral wave propagating in a deforming medium experiences a perturbation that affects its dynamics. In this way, mechanics and wave are coupled in both directions: the mechanics affects the wave, which in turn dictates the mechanics. Although the mutual interaction of spiral wave and mechanical deformation has been addressed before [NP04, CFNT08, KNP09, WNP11, CPZ+14], the problem is still not completely understood and open questions remain. We formulate an approach, based on prior studies, for numerically simulating spiral waves in isotropic deforming media from the ground up, starting from the theory of nonlinear elasticity, which is covered in some detail. Modelling issues and implementation of the elastic problem via the finite element method are also

Figure 1.4: Example of a spiral wave in a deforming excitable medium. Local contraction of the geometry is activated by the regions that are excited. In this case, excited regions are shown in red, quiescent regions are shown in pale blue and the spiral is rotating anticlockwise. Black dashes indicate the deformation of material lines.
discussed. A variety of example results from the implementation are given.

1.4 Shape selection of dielectric elastomers

For the final problem, we leave spiral waves and study pattern formation in the static equilibrium shapes adopted by a different nonlinear elastic medium. Motivated by results obtained by experimental colleagues, we model a type of capacitor made with flexible plates and an elastic material that deforms under the electrostatic forces from its stored charges. These systems, called dielectric elastomers, have received significant attention from the engineering community in recent years due to the large number of potential applications highlighted by researchers in technology, robotics and other areas [OOM08, CBDR10].

A diagram of a typical dielectric elastomer in operation is shown in Fig. 1.5(a). The elastic medium is a thin elastomer film with a circular electrodes on the top and underside. When a potential difference is applied across these electrodes, the material between them is compressed. This is coupled to a lateral stretching via incompressibility.

Under certain boundary conditions, the compressive pressure exerted by the capacitor plates leads to a buckling instability that gives rise to interesting out-of-plane deformations. An example exhibiting pattern formation on
the deformed surface of an elastomer with annular electrodes is shown in Fig. 1.5(b). In Chapter 5, we propose a novel modelling approach for these systems that produces quantitative agreement with experimental data and captures such patterns formed by the deformation. Furthermore, we use our model to obtain insight into aspects of the underlying physics that are usually neglected in modelling studies. The computation of multiple nonequivalent equilibrium shapes is also discussed.
Chapter 2

Spiral pinball reflections

Wave-particle duality is typically associated with quantum mechanical systems. However, in recent years it has been observed that some macroscopic systems commonly studied in the context of pattern formation also exhibit wave-particle duality. Two systems in particular have attracted considerable attention in this regard: drops bouncing on the surface of a vibrated liquid layer [CFGB05, CPFB05, PBC06, CF06, EFMC09, EDFC09] and waves in chemical media [BB03, Bik07, SMCS02, STS08, BBBF10, BBB10]. The second case is the focus of this chapter. We explore non-specular reflections associated with spiral waves in excitable media—reflections not of the waves themselves, but of the particle-like trajectories tied to these waves.

2.1 Introduction

Rotating spirals are a pervasive feature of two-dimensional excitable media, such as the Belousov-Zhabotinsky reaction [Bel59, Zha64, ZZ70, Win84]. Figure 2.1(a) illustrates a spiral wave from a standard model of excitable media discussed below. The wave character of the system is evident. As the spiral rotates, a periodic train of excitation is generated which propagates outward from the centre, or core, of the spiral. Much of the historical study of excitable media has focused on the wave character of the problem, as illustrated by efforts to determine the selection of the spiral shape and rotation.
Figure 2.1: Illustration of resonant drift and reflection for spiral waves in excitable media. (a) Periodically rotating spiral wave in the unforced regime. The wave rotates around a fixed core and the path of the spiral tip (white) is a circle. (b) Resonant drift. The medium is parametrically forced at the spiral rotation frequency. The core moves along a straight path and the spiral tip traces out a cycloid (white). (c) Reflection of drifting spiral from a no-flux boundary. The incident and reflected angles, $\theta_i$ and $\theta_r$, are indicated. (d) Path of a drifting spiral in a square box. The underlying spiral wave at one instant in time is shown faintly. In all cases the plotted fields are the excitation variable, $u$, of the reaction diffusion model. Details are given later in the text.

However, it is now understood that these spiral waves also have particle-like properties. This was first brought to the forefront by Biktasheva and Biktashev [BB03] and has been developed in more recent years [Bik07, BBBF10, BBB10, BBB 09, BBS11]. One of the more striking illustrations of a particle-like property is resonant drift [ADM87, DZMB88, BH93, BH95, SZM93, SBE95, MB96, ZWY 04, Bik07, BBBF10], shown in Fig. 2.1(b)-(d). Resonant drift can occur spontaneously through instability, or due to spatial inhomogeneity (meander), or as here, by means of resonant parametric forcing (periodically
varying the medium parameters in resonance with the spiral rotation frequency). As is seen, the core of the spiral drifts along a straight line. The speed is dictated by the forcing amplitude while the direction is set by the phase of the forcing, or equivalently the initial spiral orientation.

The trajectories of resonantly drifting spirals are unaffected by the domain boundaries (or other spirals should they be present) except on close approach, where often the result is a reflection of the drifting core [BH93, BH95, OS08], as illustrated in Fig. 2.1(c). Reflections are not specular—the reflected angle $\theta_r$ is not in general equal to the incident angle $\theta_i$. When placed in a square box, the drift trajectory typically will ricochet off each boundary in such a way to eventually be attracted to a unique square path where $\theta_i + \theta_r = 90^\circ$, as shown in figure 2.1(d). (This is the more common case, but others are considered herein.)

We shall refer to this situation of a reflecting, resonantly drifting spiral wave as a spiral pinball, thereby emphasising the particle-like behaviour of the system.

The primary goal of this chapter is firstly to determine accurately, through numerical simulations, the relationship between reflected and incident angles for some representative cases of spiral pinballs in excitable media, and secondly to explore the qualitative features of reflections in excitable media, particularly multiple reflections in square domains. While the numerical and theoretical study of reflecting trajectories was undertaken by Biktashev and Holden many years ago [BH93, BH95], much more extensive results are now possible and desirable, especially since phenomena strikingly similar to that seen in Figs. 2.1(c) and (d) have been observed in other macroscopic systems with both wave-like and particle-like properties [PBC06, EDFC09, SMCS02, STS08, PLTB11, Sho12].

### 2.2 Model and methods

Our study is based on the standard Barkley model describing a generic excitable medium [Bar91]. In the simplest form the model is given by the
reaction-diffusion equations

\[
\frac{\partial u}{\partial t} = \nabla^2 u + \frac{1}{\epsilon}u(1 - u)\left(u - \frac{v + b}{a}\right), \quad (2.1a)
\]

\[
\frac{\partial v}{\partial t} = u - v, \quad (2.1b)
\]

where \(u(x,y)\) is the excitation field (plotted in Fig. 2.1) and \(v(x,y)\) is the recovery field; \(a, b,\) and \(\epsilon\) are parameters. The parameters \(a\) and \(b\) collectively control the threshold for and duration of excitation while the parameter \(\epsilon\) controls the excitability of the medium by setting the fast timescale of excitation relative to the timescale of recovery.

We consider two parameter regimes—known commonly as the small-core and large-core regimes. The small-core case is shown in Fig. 2.1. As the name implies, the core region of the spiral, where the medium remains unexcited over one rotation period, is small. This is the more generic case for the Barkley model and similar models and occupies a relatively large region of parameter space in which waves rotate periodically. Small-core spirals are found in the lower right part of the standard two-parameter phase diagram for the Barkley model (see Figure 4 of Ref. [Bar08]). Large-core spirals rotate around relatively large regions [see Fig. 2.8(a) discussed below]. Such spirals occur in a narrow region of parameter space [Bar08, Bar94] near the boundary for wave propagation failure. The core size diverges to infinity near this boundary.

Parametric forcing is introduced through periodic variation in the excitability. Specifically, we vary \(\epsilon\) according to

\[
\epsilon(t) = \epsilon_0 \left(1 + A \sin(\omega_f t + \phi)\right), \quad (2.2)
\]

where \(A\) and \(\omega_f\) are the forcing amplitude and frequency. The phase \(\phi\) is used to control the direction of resonant drift. The forcing frequency producing resonant drift will be close to the natural, unforced, spiral frequency. However, one effect of perturbing \(\epsilon\) is that it induces a shift in the spiral rotation frequency that depends nonlinearly on the amplitude. Therefore \(\omega_f\) must be
adjusted with $A$ to produce resonant drift along a straight line.

We have studied reflections in two situations. The first is reflection from a no-flux boundary. This type of boundary condition corresponds to the wall of a container containing the medium. We set the reflection boundary to be at $x = 0$ and impose a homogeneous Neumann boundary condition there:

$$\frac{\partial u}{\partial x}(0, y) = 0. \quad (2.3)$$

Since there is no diffusion of the slow variable, no boundary condition is required on $v$. The medium does not exist for $x < 0$.

The second situation we have studied is reflections from a step change in excitability across a line within the medium. We locate step change on the line $x = 0$. We vary the threshold for excitation across this line by having the parameter $b$ vary according to

$$b(x, y) = \begin{cases} b_0 & \text{if } x \geq 0, \\ b_0 - \triangle b & \text{if } x < 0. \end{cases} \quad (2.4)$$

Unlike for the no-flux boundary, in this case waves may cross the line $x = 0$ and so there is no boundary to wave propagation. Nevertheless, drifting spirals may reflect from this step change in the medium and we refer to this a step boundary.

The numerical methods for solving the reaction-diffusion equations are standard and are covered elsewhere [Bar91, DMB97]. Some relevant computational details particular to this study of spiral reflections are as follows. A converged spiral for the unforced system is used as the initial condition. Simulations are started with parametric forcing and the spiral drifts in a particular direction dictated by the phase $\phi$ in Eq. (2.2). The position of the spiral tip is recorded once per forcing period (which, at resonance, exactly equals the spiral rotation period). We adopt the convention of defining the tip to be the intersection point of two $u$ and $v$ isolines, for which we choose $u(x, y) = 0.5$ and $v(x, y) = 0.5a - b$. From several successive tip samples, the direction of drift, i.e. the incident angle $\theta_i$, is determined by a least-squares fit
over an appropriate range of drift (after the initial spiral has equilibrated to a state of constant drift, both in speed and direction, but before the spiral core encounters a boundary). Likewise, from a fit to the sampled tip path after the interaction with the boundary, we determine the reflected angle $\theta_r$. By varying $\phi$ we are able to scan over incident angles.

The simulations are carried out in a large rectangular domain with no-flux boundary conditions on all sides. For reflections from a Neumann boundary [Eq. (2.3)], we simply direct waves to the computational domain boundary corresponding to $x = 0$. We also study reflections more globally from all sides of a square domain with Neumann boundary conditions, such as in Fig. 2.1(d). In the study of reflections from the step boundary [Eq. (2.4)], the computational domain extends past the step change in parameter. We have run cases with the left computational boundary both at $x = -7.5$ and $x = -15$ and these are sufficiently far from $x = 0$ that trajectory reflection is not affected by the computational domain boundary. The dimensions of the rectangular computational domain are varied depending on the angle of incidence. For $\theta_i \approx \pm 90^\circ$ we require a long domain in the $y$-direction, whereas for $\theta_i \approx 0^\circ$ a much smaller domain may be used. In all cases we use a grid spacing of $h = 1/4$. The time step is varied to evenly divide the forcing period, but $\Delta t \approx 0.019$ is typical. Except where stated otherwise, the model parameters for the small-core case are: $a = 0.8, b = 0.05,$ and $\epsilon_0 = 0.02$. For the large-core case they are $a = 0.6, b = 0.07,$ and $\epsilon_0 = 0.02$. For the step boundary $b_0 = 0.05$ and $\Delta b = 0.025$. Different values of the forcing amplitude and period, $A$ and $\omega_f$, are considered. Given the desire to measure incident and reflected angles precisely, we have required drift be along straight lines to high precision and in turn this has required high accuracy in the imposed forcing amplitude and period. In the Appendix to this chapter we report the exact values for the forcing parameters used in the quantitative incidence-reflection studies.
2.3 Results

Before presenting results from our study of reflections, it is important to be precise about the meaning of incident and reflected angles. As is standard, angles are measured with respect to the boundary normal. This is illustrated in Fig. 2.1(c). What needs to be stressed here is that spirals have a chirality—right or left handedness—and this implies that we need to work with angles potentially in the range \([-90^\circ, 90^\circ]\), rather than simply \([0^\circ, 90^\circ]\).

Specifically, we consider clockwise rotating spiral waves and define \(\theta_i\) to be positive in the clockwise direction from the normal. We define \(\theta_r\) to be positive in the anticlockwise direction from the normal. Both \(\theta_i\) and \(\theta_r\) are positive in Fig. 2.1(c) and for specular reflections \(\theta_r = \theta_i\).

2.3.1 Small-core case

We begin with the small-core case already shown in Fig. 2.1. Figures 2.2 and 2.3 illustrate the typical behaviour we find in reflections from both types of boundaries. In both figures the upper plot shows measured reflected angle \(\theta_r\) as a function of incident angle \(\theta_i\) over the full range of incident angles. The lower plots show representative trajectories for specific incident angles indicated. Here and throughout, the no-flux nature of the Neumann boundary is indicated with shading (\(x = 0\) is the at the rightmost edge of the shading), while the step in excitability at a step boundary is indicated with sharp lines. All parameters are the same for the two cases; they differ only in the type of boundary that trajectories reflect from.

The reflections are far from specular. This is particularly striking for \(\theta_i < 0\) where the incoming and outgoing trajectories lie on the same side of the normal. The reflected angle is nearly constant, independent of the incident angle, except for incident angles close to \(\theta_i = 90^\circ\). There is a slight variation in the reflected angle, seen as undulation in the upper plots, but the amplitude of the variation is small.

One can also observe in the lower plots that the point of closest approach is also essentially independent of incident angle, except close to \(\theta_i = 90^\circ\)
Figure 2.2: Illustrative results for reflection from a no-flux boundary, i.e. Neumann boundary condition. (a) Reflected angle $\theta_r$ versus incident angle $\theta_i$. (b) Representative tip trajectories showing reflections at the incident angles marked with crosses in (a). The reflected angle is nearly constant for the full range of incident angles. The forcing amplitude is $A = 0.072$. 
Figure 2.3: Illustrative results for reflection from a step boundary, i.e. a step change in the excitability of the medium. (a) Reflected angle $\theta_r$ versus incident angle $\theta_i$. (b) Representative tip trajectories showing reflections at the incident angles marked with crosses in (a). The reflected angle is nearly constant for the full range of incident angles. The forcing amplitude is $A = 0.072$. 
where the distance grows. Spiral pinball trajectories come much closer to the step boundary than to the Neumann boundary.

It is worth emphasising that there is no effect of rotation phase in the results presented in Figs. 2.2 and 2.3. As the incident angle is scanned, the instantaneous rotation phase of the spiral wave as it reaches a given distance to the boundary will be different for different incident angles. In particular, the phase will differ at the point of closest approach, where the boundary effect is strongest. While this could have an effect on the reflected angle, we have verified that there is no such effect for the small-core cases we have studied, except at large forcing amplitudes near where spirals annihilate at the boundary (discussed later).

\[ \theta_i \approx 0^\circ, \] including faint visualization of the \( u \)-field at a particular time instance, while lower plots show \( \theta_i \approx -70^\circ \). The reflected angle is nearly constant independently of incident angle, parameter values, and boundary type. Model parameters span a substantial range of the non-meandering small-core region: in (a) and (b) \( a = 0.7, b = 0.01 \); in (c) and (d) \( a = 0.95, b = 0.01 \); in (e)–(h) \( a = 0.95, b = 0.08 \). Cases (g) and (h) are inhomogeneous boundary, the others are all Neumann boundaries. \( A = 0.072 \) throughout.

Figure 2.4: Illustration of the insensitivity of reflections throughout the small-core region of parameter space. Upper plots show \( \theta_i \approx 0^\circ \), including faint visualization of the \( u \)-field at a particular time instance, while lower plots show \( \theta_i \approx -70^\circ \). The reflected angle is nearly constant independently of incident angle, parameter values, and boundary type. Model parameters span a substantial range of the

While we have not conducted detailed studies at other parameter values, we have explored the small-core region of parameter space. Figure 2.4 shows representative results at distant points within the small-core region. The figure indicates not only a qualitative robustness, but also a quantitative insensitivity to model parameter values throughout the small-core region. In each case the upper plot shows $\theta_i \approx 0^\circ$ while the lower plot shows $\theta_i \approx -70^\circ$. The reflected angle varies by only a few degrees throughout all cases shown in the figure. Case (a)–(b) is close to the meander boundary while (c)–(d) is far from the meander boundary and corresponds to a very small core. Cases (e)–(h) are relatively large values of parameters $a$ and $b$, both with Neumann and step boundary conditions.

In the step boundary case, there is also the effect of $\Delta b$ to consider. Across a number of representative incident angles, we observed that as $\Delta b$ is incremented from 0.025 up to 0.05, the closest approaches of the spiral wave tips occur further from the boundary. We also find a slight reduction in the angle of reflection. Decrementing $\Delta b$ has the opposite effect. However, if $\Delta b$ is too small then the repulsive effect at the boundary will be too small and the spiral cores will cross the boundary.

We have examined the effect of forcing amplitude $A$. Figures 2.5 and 2.6 show reflected angle as a function of incident angle for various values $A$ as indicated. There is a decrease in the reflected angle with increasing forcing amplitude, or equivalently increasing drift speed. Generally there is also an increase in the oscillations seen in the dependence of reflected angle on incident angle. The solid curves are from the Biktashev-Holden theory discussed in Sec. 2.4.1.

For sufficiently large forcing amplitudes small-core spirals may be annihilated as they drift into Neumann boundaries. In such cases no reflection occurs. We have not investigated this in detail as it is outside the main focus of our study on reflections. Nevertheless, we have examined the effect of increasing the forcing amplitude through the point of annihilation for the case of a fixed incident angle $\theta_i = 0^\circ$. The results are summarised in Figure 2.7. The reflected angle reaches a minimum for $A \approx 0.11$, and thereafter increases slightly, but does not vary by more than 4° up to the amplitude where annihi-
Figure 2.5: Effect of forcing amplitude on reflection of small-core spirals pinballs for the case of a Neumann boundary. Points are measured reflected angle as function of incident angle at forcing amplitudes $A$ indicated. Solid curves are from Biktashev-Holden theory discussed in Sec. 2.4.1.

Figure 2.6: Effect of forcing amplitude on reflection of small-core spiral pinballs for the case of a step boundary. Points are measured reflected angle as function of incident angle at forcing amplitudes $A$ indicated. Solid curves are from Biktashev-Holden theory discussed in Sec. 2.4.1.
Figure 2.7: (a) Reflected angle as a function of forcing amplitude $A$ up to the point of annihilation at a Neumann boundary for small-core spiral pinballs. The incident angle is fixed at $\theta_i = 0^\circ$. (b) Tip trajectories a little below ($A = 0.215$) and a little above ($A = 0.235$) the forcing amplitude resulting in annihilation of the spiral at the boundary. (c) Tip trajectory at $A = 0.5$ showing annihilation at very large forcing amplitude.

Annihilation occurs, $A \approx 0.225$, as indicated in Figure 2.7(a). The forcing amplitude at which annihilation first occurs is rather large in that it corresponds to displacing the spiral wave considerably more than one unforced core diameter per forcing period. Figure 2.7(b) shows tip trajectories on either side of the amplitude where annihilation occurs, while Figure 2.7(c) shows annihilation at much larger forcing amplitude. We note that the exact amplitude at which annihilation first occurs depends slightly on the rotational phase of the spiral as it approaches the boundary. (Annihilation first occurs in the range $0.22 \lesssim A \lesssim 0.23$ depending on phase.) Likewise, the spiral phase can affect the reflected angle by nearly $1^\circ$ for $A \gtrsim 0.16$. The influence of phase is nevertheless small for the small-core spiral pinballs. It is, however, more pronounced in the large-core case which we shall now discuss.

### 2.3.2 Large-core case

We now turn to the case where unforced spiral waves rotate around a relatively large core region of unexcited medium. This case is illustrated in Fig. 2.8(a) where a rotating spiral wave and corresponding tip trajectory are shown in a
Figure 2.8: Illustration of a large-core spiral wave. (a) A portion of a rotating spiral and corresponding tip trajectory in a square region $40 \times 40$ space units [same size as Fig. 2.1(a)]. (b) Resonant forcing and reflection for a large-core spiral pinball shown in comparison to that of a small-core spiral pinball. The forcing amplitude is $A = 0.05$ in both cases.

region of space the same size as in Fig. 2.1. The larger tip orbit and unexcited core, as well as the longer spiral wavelength, in comparison with those of Fig. 2.1(a) are clearly evident. While such spirals occupy a relatively narrow region of parameter space, they are nevertheless of some interest because asymptotic treatments have some success in this region [MDZ94, HK99] and because this is nearly the same region of parameter space where wave-segments studies are performed [SMCS02, ZS05, STS08].

Figure 2.8(b) shows a typical case of non-specular reflection for a large-core spiral compared with a small-core spiral forced at the same amplitude. While many features are the same for the two cases, large-core spiral pinballs are found often to reflect at smaller $\theta_r$ and moreover, they can exhibit different qualitative phenomena.

Figures 2.9 and 2.10 summarise our findings for large-core spiral pinballs. Reflected angle as a function of incident angle for three forcing amplitudes is shown in Fig. 2.9. One sees the overall feature, as with the small-core case, that reflected angle is approximately constant over a large range of incident angles. This is particularly true of low-amplitude forcing, $A = 0.022$. However, there are also considerable differences with the small-core case.

For large-core spiral pinballs the reflected angle increases with forcing amplitude. This is opposite to what is found for the small-core spirals in Figs.
Figure 2.9: Summary of results for large-core spiral pinballs. Reflected angle is plotted versus incident angle for three forcing amplitudes as labelled. Neumann boundary conditions are used. For $A = 0.05$ spirals are frequently annihilated at the boundary, [Fig. 2.10(d)], over the range of incident angles indicated. For $A = 0.05$ and $\theta_i \leq -52^\circ$ trajectories glance from the boundary [Fig. 2.10(a)], leaving at $\theta_r \approx -85^\circ$ (not plotted). For $A = 0.036$ and $\theta_i \leq -52^\circ$ trajectories become permanently bound to the boundary. [See text and Fig. 2.10(b).] Wiggles in the higher amplitude results are the effect of incident phase.
Figure 2.10: Catalogue of interesting trajectories for large-core spiral pinballs. (a)–(c) show impacts with $\theta_i \approx -60^\circ$ at different forcing amplitudes. (a) With $A = 0.05$ the trajectory glances from the boundary and moves off nearly parallel to it ($\theta_r \approx -85^\circ$). (b) With lower amplitude $A = 0.036$, the trajectory becomes bound to the boundary. (c) With yet lower amplitude $A = 0.022$, the trajectory hugs the boundary for a while then leaves abruptly at an oblique angle ($\theta_r \approx 23^\circ$). In (d) the incoming spiral with large forcing, $A = 0.05$, is annihilated at the boundary. In (e) the effect of phase is seen with two approaching trajectories shifted by half a core diameter. Otherwise the conditions are identical, $A = 0.05$. The resulting reflected angles differ slightly.
2.5 and 2.6. Moreover, the reflected angles are noticeably smaller than for the small-core case, as was already observed in Fig. 2.8(b).

We now focus in more detail on what happens in various circumstances. The left portion of Fig. 2.9 indicates the different dynamics we observe, depending on forcing amplitude, at large negative incident angles ($\theta_i \leq -52^\circ$), and Figs. 2.10(a)-2.10(c) show representative trajectories with $\theta_i \approx -60^\circ$. At $A = 0.05$ trajectories glance off the boundary. That is, they remain close for short while before moving off with a well defined large negative reflected angle. The reflected angle is nearly constant at $\theta_r \approx -85^\circ$ for incident angles $\theta_i \leq -52^\circ$. At $A = 0.036$, $\theta_i \leq -52^\circ$, trajectories become bound to the boundary and move parallel to it indefinitely. In Fig. 2.10(c), with $A = 0.022$, one observes the trajectory moving along the boundary for a distance before abruptly leaving the boundary at a well-defined, relatively small positive reflected angle. This behaviour is not restricted to $\theta_i \leq -52^\circ$ and is observed until $\theta_i \approx +20^\circ$. In fact, this type of reflection is also observed for the other two forcing amplitudes studied for $\theta_i$ in a range above $-52^\circ$. For $A = 0.036$ this occurs until $\theta_i$ is approximately $-15^\circ$, while for $A = 0.05$ this is seen only until $\theta_i$ is about $-45^\circ$.

At the higher forcing amplitudes, as indicated for the case $A = 0.05$ in Fig. 2.9, large-core spirals are frequently annihilated when they come into contact with the boundary. Figure 2.10(d) shows a typical example. Whether or not a spiral is annihilated depends very much on the spiral phase on close approach to the boundary. The points shown in Fig. 2.9 with $A = 0.05$ are those where the trajectory reflected; the absence of points indicates annihilation. However, these results are for spiral pinballs all initiated a certain distance from the boundary. Changing that distance would affect the spiral phase at close approach and hence a different set of points would be obtained. Nevertheless, the marked range of frequent annihilation is indicative of what occurs at this forcing amplitude.

Finally, we address the wiggles in the reflected angle curves in Fig. 2.9, most evident at large forcing amplitudes. These wiggles are also due to the fact that the phase of spirals on close approach varies with incident angle. Figure 2.10(e) illustrates how the reflected angle depends on phase by showing
two trajectories shifted by half a core diameter. This shifts the spiral phase upon approach to the boundary and results in slightly different reflected angles. Rather than eliminating these wiggles by averaging over various initial spiral distances, we leave them in as an indication of the variability due to this effect. In general, reflections of large-core spiral pinballs are much more sensitive to phase than reflections of small-core spirals, and one should understand that the data in Fig. 2.9 will vary slightly if similar cases are run with spirals initiated at different distances from the boundary.

While we have not studied the step boundary in detail for large-core spiral pinballs, we have carried out a cursory investigation for such a boundary with $\Delta b = 0.035$. With the exception that there is no annihilation at the step boundary, we observe qualitatively similar behaviour to that just presented for the Neumann case. Most notably we find both glancing and bound trajectories.

2.4 Discussion

2.4.1 Biktashev-Holden theory

Many years ago Biktashev and Holden [BH93, BH95] carried out a study very similar in spirit to that presented here. Moreover, they understood that a primary cause for the reflection from boundaries was the small changes in spiral rotation frequency occurring as spiral cores came into interaction with boundaries. Based on this they proposed an appealing simple model to describe spiral pinball reflections. The model is based on the assumption that both the instantaneous drift speed normal to the boundary and spiral rotation frequency are affected by interactions with a boundary, with the interactions decreasing exponentially with distance from the boundary. While the actual interactions between spiral waves and boundaries are now known to be more complex (see below), it is worth investigating what these simple assumptions give. The beauty of the simple model is that it can be solved to obtain a relationship between reflected and incident angles, depending on only a single combination of phenomenological parameters. (They called this combination $\theta$, but we shall call it $p$. They also used different definitions for
incident and reflected angles.)

The model naturally predicts large ranges of approximately constant reflected angle depending on the value of $p$. What is nice is that while fitting the individual phenomenological parameters in their model would be difficult, it is also unnecessary. The value of $p$ can be selected to match the plateau value of $\theta_r$ observed in numerical simulations. Then the entire relationship between $\theta_r$ and $\theta_i$ from the theory is uniquely determined.

Curves from the Biktashev-Holden theory are included in Figs. 2.5 and 2.6. While there are obvious limitations to the theory, it is nevertheless interesting to see that some of the features are reproduced just from simple considerations. The theory would be expected to work best where the drift speed is small: low amplitude forcing. For the large-core spiral pinballs the theory does not apply and so the corresponding curves are not shown in Fig. 2.9.

### 2.4.2 Multiple reflections

As noted in the introduction, Fig. 2.1(d), when placed within a square domain the trajectory of a spiral pinball will typically approach a square, reflecting from each domain boundary such that $\theta_i + \theta_r = 90^\circ$. The reasons for this are simple (see for example Prati, et al. [PLTB11]), but a brief analysis is useful, particularly for understanding when square orbits become unstable.

![Figure 2.11](image-url)  
**Figure 2.11**: Sketch showing the geometry of multiple reflections in a portion of a square box of normalised length. $d^n$ is the location, relative to the length of a side, of the $n$-th reflection.

Figure 2.11 shows the geometry of a consecutive pair of reflections in the case where the reflected angle is larger than 45°. In this case the path will
necessarily strike consecutive sides of the domain. Consider first the path in terms of angles and let $\theta_i^n$ and $\theta_r^n$ denote, respectively, the $n$-th incident and reflected angles, starting from the initial reflection $\theta_i^0, \theta_r^0$. Trivially, the geometry of the square domain dictates that $\theta_i^{n+1} + \theta_r^n = 90^\circ$. Then, if the trajectory approaches an attracting path with constant angles, $\lim_{n \to \infty} \theta_i^n = \theta_i^*$, $\lim_{n \to \infty} \theta_r^n = \theta_r^*$, it must be that this path satisfies $\theta_i^* + \theta_r^* = 90^\circ$. That is, it must be a square or a rectangle. Denoting the relationship between incident and reflected angle by $\theta_r = \Theta(\theta_i)$, then a necessary condition for the square path to be attracting is that $|\Theta'(\theta_i^*)| < 1$. For the cases we have studied this is true since $|\Theta'(\theta_i^*)| = 0$.

Turning now to the points at which the path strikes the edge of the domain, we let $d^n$ denote the position of the $n$-th reflection along a given side, relative to the length of a side. One can easily see from the geometry that $d^{n+1} = (1 - d^n) \cot \theta_r^n$. Now, since $\theta_r \to \theta_r^*$, the fixed point $d^*$ is given by $d^* = (1 - d^*) \cot \theta_r^*$, or $d^* = 1/(1 + \tan \theta_r^*)$. This corresponds to a square trajectory. For example, from Fig. 2.5 with a forcing amplitude $A = 0.072$ one can see that $\theta_r^*$ will necessarily be about $74^\circ$, giving $d^* \approx 0.22$. These are the values seen in the simulation in Fig. 2.1(d). A necessary condition for this fixed point to be stable is $|\cot \theta_r^*| < 1$. For small-core spirals $\theta_r^* > 45^\circ$, so $\cot \theta_r^* < 1$, and hence their square paths are stable.

While square trajectories occur for small-core spiral pinballs, for large-core spirals other trajectories are possible. Examples are shown in Fig. 2.12. These occur when the reflected angle is smaller than $45^\circ$. (It is possible that $\theta_r < 45^\circ$ might occur for small-core spirals in some regimes, although we have not observed them.) When $\theta_r < 45^\circ$ it is not necessarily the case that trajectories will strike consecutive sides of a square box. This is seen in Fig. 2.12(a) where the spiral pinball reflects between opposite sides of the domain. The reflections satisfy $\theta_r = -\theta_i$.

The more interesting case is when $\theta_r$ is only slightly less than $45^\circ$ as is seen in Fig. 2.12(b). The square trajectory is unstable. While $|\Theta'(\theta_i^*)| < 1$ and the angles converge quickly to $\theta_i^n + \theta_r^n \approx 90^\circ$, the equation $d^{n+1} = (1 - d^n) \cot \theta_r^n$ exhibits growing period-two oscillations for $\cot \theta_r^n$ slightly larger than 1. Period-two oscillations in $d^n$ with $\theta_i^n + \theta_r^n \approx 90^\circ$ correspond to approximately
Figure 2.12: Examples of non-square paths for large-core spiral pinballs. (a) $A = 0.022$. The reflected angle is considerably smaller than $45^\circ$ and the resulting trajectory bounces between opposite sides of the domain. The spiral is shown faintly at one time instance. (b) $A = 0.034$. The square trajectory is unstable. For the first circuit around the nearly square path the full tip trajectory is plotted. Subsequently, for clarity only, the tip path sampled once per forcing period is shown. The trajectory collapses towards the diagonal. The final portion of the trajectory before the spiral pinball approaches the corners is shown in bold. The spiral undergoes a complicated reflection from the corner (grey, dotted).

rectangular trajectories that approach a diagonal. This ultimately leads the spiral into a corner of the domain where it may reflect in complicated manner.

The analysis just presented should not be viewed as a model for spiral pinball trajectories. Rather it just shows what global dynamics can be deduced simply from a measured relationship between incident and reflected angles. Essentially this same analysis appears as part of a study of cavity solitons [PLTB11] which also undergo non-specular reflections from walls and hence exhibit square orbits similar to Figure 2.1(d). Our simple analysis should be contrasted with the situation for drops bouncing on the surface of an oscillating liquid, so called walkers. Here physical models of the liquid surface and drop bounces account for many varied features of the system [PBC06, CF06, PBC08, FEB+10, ESM+11, Shi13]. Likewise, a predictive theory for spiral pinballs may be formulated using adjoint modes of the reaction-diffusion system known as response functions [BBBF10, BBB10]. This approach is presented in the following chapter. Memory effects are
important for walkers because bouncing drops interact with surface waves generated many oscillations in the past, and models necessarily take this into account [CF06, FEB+10, ESM+11, Shi13]. However, path memory is absent from spiral waves in excitable media and this constitutes a significant difference between the two systems.

2.4.3 Concluding remarks

We have reported some quantitative and some qualitative features of resonant-drift trajectories in excitable media. The main message is that reflections are far from specular—the reflected angle generally depends only weakly on the incident angle and typically is nearly constant over a substantial range of incident angles (particularly negative incident angles). Biktashev-Holden theory [BH93, BH95] accounts for some of the observed features, but a more detailed theory based on response functions [BB03, BBBF10, BBB10] is required. This follows in Chapter 3. We have seen that the behaviour of large-core spiral pinballs is more varied than that for small-core ones. Rather than simply reflecting from a boundary, large-core spirals may sometimes become bound to, or glance from, or be annihilated at a boundary, even at moderate forcing amplitudes. Finally we have considered what can occur as spiral pinballs undergo multiple reflections within a square domain, and in particular have shown that while small-core spirals are observed to meet the conditions of stable square trajectories, large-core spirals may fail to meet these conditions and exhibit more interesting dynamics.

We motivated this study with a broader discussion of macroscopic systems with wave-particle duality. A large number of analogues to quantum mechanical systems have been reported for walkers on the surface of a vibrated liquid layer [CFGB05, CPFB05, PBC06, CF06, EFMC09, EDFC09]. As far as we are aware, this is less the case for the propagating wave segments studied by Showalter et al. [SMCS02, ZS05, STS08] or the drifting spiral waves in excitable media considered here. (We examined briefly small-core drift trajectories through a single slit, but did not observe diffraction-like behaviour.) Nevertheless, for the reflection problem, spiral pinball trajectories, propa-
gating wave segments, cavity solitons, and walkers all share the feature of non-specular reflections [PBC06, STS08, PLTB11, CF12] and as a result these systems can show similar dynamics when undergoing multiple reflections within a bounded region [PBC06, EDFC09, STS08, PLTB11, Sho12]. It will be of interest to make further quantitative comparisons between these different systems in the future and to explore theoretical basis of this behaviour.

2.A Appendix: Resonant forcing parameters

In this Appendix we report the exact values for the forcing parameters used in the detailed quantitative incidence-reflection studies, since obtaining high-precision values for resonant drift can be time consuming. The values stated in the body of the chapter are reported only to two significant figures.

Table 2.1: Parameter values used to produce resonant (straight line) drift in the small-core case ($a = 0.8, b = 0.05$).

<table>
<thead>
<tr>
<th>$A$</th>
<th>$\omega_f$</th>
<th>$\Delta t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.044462</td>
<td>1.82</td>
<td>0.0187625</td>
</tr>
<tr>
<td>0.071868</td>
<td>1.792</td>
<td>0.0188508</td>
</tr>
<tr>
<td>0.102609</td>
<td>1.75</td>
<td>0.0188968</td>
</tr>
<tr>
<td>0.196132</td>
<td>1.63</td>
<td>0.0188957</td>
</tr>
</tbody>
</table>

Table 2.2: Parameter values used to produce resonant (straight line) drift in the large-core case ($a = 0.6, b = 0.07$).

<table>
<thead>
<tr>
<th>$A$</th>
<th>$\omega_f$</th>
<th>$\Delta t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.022</td>
<td>1.025</td>
<td>0.0188035</td>
</tr>
<tr>
<td>0.035863</td>
<td>1.003</td>
<td>0.0187557</td>
</tr>
<tr>
<td>0.050144</td>
<td>0.989</td>
<td>0.0187961</td>
</tr>
</tbody>
</table>
Chapter 3

Asymptotic dynamics of spiral pinballs

3.1 Introduction

In the past decade an intrinsic wave-particle dualism in spiral waves has been highlighted [BB03, BHB06, BBB10, BBBF10, BBS11]. This invites comparison with a growing number of macroscopic systems in which waves propagating in a nonlinear medium are associated with some degree of spatial localization [PLM+14], including liquid ‘walker’ droplets bouncing on a vibrated bath [CFGB05, CPFB05], various optical solitons [SCS00, GA12] and chemical wave segments [SMCS02]. Among other common properties, each of these examples exhibits nonspecular reflections from obstacles or medium perturbations [PBC06, EFMC09, Shi13, PLTB11, STS08] and the dynamics involved in the reflection process can be quite complex. It is within this context that we have undertaken the present investigation.

Our study focuses on rotating spiral waves in a system with excitable dynamics. First witnessed experimentally in the Belousov-Zhabotinsky chemical oscillator [Bel59, ZZ71, Win72], they have since been discovered in diverse biological [TD81, TAMM89, GB83, DPS+92, PDS+93], chemical [JRE+90, NvORE93, AS00] and physical [FRCG94] contexts. Within two-dimensional homogeneous excitable media, spiral waves typically rotate about an unexcited core
of fixed radius and centre. These are so-called rigidly rotating spirals. The rotation frequency is determined solely by medium properties, while the centre of rotation and phase are determined by initial conditions. However, applying spatial or temporal perturbations to an otherwise homogeneous medium can cause the wave pattern to undergo a spatial displacement or drift [Bik07, BBBF10]. By tracking either the local rotation centre, or the closely related wave tip, one may observe interesting trajectories as drifting spirals move through a medium.

A noteworthy case is resonant drift [ADM87, DZMB88, SZM93, BH93, ZSM94, MB96, ZWY+04, KJK05, NJBWHP06, XLQD12] in which spatially uniform periodic driving is applied in resonance with the spiral rotation frequency. In this case the spiral core travels in a straight line with constant velocity. In a typical experimental domain, such a spiral will inevitably come close to a boundary, which may lead to a reflection in the drift trajectory [BH93, OS08], as illustrated in Fig. 3.1. We refer to this system as a spiral pinball. Reflections are in general nonspecular: the incidence angle rarely equals the reflection angle. Furthermore, the character of individual reflection trajectories depends on the medium in which the wave propagates, the properties of the boundary and the spiral’s resonant drift velocity.

Numerical simulations of spiral pinball reflections were undertaken some time ago by Biktashev and Holden [BH93], who laid the foundations of the asymptotic approach in a subsequent study [BH95]. In Chapter 2, we updated their numerical work with more extensive simulations and the calculation of a large catalogue of reflection trajectories. A key feature of spiral pinball reflections in these two studies is that the angle of reflection is essentially independent of the angle of incidence for a large range of incident angles. Indeed, the reflection angle instead depends more strongly on the characteristics of the medium than on incident angle. This was predicted by Biktashev and Holden using an ODE model based on the simplifying assumption that the component of the spiral’s drift velocity caused by interaction with the boundary decays exponentially with distance from the boundary [BH93, BH95]. However, a more detailed theoretical treatment is required to fully understand the mechanism behind spiral pinball reflection. While separate theoretical accounts
Figure 3.1: Two examples of spiral pinballs reflecting in the Barkley model of a generic excitable medium. The trajectories of the wave tips are drawn in black. Arrows indicate the overall direction of drift. The spiral waves at the final point in the plotted trajectory are visualized by the $u$-field of the Barkley model. Both plots use the same length scale. The boundaries are generated by a step change in medium properties, indicated by grey shading at the left-hand edges. (a) A ‘small-core’ spiral pinball approaches a boundary and doubles back on itself; its reflection angle lies on the same side of the boundary normal as its incidence angle. (b) A ‘large-core’ spiral pinball speeds up close to the boundary and travels alongside it for a short while before reflecting sharply away. (The plots were cropped to $25 \times 40$ space units from simulations performed on a $50 \times 50$ square domain, discretised in space with grid spacing $h = 1/12$ and in time with time step $\Delta t = 2.09 \times 10^{-3}$. The step change was located 12 space units from the left-hand domain wall. Parameters: (a) $a = 0.8$, $b = 0.05$, $c = 0.02$, $\epsilon_s = 0.035$, $\epsilon_f = 1.44 \times 10^{-3}$, $\omega_f = 1.7893$; (b) $a = 0.6$, $b = 0.07$, $c = 0.02$, $\epsilon_s = 0.035$, $\epsilon_f = 4.4 \times 10^{-4}$, $\omega_f = 0.9504$. Details concerning these parameters and the methods used are given in Sec. 3.3.)
of both resonant drift [BH93, BH95, BEB99, BBBF10, XLQD12] and spatial medium inhomogeneities [EP86, AKM95, Bik00, XQD09, BBBF10] (which may act as boundaries to drift) already exist, it is the combination and interaction of these two phenomena which we must consider here.

A good candidate for an updated approach is to use the theory of response functions [BH95, BEB99, Bik00, BB03, BHB06, BBB+09, BBBF10] which has developed and matured in the years since the Biktashev-Holden study. Response functions are adjoint modes to the neutral symmetry modes of a spiral wave which characterize how the position and rotation phase of a spiral react to asymptotically small perturbations. In practical terms, response functions allow us to reduce the PDEs governing the whole medium to the dynamics of just three real variables—the two spatial coordinates of the wave rotation centre and the rotational phase.

In this chapter we bring the reflection of spiral pinballs into this asymptotic framework by considering the superposition of two small perturbations: one corresponding to resonant forcing generating drift and the other corresponding to a step change in a medium parameter acting as a boundary to drift. Previous studies addressed both effects independently using response functions [BEB99, BBBF10]. While the approach is strictly applicable only in the limit of slow resonant drift and weak boundary effects, we show that it nevertheless can capture, and thereby explain, most of the important features of spiral pinball reflections outside of this asymptotic limit.

### 3.2 Theory

The underlying dynamics of the excitable medium are well described by models in the class of reaction-diffusion PDEs on the plane:

\[
\partial_t u = D \nabla^2 u + f(u, p),
\]  

(3.1)

where \( u(x, t) \in \mathbb{R}^\ell \) is a vector of \( \ell \geq 2 \) state variables for the medium, \( f(u, p) \in \mathbb{R}^\ell \) describes the excitable dynamics at each point in space dependent on a vector of \( m \) parameters \( p \in \mathbb{R}^m \) and \( D \in \mathbb{R}^{\ell \times \ell} \) is a diffusion matrix.
We are interested in models that admit solutions rotating with angular frequency $\omega$ about a centre point $R = (X, Y)$. That is, rigidly rotating waves of the form

$$u = U(\rho, \theta + \omega t - \Phi),$$

(3.2)

where $(\rho, \theta)$ are polar coordinates centred at $R$ and $\Phi$ is the fiducial phase of the spiral at $t = 0$. Note that due to symmetries of the plane, if Eq. (3.1) admits a solution of the form in Eq. (3.2), then there are infinitely many such solutions related by symmetry, and this is captured by the fact that $R$ and $\Phi$ are arbitrary constants. We refer to $\omega$ as the natural frequency since it is an intrinsic property of the medium, whereas $R$ and $\Phi$ depend on initial data.

Suppose we perturb the medium slightly. In the limit of weak perturbations, this induces small shifts in the rotation centre $R$ and the phase $\Phi$, leaving the shape of the spiral otherwise unchanged. Thus the response of the spiral to weak perturbations is a trajectory through the space of solutions of the form Eq. (3.2), where $R$ and $\Phi$ depend on time.

Mathematically, we treat such a perturbation as the addition of a vector $|\epsilon h(x, t)| \ll 1$ to the right-hand side of Eq. (3.1). It can be shown using perturbation methods [BH95, BEB99, BHB06] that to first order in $\epsilon$, the time derivatives of $R(t)$ and $\Phi(t)$ are proportional to the $L^2$ inner products $\langle \cdot, \cdot \rangle$ of the spiral’s response functions $W_0$ and $W_1$ with the perturbation vector, averaged over one full rotation period $T = \frac{2\pi}{\omega}$:

$$\dot{\Phi}(t) = \frac{\epsilon}{T} \int_{t-T/2}^{t+T/2} \langle W_0, h \rangle \, d\tau + O(\epsilon^2),$$

(3.3a)

$$\dot{R}(t) = \frac{\epsilon}{T} \int_{t-T/2}^{t+T/2} e^{i(\Phi - \omega \tau)} \langle W_1, h \rangle \, d\tau + O(\epsilon^2),$$

(3.3b)

where we use the identification $R = (X, Y) \equiv X + iY$.

Technical details can be found in Appendix 3.A and elsewhere [BH95, BEB99, BB03, BHB06, BBB+09, BBBF10], but the essence of these equations is the following. The response functions are adjoint fields corresponding to the symmetries of the reaction-diffusion system [Eq. (3.1)]. $W_0$ is $\mathbb{R}^\ell$-valued and corresponds to the presence of rotational symmetry. One can think of the
perturbation, $\epsilon h$, as providing an infinitesimal impulse $\epsilon \langle W_0, h \rangle$ along the direction of the symmetry (phase $\Phi$ in this case), at each time $\tau$. Equation (3.3a) captures the effect of all such impulses over one rotation period to give the rate of change in $\Phi$.

The response function $W_1$ is $C^\ell$-valued and corresponds to the two translational symmetries. Here the perturbation at each time $\tau$ provides the spiral with an infinitesimal impulse in the direction $\arg \langle W_1, h \rangle$ rotated by $e^{i(\Phi - \omega \tau)}$ due to the underlying natural rotation of the spiral. These contributions, averaged over one rotation period, give the drift velocity. Importantly, a change in $\Phi$ typically implies a change in the direction of drift.

Response functions have been computed numerically for a variety of spiral waves in previous studies. For various cases, including that of the spiral waves we study here, the support of these functions was found to be highly localized around the spiral rotation centre [BB03, BHB06, BBBF10]. Thus, a spiral wave drifts only in response to perturbations very close to the core. That is, it behaves as a particle whose position may be identified with the rotation centre $R$.

We are interested in the case where a resonantly forced spiral moves towards, and reflects from, a boundary in the medium. This is a combination of two perturbations to the original reaction-diffusion equations—a homogeneous, time-periodic one that causes resonant drift of the spiral and a spatial one that imposes a boundary to the drifting spiral. Let us suppose the resonant forcing can be described by some $h_f(t)$. In practice we will consider the simple case of harmonic forcing of one of the medium parameters at the natural frequency $\omega$. Likewise, suppose that the effect of a boundary may be formulated in $h_s(x)$. The type of boundary we shall consider is a sharp interface along the line $x = 0$ between two media with different excitability properties. Although this is not a physical barrier to wave propagation, a drifting spiral core may nevertheless reflect from the spatial inhomogeneity; see Fig. 3.1 and Chapter 2. We refer to this as a step boundary. It may be considered as a weak perturbation provided that the step change in medium parameters is small. In previous studies a Neumann or ‘no-flux’ boundary was also considered. While this type of boundary condition cannot be treated
as a weak perturbation, we saw in the previous chapter that reflections from
a step inhomogeneity are qualitatively similar to the no-flux case.

The total perturbation to the medium can be written as $h(x,t) = \epsilon_s h_s(x) + \epsilon_f h_f(t)$, where $0 < \epsilon_s, \epsilon_f \ll 1$ represent the strengths of the respective ‘step’ and ‘forcing’ perturbations. One can immediately see from Eqs. (3.3a) and (3.3b) that the effects of the two perturbations on $\dot{\Phi}$ and $\dot{R}$ are a linear superposition and may therefore be considered separately. It may consequently be shown (see Appendix 3.A) that the equations of motion for the spiral centre $R = (X,Y)$ and phase $\Phi$ are of the form:

\[
\begin{align*}
\dot{X} &= \epsilon_s S_X(X) + \epsilon_f F_X(\Phi), \\
\dot{Y} &= \epsilon_s S_Y(X) + \epsilon_f F_Y(\Phi), \\
\dot{\Phi} &= \epsilon_s S_\Phi(X),
\end{align*}
\]

where $S_X, S_Y, S_\Phi$ are contributions due to the step boundary and $F_X, F_Y$ are contributions due to the resonant forcing. These are given by integrals of the form in Eqs. (3.3a) and (3.3b). While the functions depend in detail on the specific model used and the particular spiral wave under consideration, their general form, in particular their respective dependence on $X$ and $\Phi$ as indicated, is independent of these details.

Since the step boundary is located along the line $x = 0$ in the original PDE, the dynamics of the spiral depends only on the distance $X$ of the spiral centre from step boundary and does not depend on $Y$. Likewise, since the step perturbation is time independent, its effect, when averaged over a full spiral rotation, cannot depend on the spiral’s phase $\Phi$.

The form of the functions $F_X$ and $F_Y$ and the role of $\Phi$ are quite important. In Appendix 3.A.1 we show that for sinusoidal resonant forcing of a medium parameter:

\[
F(\Phi) = A e^{i\Phi},
\]

where $F \equiv F_X + iF_Y$ and $A$ is a real constant for each model and set of parameter choices. Hence, for a given spiral wave and given forcing amplitude, the drift velocity due to resonant forcing is, in the asymptotic limit, constant with
direction determined by the phase $\Phi$. This direction of drift can change as a result of interaction with the boundary, i.e., the function $S_\Phi$, but not due to periodic forcing alone.

Equations (3.4a), (3.4b) and (3.4c) reduce the spiral dynamics from a set of nonlinear PDEs to three coupled autonomous nonlinear ODEs. The functions $S_X$, $S_Y$, $S_\Phi$, $F_X$, and $F_Y$ on the right-hand sides must in practice be obtained numerically by taking appropriate inner products with numerically computed response functions. Nevertheless, evaluating the right-hand sides and then numerically solving the ODEs can be done quickly with minimal computational resources. It is worth noting that the essential dynamical quantities $X$, $Y$, and $\Phi$ are the same variables that Biktashev and Holden used in their asymptotic theory of spiral reflections [BH93, BH95]. Moreover, we stress that while the variable $\Phi$ was introduced as the phase of the spiral wave, its role in the reduced system becomes the direction of drift due to periodic forcing.

3.3 Model and methods

The previous discussion of response functions did not depend on any specific model. As in Chapter 2, we consider spiral wave solutions in the standard Barkley model [Bar91, Bar08], for which $\ell = 2$:

$$
\frac{\partial u}{\partial t} = \nabla^2 u + \frac{1}{c} u (1 - u) \left( u - \frac{v + b}{a} \right),
$$

(3.6a)

$$
\frac{\partial v}{\partial t} = u - v.
$$

(3.6b)

The two state variables $u(x, y, t)$ and $v(x, y, t)$ capture, respectively, the excitation and recovery processes of the medium. Parameters $a, b > 0$ control the threshold for excitation and $0 < c \ll 1$ sets the timescale of the fast excitation process, relative to recovery. The excitation parameter $c$ is usually called $\epsilon$. In this chapter only, we avoid calling it $\epsilon$, so that it is not confused with the standard notation for perturbation strengths. For fixed parameter $c$ and variable $a, b$, the section of parameter space which admits rigidly rotating
spiral wave solutions is divided roughly into two regimes distinguished by the size of the rotation core. We showed in Chapter 2 that the reflections of so-called small- and large-core spiral pinballs markedly differ and we therefore divide our study along these lines.

Throughout our study we have varied the \( b \) parameter to create the step inhomogeneity by considering \( b(x) = b_0 + \epsilon_s(H(x) - 1) \), where \( H \) is the Heaviside step function. Resonant forcing has been applied homogeneously by varying the excitability \( c \) as \( c(t) = c_0 + \epsilon_f \cos(\omega_f(t - t_0)) \), where \( \omega_f \) is the forcing frequency required to obtain resonant drift and \( t_0 \) is some initial forcing time (the choice of which is discussed in Appendix 3.A.1). For our results in Sec. 3.4, \( \omega_f = \omega \). In all numerical simulations, the values of \( \epsilon_s \) and \( \epsilon_f \) have been chosen small enough that the perturbed medium remains in the same parameter regime (of small- or large-core rigid rotation) as the unperturbed parameters.

The response functions and natural rotation frequencies for various small- and large-core spirals in the Barkley model were calculated on a polar grid using the software \textsc{DSpiral} [BBB+08]. The numerical methods are detailed in Ref. [BBB+09]. A disk of radius 15 was used in the small core with 64 angular grid points and 1875 radial grid points. In the large core the radius size was increased to 20 and the number of radial grid points used was 2500. The resulting response function discretisations were used to numerically compute the right-hand sides of Eqs. (3.4a), (3.4b), and (3.4c) (see the Appendix to this chapter for the specific integrals), again using \textsc{DSpiral}. Reflection trajectories were calculated by time-stepping the resulting three dynamical variables from chosen initial conditions.

Direct numerical simulations of the Barkley model PDEs were also performed for comparison with the response function predictions. These were computed using standard finite-difference techniques described in Refs. [Bar91, DMB97]. These simulations use unusually high precision to ensure that they correctly capture the spiral rotation frequency [BBBF10, Sec. IV B]. (The simulations involve forcing at the natural frequency, i.e. \( \omega_f = \omega \), obtained very accurately from \textsc{DSpiral}. Small inaccuracies in the simulations, which would normally be irrelevant, result in artificial frequency mismatches which
then lead to artificially curved trajectories.) In the small core (Fig. 3.12) a 20 × 20 square domain was used with grid spacing \( h = 0.0125 \) and time step \( \Delta t = 2.3 \times 10^{-5} \). The step inhomogeneity was located 5 space units from the left-hand domain edge. In the large core (Fig. 3.13) a larger 40 × 40 square domain was used, with the step inhomogeneity located 10 space units from the left-hand edge, in order to avoid interaction of the spiral wave with the no-flux domain walls. The grid spacing was \( h = 0.025 \), with corresponding time step \( \Delta t = 9.4 \times 10^{-5} \). Model parameter values are given later in the text.

### 3.4 Results

Before presenting our response function calculations, we make a note concerning incident and reflected angles. As is standard, we define both the angles of incidence \( \theta_i \) and reflection \( \theta_r \) to be measured from the boundary normal. In the case of light paths in classical optics, one considers incident angles only in the range \([0°, 90°]\), since, due to symmetry in the \( y \)-direction, trajectories at equal angles either side of the normal correspond to physically identical situations. However, since spirals possess a chirality, this symmetry is not present and we must consider both incident and reflected angles in the range \([-90°, 90°]\).

In Sec. 3.2 and the Appendix to this chapter we have implicitly set \( \omega > 0 \) to correspond to clockwise rotation. We consider spirals of this chirality only. Our convention is to define \( \theta_i \) to be positive in the clockwise direction from the normal and \( \theta_r \) to be positive in the anticlockwise direction from the normal. That is, incident and reflected angles on opposite sides of the normal have the same sign.

#### 3.4.1 Small-core case

Our study begins by considering spiral waves in the small-core region of parameter space. We set \( a = 0.8, b = 0.05, \) and \( c = 0.02 \). Figure 3.2 shows the step boundary functions \( S_X, S_Y, \) and \( S_\Phi \) for these parameters. These curves represent the intrinsic character of the boundary influence. Let us
Figure 3.2: $S_X$, $S_Y$, and $S_\Phi$ for a small-core spiral with $a = 0.8$, $b = 0.05$, and $c = 0.02$. Also plotted in dotted grey are the vertical lines $x = \pm 2.9$, which enclose the effective boundary region. [For $|x| > 2.9$, $S_X(x)$ and $S_Y(x)$ are less than 0.1% of $S_X(0)$ and $S_Y(0)$ respectively.]

first consider the effects of this boundary in the absence of resonant forcing. The dynamics of the spiral rotation centre in this case are governed simply by the $S_X$ and $S_Y$ curves, scaled by the size of the step:

$$\dot{R} = \epsilon_s S(X),$$

(3.7)

where $S \equiv S_X + iS_Y$. We see, as expected, that $S_X$ and $S_Y$ are zero outside a relatively small neighbourhood of $x = 0$ and thus spirals outside this region are unaffected by the step boundary. Since $S_X(X)$ is positive inside the boundary region, spirals to the right of $x = 0$ are repelled away from the step. Furthermore, since $S_Y(X)$ is also positive in this region, the boundary acts to intrinsically drive spirals in the positive $y$-direction. Note also the form of $S_\Phi$: far to the left of the boundary, $S_\Phi(X)$ tends to a non-zero (negative in this case) constant. This is because the spiral’s rotation frequency in the left half-plane, with the perturbed model parameter $b_0 - \epsilon_s$, differs from the ‘natural’ frequency $\omega$ of the unperturbed spiral in the right half-plane.
Now let us add in the effect of periodic forcing. The rotation centre in this case moves according to

\[ \dot{R} = \epsilon S(X) + \epsilon f \Phi, \]  

(3.8)

where \( F(\Phi) = A e^{i\Phi} \), from Eq. (3.5). Thus, the velocity at each instant is the superposition of the step component and a vector of fixed magnitude due to the resonant forcing, whose direction is set by the spiral’s phase \( \Phi \). Far from the boundary, the velocity is constant, since \( S(X) = 0 \) and \( S_\Phi(X) = 0 \) for \( X \gg 0 \). Close to the boundary, if the step perturbation is large enough relative to the resonant forcing perturbation, the boundary effects dominate and spirals in the positive half-plane are repelled from the step. Furthermore, since \( S_\Phi(X) < 0 \) for \( X \lesssim 2.9 \), the forcing component rotates clockwise in time while the spiral is in the boundary region.

This suggests a mechanism for reflection. Consider a resonantly forced spiral wave travelling towards the step from the right half-plane. Far from the boundary, the spiral drifts with constant velocity at some incident angle \( \theta_i \) (set by initial conditions). On entering the boundary region, the spiral is repelled by the inhomogeneity, causing it to slow down and preventing it from passing through \( x = 0 \). This effect itself does not cause the subsequent reflection from the boundary. The motion away from the boundary is rather due to the \( \Phi \) dynamics. As the spiral approaches the boundary, \( \Phi \) decreases bringing about a rotation in the resonant forcing component \( F(\Phi) \). After a time, this component inevitably rotates around to the positive \( x \)-direction and this drives the spiral away from the step. Consequently, the spiral leaves the boundary at some reflection angle \( \theta_r \), dictated by the phase on exiting the boundary region.

We see this mechanism at work in Fig. 3.3, which displays a typical theoretical reflection trajectory in the small-core regime. (One should note that the lengths of vectors in Fig. 3.3 have been scaled nonlinearly so their directions far from the step are discernible—the magnitude of the forcing component is comparatively much weaker than depicted.) After entering the boundary region, the spiral undergoes a rapid change in direction and phase.
Figure 3.3: Theoretical trajectory of a small-core spiral pinball reflection with $\theta_i = 0^\circ$ and $\epsilon_f/\epsilon_s = 1/25$. Initial conditions: $X_0 = 6$, $Y_0 = 0$, $\Phi_0 = \pi$. Each horizontal row of vectors plots the velocity field at the instant at which the spiral centre attained the given $Y$. These vectors depend on $X$ and the phase $\Phi$. The value of $\Phi$ at each horizontal slice is indicated on the right-hand side. Vector magnitudes have been scaled nonlinearly for visual clarity. The ratio of the $X:Y$ scales is $1:4$.

and its speed in the $x$-direction slows considerably. As the resonant forcing component $F(\Phi)$ (depicted in the rightmost vectors of Fig. 3.3) rotates with the decreasing phase, its $x$-component diminishes and consequently the boundary effects push the spiral centre further away from the step. This process is slow and the spiral travels far in the $y$-direction in this time. Eventually, the evolving phase turns the resonant drift direction towards the positive half-plane, i.e., $F_X(\Phi)$ changes sign and becomes positive. As a result, the spiral centre leaves the boundary. The reflected angle is close to $+90^\circ$, since $S_\Phi(X)$ is very near zero when this sign change occurs and therefore the phase changes only by a small amount after this.

We observe that the situation is similar across the full range of incident an-
Figure 3.4: Two theoretical trajectories in the small-core regime, initiated at $X_0 = 6$, $Y_0 = 0$. $\epsilon_f/\epsilon_x = 1/25$. The filled points plotted along the trajectories are equally spaced in time to indicate drift speed. Incident angles are (a) $\theta_i \approx -70^\circ$ and (b) $\theta_i \approx +70^\circ$. Both spiral pinballs reflect with angle $\theta_r \approx +88^\circ$. The ratio of the $X : Y$ scales is $1 : 1$.

Angles $\theta_i \in [-90^\circ, 90^\circ]$. Figure 3.4 displays two theoretical reflection trajectories which approach the boundary at different angles, either side of the normal, reflecting in the same direction. Regardless of incident angle, the spiral centre may only leave the boundary once $F(\Phi)$ points away from the step. Each spiral pinball reaches this sign change of $F_X(\Phi)$ in essentially the same state: with $\Phi = \pi/2$ and $X$ close to the edge of the boundary region. This is because the $\Phi$ dynamics are sufficiently slow that the spiral centre is pushed almost completely out of the boundary region by the time that $\Phi = \pi/2$. Therefore, each spiral pinball changes direction by only a small amount after this point and reflects with $\theta_r$ close to $+90^\circ$. On close inspection, another interesting feature of the trajectory in Fig. 3.4(a) is the change of curvature sign as the spiral enters and leaves the boundary region. (See also Fig. 3.5.)
Figure 3.5: $S_X$, $S_Y$, and $S_\Phi$ curves, together with representative theoretical reflection trajectories for three different small-core spiral pinballs. $\epsilon_f/\epsilon_s = 1/50$. Each pair of reflection trajectories is plotted below the corresponding boundary curves. The left- and right-hand trajectories are $\theta_l \approx 0^\circ$ and $\theta_l \approx -70^\circ$, respectively, and include filled points, matched to the time step of the corresponding points in Fig. 3.4, indicating drift speed. Model parameters: in (a) and (b) $a = 0.7$, $b = 0.01$; in (c) and (d) $a = 0.95$, $b = 0.01$; in (e) and (f) $a = 0.95$, $b = 0.08$. In all cases $c = 0.02$. These span a substantial extent of the small-core regime.
This can be attributed to the fact that the repulsion of the spiral in the \( X \) direction (which causes the spiral path to turn anticlockwise on entering and leaving) is felt in a slightly wider region than the effect of the boundary on the phase (which causes the spiral path to turn clockwise).

It is worth noting that in addition to the invariance of reflection angle, these theoretical trajectories exhibit qualitative features observed in numerical simulations. In particular, the nontrivial shape of Fig. 3.4(a), the sharp change of direction at the boundary in Fig. 3.3 and the decrease in the closest distance to the boundary reached by the spiral centre as \( \theta_i \) increases. For comparison see Figs. 2.3(b), 2.4(g) and 2.4(h) from the previous chapter.

Across the small-core parameter regime, we see that the curves \( S_X \), \( S_Y \) and \( S_\Phi \) vary in magnitude and shape. However, the qualitative differences in the theoretical reflection trajectories are only subtle and the reflection mechanism in each case is the same. Representative curves and trajectories are plotted in Fig. 3.5.

### 3.4.2 Large-core case

We now turn to the large-core case, setting \( a = 0.6 \), \( b = 0.07 \), and \( c = 0.02 \). As before, we begin by plotting the \( x \)-dependence of the key functions \( S_X \), \( S_Y \), and \( S_\Phi \), shown in Fig. 3.6. At first glance these do not appear to differ too much from the corresponding curves in the small core (see Figs. 3.2 and 3.5). Nevertheless, there are differences, some of which are quite important. The region of boundary influence is wider than in the small-core, extending to roughly a distance of five space units from the step inhomogeneity. This is expected: spiral waves propagate outwards from their tips, which rotate around a circle of much larger radius in the large-core. Furthermore, \( S_X \) has roots within this boundary region, at \( x \approx \pm 2.5 \). The root at positive \( x \) is attracting (in the absence of resonant forcing). Also, the magnitudes of the curves are (pointwise) greater than those in the small-core case. For the set of parameters we consider, this is particularly true for \( S_\Phi \). Finally, notice that \( S_Y \) has changed sign with respect to the small-core case.

These differences have a significant impact on the character of reflections
Figure 3.6: $S_X$, $S_Y$ and $S_\Phi$ for a large-core spiral with $a = 0.6$, $b = 0.07$, $c = 0.02$. Also plotted in dotted grey are the vertical lines $x = \pm 5.0$, which enclose the effective boundary region. [For $|x| > 5.0$, $|S_X(x)|$ and $|S_Y(x)|$ are less than 0.1% of $S_X(0)$ and $S_Y(0)$ respectively.]

for spiral pinballs in the large-core region. Figure 3.7 demonstrates a typical theoretical trajectory. Approaching at $\theta_i = 0^\circ$, the spiral changes direction as it enters the boundary region as before, but turns to move in the negative rather than the positive $y$-direction, since $S_Y$ is large and negative inside the boundary region. While $\pi/2 < \Phi < \pi$, the resonant forcing has negative $x$-component and the spiral remains near the positive root of $S_X$. Since $\Phi$ decreases to less than $\pi/2$, the forcing acts to push the spiral away from the boundary. As it exits, $\Phi$ continues to decrease causing the resonant forcing direction to turn further clockwise. Finally, the spiral leaves the boundary at the constant angle dictated by $\Phi = -0.17$ ($\theta_r \approx -9.5$ in this case). Qualitatively similar trajectories for low amplitude resonant forcing in the large core were observed in the previous chapter for Neumann boundary conditions: see Fig. 2.10(c).

The key difference between this large-core case and the small-core theoretical trajectories in Sec. 3.4.1 is the attracting root of the $S_X$ curve, which importantly occurs within the boundary region. While the spiral is in the
boundary region, the phase evolves, causing the resonant forcing component to rotate, just as with small-core spirals. Once $F_X(\Phi)$ changes sign, the resonant forcing turns to impel the spiral away from the boundary. While in the small-core cases this occurs when the spiral centre is near to the end of the boundary region, in the large-core case the spiral remains close to the attracting root of $S_X$ prior to the sign change. Since the magnitude of $S_\Phi$ is non-negligible near the attracting root of $S_X$, $\Phi$ continues to evolve, changing the sign of $F_X$ and decreasing for some time thereafter as the spiral exits the boundary. Consequently, the final direction of the spiral differs greatly from $+90^\circ$.

In the large-core regime, we see a notable effect of incident angle on reflection angle. Using the same parameters, we demonstrate this in Fig. 3.8. Spiral pinballs approaching the boundary at higher incidence angles have
Figure 3.8: Effect of incident angle $\theta_i$ for a large-core spiral pinball. Various theoretical trajectories are shown with different initial $\Phi_0$ and $\epsilon_f/\epsilon_s = 1/87.5$. The filled points plotted along the trajectories are equally spaced in time to indicate drift speed. Incident angles: (a) $\theta_i = 60^\circ$, (b) $\theta_i = 67.5^\circ$, (c) $\theta_i = 75^\circ$.

Figure 3.9: Phase dynamics for large-core spiral pinballs approaching the boundary with different incident angles. $\epsilon_f/\epsilon_s = 1/87.5$. The top plot shows the curve $S_X$, for reference. The bottom plot shows the theoretical ‘trajectory’ of the spiral phase as the spiral moves in and out of the boundary region, for various incident angles. Incoming trajectories have $\Phi \in (\pi/2, 3\pi/2)$ and outgoing trajectories have $\Phi \in (-\pi/2, \pi/2)$. The solid black trajectory corresponds to the reflection in Fig. 3.7 and the dotted and dashed trajectories correspond to the reflections in Fig. 3.8.
lower initial phase and consequently reach the sign change of $F_X(\Phi)$ (at $\Phi = \pi/2$) sooner. Therefore, at high incident angles the sign change occurs much further from the step than at low incident angles, since $\Phi$ reaches $\pi/2$ before the spiral centre reaches the attracting root of $S_X$. This means these spirals necessarily leave the boundary region sooner and with a greater $\Phi$, i.e., greater reflected angle. This can be visualized more clearly by plotting the trajectory of the phase with respect to the distance from the boundary, as we have done in Fig. 3.9.

The change in sign of the $S_Y$ curve between the large- and small-core parameter regimes has no effect on reflection angle, since the dynamics of the spiral centre far from the boundary depends only on $\Phi$ and $X$. However, it is relevant to the overall qualitative shape of trajectories at the boundary. This difference in sign can be qualitatively explained by referring to arguments originally given by Krinsky et al. [KHV96] for the case of spiral wave drift in electric fields, which were later studied by Xu et al. [XQD09] for medium inhomogeneities. Drift of the spiral rotation centre may be caused by changes to the radius of the rotation core and also by changes to the rotation frequency. In the Barkley model, decreasing the $b$ parameter, as we have done to form the step boundary, decreases the core size and increases the rotation frequency. The effect of our step inhomogeneity on the core radius causes the spiral to drift in the negative $y$-direction. However, the effect on the rotation frequency causes the spiral to drift in the positive $y$-direction. For small-core parameters, the core radius changes little and the effect of the step boundary on the rotation frequency dominates. In the large-core parameter region, it is instead the changes in the core radius which dominate. Therefore the vertical component of drift due to the boundary changes sign between the two parameter regions.

We may also consider the effects of altering the ratio $\epsilon_f/\epsilon_s$. Let us fix $\epsilon_s$ and vary $\epsilon_f$. Higher $\epsilon_f$ corresponds to higher amplitude resonant forcing, meaning that the drift speed due to resonant forcing is greater. Figure 3.10 plots some illustrative theoretical reflection trajectories at different amplitudes. We see that as resonant forcing amplitude increases, reflected angle increases. This is because higher amplitude forcing impels spirals with greater drift speed. Faster spirals leave the boundary more quickly after $F_X(\Phi)$ changes sign and
**Figure 3.10:** Effect of forcing amplitude on large-core spiral pinballs. Three theoretical trajectories are shown in order of increasing amplitude and include filled points, matched to the time step of the corresponding points in Fig. 3.8, indicating drift speed. The perturbation ratio $\epsilon_f/\epsilon_s$ in each case equals (a) 1/75, (b) 1/50, and (c) 1/25.

**Figure 3.11:** Reflected angle $\theta_r$ versus incident angle $\theta_i$ for large-core spiral pinballs at different forcing amplitudes. The angles were measured from the theoretical response function trajectories at the given $\epsilon_f/\epsilon_s$ ratios.

therefore leave with a greater $\Phi$. (Note that they also approach closer to the step, which acts to decrease reflection angle, but this effect is not significant relative to the effect of increased drift speed.)

The combined effects of incidence angle and forcing amplitude are illustrated in Fig. 3.11, where we plot reflected angle $\theta_r$ versus incident angle $\theta_i$ for the three forcing amplitudes used in Fig. 3.10. These theoretical incidence-reflection data are qualitatively close to the large-core results from direct numerical simulation presented in the previous chapter (albeit with Neumann boundary conditions): see Fig. 2.9.
3.4.3 Comparison with direct numerical simulation

Figures 3.12 and 3.13 show comparisons between the reflection trajectory predicted by our response function calculations and results from direct numerical simulation (DNS) of the full Barkley model PDEs using the same parameters.

A thorough study of the numerical convergence of the asymptotic theory in the separate cases of resonant parameter forcing and step inhomogeneity has previously been conducted [BBBF10] and consequently we do not repeat such a study here. Instead, the cases presented have been chosen to demonstrate various phenomena predicted theoretically in the preceding sections. Excellent agreement is seen between theory and full DNS of spiral pinballs over a broad range of parameters and conditions.

In the small-core cases, Fig. 3.12, the spiral pinball drift direction, drift speed, and point of closest approach to the boundary are in very close correspondence with theoretical predictions. Note that speed is gauged from the distance travelled between successive points (open circle for DNS and filled circles for theory). Most of the (very small) differences between DNS and theory arise in the vicinity of the boundary where the effects of both perturbations are felt. Since points are plotted at fixed time intervals over the full trajectory, small speed differences can nevertheless give rise to an accumulated shift between points from DNS and theory. The most striking feature in the small-core regime is the correct theoretical prediction at large negative incident angles: Figs. 3.12(d) and 3.12(f). Theory correctly predicts that the spiral centre first moves downward near the boundary for a large number of rotation periods before turning, moving upward, and slowly leaving the boundary.

In the case of large-core spiral pinballs, Fig. 3.13, the considerable variation in the reflected angle predicted by theory is seen to hold in the full DNS. In particular, for fixed parameter values, as the incident angle is changed from near zero, Fig. 3.13(a), to large positive angles, Figs. 3.13(b) and 3.13(c), the drift trajectory spends less time in the vicinity of the boundary and develops a loop as the reflected angle changes from negative (moving down and to the right in
Figure 3.12: Comparison between theory and DNS of the Barkley model for a variety of parameter values and incident angles in the small-core regime. In each case the rotation centre of the spiral wave in the DNS is plotted (open circles) every 30th rotation period. The theoretical trajectories (curves with solid dots) use an initial condition selected such that they agree with the DNS trajectory at a point close to the boundary. Solid dots are separated by a time corresponding to 30 rotation periods. Also shown are the rotating spiral tip trajectories, dotted in grey, and the step boundary at $x = 0$, dashed in grey. Each of the three columns corresponds to a different choice of model parameters broadly spanning the small-core parameter regime. Within each column two incident angles are shown: one normal and one oblique to the boundary. Parameters in (a) and (b): $a = 0.8$, $b = 0.05$, $c = 0.02$, $\omega_f = 1.850564$, $\epsilon_f/\epsilon_s = 4 \times 10^{-5}/3.5 \times 10^{-3} = 1/87.5$; in (c) and (d): $a = 0.7$, $b = 0.01$, $c = 0.02$, $\omega_f = 2.043489$, $\epsilon_f/\epsilon_s = 4 \times 10^{-5}/2 \times 10^{-3} = 1/50$; in (e) and (f): $a = 0.95$, $b = 0.08$, $c = 0.02$, $\omega_f = 1.768359$, $\epsilon_f/\epsilon_s = 4 \times 10^{-5}/2 \times 10^{-3} = 1/50$. 
Figure 3.13: Comparison between theory and DNS of the Barkley model in the large-core regime verifying the theoretical predictions for the role of incident angle and forcing amplitude. Plot (a) shows a simulation of the case explained theoretically in Fig. 3.7, in which $\theta_i \approx 0^\circ$. Between (a), (b) and (c), the incident angle was varied from $0^\circ$ to approximately $60^\circ$ and $70^\circ$, respectively, keeping all other parameters fixed. In (d), the resonant forcing perturbation used was twice that of (a), while the incident angle and all other parameters remained fixed. Thus the effects of incident angle and forcing amplitude are seen to agree with those predicted in Figs. 3.8 and 3.10 and explained theoretically in Sec. 3.4.2. In each case the rotation centre of the spiral in the DNS is plotted (open circles) every 10th rotation period. The theoretical trajectories (curves with solid dots) use an initial condition selected such that they agree with the DNS trajectory at a point close to the boundary and are plotted with a time step (time between successive points) corresponding to 10 rotation periods of the simulation. Also shown are the rotating spiral tip trajectories, dotted in grey, and the step boundary at $x = 0$, dashed in grey. Parameters: $a = 0.6$, $b = 0.07$, $c = 0.02$, and $\omega_f = 0.9164372$; in (a)–(c) $\epsilon_f/\epsilon_s = 4 \times 10^{-5}/3.5 \times 10^{-3} = 1/87.5$; in (d) $\epsilon_f/\epsilon_s = 8 \times 10^{-5}/3.5 \times 10^{-3} = 1/43.75$. 
the figure) to positive (moving up and to the right). (See also for comparison Fig. 3.8.) Furthermore, as the forcing amplitude is increased for otherwise fixed conditions [Fig. 3.13(a) and Fig. 3.13(d)], the time at the boundary decreases and the reflection angle increases. (See also for comparison Fig. 3.10.)

The agreement between asymptotics and DNS is not quite as good in the large-core results, Fig. 3.13, as in the small-core results, Fig. 3.12. The main visible difference between theory and DNS in the large core regime is the point at which the spiral centre leaves the boundary. Other features, such as the reflected angle and the point of closest approach are predicted well. Discrepancies between theory and DNS are due to slight frequency mismatches. Large-core spiral waves are particularly susceptible to this as their rotation frequencies and tip orbits vary rapidly with parameters [Win91]. In the DNS there is a shift from the unperturbed rotation frequency \( \omega \) (as calculated to high accuracy by \textsc{dxspiral}) due to small but finite spatial discretisation errors, as well as weak nonlinear effects at finite perturbation strength. As the perturbation magnitudes and the computational grid spacing tend to zero, the theoretical and DNS trajectories do converge [BBBF10].

### 3.5 Discussion

We have applied the theory of response functions to the reflection of spiral pinball trajectories from boundaries. Via numerical computation of response functions, we have studied reflections in the asymptotic limit of slow drift and weak boundary effects. In this limit the approach is quantitatively accurate, as we have demonstrated for a variety of cases by comparing direct simulations of spiral pinballs in a full reaction-diffusion model with the theoretical predictions from response functions. However, the main value of the response function approach is the *qualitative* understanding it brings to how interactions with a boundary lead to different types of reflections in various situations. Several of the most significant features of spiral pinball reflections, observed in simulations at higher drift speeds and greater step inhomogeneities in the previous chapter, are nevertheless captured qualitatively.
by the asymptotic analysis. Consequently, we have been able to understand
the essential causes of many interesting aspects of spiral wave reflections.

As stated in Section 3.1, the primary characteristic of spiral pinball re-
fections is that across a wide range of model parameters, the reflected angle
is approximately constant for large ranges of incident angle. This reflection
angle ‘plateau’ is present in the response function results in both small- and
large-core cases. In the small-core case, we demonstrated numerically in
Chapter 2 that the value of this constant angle increases toward \( \theta_r = +90^\circ \)
as the resonant drift velocity decreases. The asymptotic results reveal the limit-
ing case of this trend, yielding only reflected angles very close to \( \theta_r = +90^\circ \)
and we have shown exactly why the reflected angle is essentially constant
across a wide range of parameter space.

Another significant feature observed in the Chapter 2 results is that, un-
lke the small-core case, for large-core spirals the reflected angle increases
with increasing drift velocity. This effect is clearly present in the asymptotics
(Figs. 3.10 and 3.11) and in the comparison with DNS [Figs. 3.13(a) and 3.13(d)].
The qualitative form of the reflection angle data in Fig. 3.11—a plateau for neg-
ative \( \theta_i \), then monotonically increasing at high \( \theta_i \)—is familiar to all previous
numerical results and emerges naturally from the response function model
by considering Fig. 3.9. Furthermore, general consideration of the differences
between small- and large-core spiral pinballs at the asymptotic level has led
to explanations of the diversity of behaviours between the two cases. Finally,
we note that the non-trivial shape, closest boundary approach distance, and
relative drift speeds that are obtained and explained via the response function
analysis were all observed qualitatively beyond the asymptotic limit, in both
the small- and large-core cases.

The work presented in this chapter fits comfortably with that which is
already known about spiral pinball reflections. Biktashev and Holden [BH93]
recognized many years ago that reflections are caused by small deviations
from the natural rotation frequency on close approach to the boundary,
which alter the direction of drift. They proposed asymptotic equations of
motion for the rotation centre and phase, positing that the boundary effects
(corresponding to \( S_X \), \( S_Y \), and \( S_\Phi \) in our notation) decay exponentially with
distance from the boundary. These simple assumptions ably capture the
overriding feature of spiral pinball reflections—large ranges of approximately
constant reflected angle—but beyond that the predictive qualities of the model
are limited. Our application of response functions to the reflection problem
can be viewed as an extension of their efforts, removing the phenomenology
for the case of a step boundary and allowing the boundary effects to be
calculated accurately for any spiral wave. This extra information yields a much
more detailed picture of the reflection dynamics, capturing the behaviour near
to the boundary as well as far from it and producing qualitatively meaningful
reflection trajectories across a wide range of parameters. Furthermore, we
have calculated response functions in the large-core regime, which was not
considered by Biktashev and Holden. Here, we observe that the repulsive
effect on spirals’ velocity normal to boundary ($S_X$) decays more rapidly than
the effect on the phase ($S_\Phi$)—a finding which accounts for the differences
between small- and large-core reflection angle results. This could not have
been captured by the original Biktashev-Holden theory which for simplicity
assumed that all boundary effects decay with respect to the same length scale.

Beyond the features of spiral pinball reflections considered here, there are
phenomena outside the asymptotic limit of small perturbations that are not
predicted by the linear order response function approach. In the small-core
regime, a wider range of reflection angles are observed at higher forcing
amplitudes than is captured by the asymptotic analysis. In the small-core
regime, as observed in Chapter 2, there exist so-called ‘glancing’ and ‘bind-
ing’ trajectories in which spiral waves respectively become temporarily and
permanently attached to the boundary. In order to predict these behaviours,
a mechanism is needed through which spirals may become attached. While
the function $S_X$ in the large core (see Fig. 3.6) does imply that there exists an
attracting point near to the boundary, only very weakly forced spiral waves
can become trapped there. This is contrary to our Chapter 2 DNS results for
finite amplitude large-core spiral pinballs, in which the attachment of the
spiral to the boundary ‘strengthens’ (from reflection, to glancing, to binding)
as forcing amplitude increases (see Fig. 2.9).

It would be desirable to address these phenomena theoretically—especially
the attachment behaviours which are particularly at odds with what we have seen in the asymptotics. One potential approach could be to use a kinematic model, similar to the one introduced by Di et al. in Ref. [DQWG03]. The principal idea is to split the motion of the spiral tip into angular and radial components, which depend on the tip rotation radius \( R_c \) and rotation period \( T \).

The dependence of \( R_c \) and \( T \) on the medium parameters (or on some external perturbation) may be determined empirically by direct simulation and thus used to model drift in a given scenario. Recent papers have employed this method to reproduce the tip dynamics of small- and large-core spirals in the presence of a step inhomogeneity [XQD09] and under periodic forcing of excitability [XLQD12]. This suggests that a similar approach could be used to model spiral wave reflections. It remains to be seen whether, given suitable modelling assumptions, predictive power outside the limit of small perturbations could be obtained.

3.A Appendix: Response function theory derivations

In this Appendix we present the derivation of the response function inner products that make up the differential equations in Eqs. (3.4a), (3.4b) and (3.4c).

The perturbations we have considered above are small temporal and spatial variations in the medium parameters. (While we could have perturbed the PDE fields directly instead, parameter variation is preferred since it is directly analogous to the way in which experiments on excitable media are often conducted [ADM87, SZM93, ZSM94, KJK05].)

Let us consider such a parameter \( p \) and take its dependence on \((x, t)\) to be of the form \( p(x, t) = p_0 + \epsilon p_1(x, t) \) for some constants \( p_0 \) and \( 0 < \epsilon \ll 1 \). Taylor expansion of Eq. (3.1) to first order in \( \epsilon \) establishes that parameter variations of this form may be considered as additive perturbations to the reaction diffusion system:

\[
\partial_t u = D \nabla^2 u + f(u, p_0) + \epsilon h(u, x, t),
\]  

(3.9)
where \( h(u, x, t) = \partial_p f(u, p_0)p_1(x, t) \).

### 3.A.1 Resonant forcing

Sinusoidal variation of a parameter at the natural frequency \( \omega \) induces resonant drift. Consider \( p \) varying as \( p(t) = p_0 + \varepsilon_f \cos(\omega(t - t_0)) \), where \( t_0 \) is some initial time whose role will become apparent below. Then the perturbation \( h_f \), in the form depicted in Eq. (3.9), is 
\[
    h_f(u, t) = \partial_p f(u, p_0) \cos(\omega(t - t_0)).
\]

To derive the dynamical equations for \( \Phi \) and \( R \), we must perform the integrations in Eqs. (3.3a) and (3.3b). Note that since the sinusoidal term does not depend on space:
\[
    \langle W_n, h_f \rangle = \cos(\omega(t - t_0))\langle W_n, \partial_p f(u, p_0) \rangle,
\]
for \( n = 0, 1 \). Furthermore, both \( W_n \) and \( \partial_p f \) depend on time only via their dependence on the wave field \( u \). Since \( u \) is stationary in a reference frame centred at \( R \) and rotating with frequency \( \omega \), the inner products \( \langle W_n, \partial_p f \rangle \) are time independent. Therefore, we have
\[
    \int_{t-T/2}^{t+T/2} \langle W_0, h_f \rangle \, d\tau = 0 \tag{3.11}
\]
and
\[
    \int_{t-T/2}^{t+T/2} e^{i(\Phi - \omega \tau)}\langle W_1, h_f \rangle \, d\tau = \frac{1}{2}Te^{i(\Phi - \omega t_0)}\langle W_1, \partial_p f \rangle. \tag{3.12}
\]
We set the initial forcing time \( t_0 \) such that \(-\omega t_0 + \arg\langle W_1, \partial_p f \rangle = 0\). Therefore the equations of motion for a sinusoidally forced spiral are, due to Eqs. (3.3a), (3.3b), (3.11), and (3.12):
\[
    \dot{\Phi} = 0, \quad \dot{R} = \varepsilon_f A e^{i\Phi} = \varepsilon_f F(\Phi), \tag{3.13}
\]
where \( A(u, p_0) := \frac{1}{2}|\langle W_1, \partial_p f(u, p_0) \rangle| \) is a real constant with respect to space and time for a given model and set of parameters. We can thus unambiguously identify the phase variable \( \Phi \) with the direction of drift due to resonant forcing and it is for this reason that \( t_0 \) was introduced.
3.A.2 Step boundary

The step boundary is a step inhomogeneity in a medium parameter that for convenience we locate at $x = 0$. Therefore, the parameter $p$ varies in space as $p(x) = p_0 + \epsilon_s (H(x) - 1)$, where $H$ is the Heaviside step function. The perturbation $h_s$ is thus $h_s(u, x) = \partial_p f(u, p_0)(H(x) - 1)$.

The integrals in Eqs. (3.3a) and (3.3b) are considered here in a co-ordinate system that rotates with the spiral wave at its natural frequency and is centred at $R(t)$ [BH95, BB03, BBBF10]. Let $(\rho, \vartheta)$ be polar co-ordinates centred at $R(t)$. Then define the rotating angular co-ordinate $\theta = \vartheta + \phi(t)$, where $\phi(t) := \omega t - \Phi(t)$ is the angle that the spiral turns through in time $t$. The co-ordinates $(\rho, \theta, \phi)$ define a frame in which the spiral wave $U$ [see Eq. (3.2)] and its response functions $W_0$ and $W_1$ are constant.

In this frame the time-averaging integration in Eqs. (3.3a) and (3.3b) becomes averaging over $\phi$. [Note that since the perturbation $h_s$ does not depend on time this averaging need not be centred about $\phi(t)$ and hence we take the range of integration to be simply $[0, 2\pi]$]. We obtain

$$\frac{1}{T} \int_{t-T/2}^{t+T/2} e^{i n (\Phi - \omega t)} \langle W_n, h_s \rangle \, d\tau =$$

$$= \frac{1}{2\pi} \int_0^{2\pi} e^{-i n \phi} \int_0^{2\pi} \int_0^{\infty} w_n(\rho, \theta) \tilde{p}_1(\rho, \theta, \phi) \rho \, d\rho \, d\theta \, d\phi,$$

(3.14)

for $n = 0, 1$, where $\tilde{p}_1$ represents the spatial variation of $p$ written in the co-rotating frame, which is

$$\tilde{p}_1(\rho, \theta, \phi) = H(X + \rho \cos(\theta - \phi)) - 1$$

(3.15)

and we have made use of the shorthand $w_n := [W_n(\rho, \theta)]^* \cdot \partial_p f(U, p_0)$.

We can compute the integral over $\phi$ explicitly. Changing the co-ordinate
to $\vartheta$ and rescaling the step function, we have

$$
\frac{1}{2\pi} \int_0^{2\pi} e^{-in\vartheta} \tilde{p}_1(\rho, \vartheta, \varphi) \, d\varphi = 
\frac{1}{2\pi} e^{-in\theta} \int_0^{2\pi} e^{in\vartheta} (H(X/\rho + \cos(\vartheta)) - 1) \, d\vartheta.
$$

(3.16)

As discussed in the main text, we see that the integral depends on the distance of the spiral centre to the step inhomogeneity. There are three cases to consider:

1. $|X| > \rho$ and $X > 0 \implies H(X/\rho + \cos(\vartheta)) - 1 = 0$

2. $|X| > \rho$ and $X < 0 \implies H(X/\rho + \cos(\vartheta)) - 1 = -1$

3. $|X| < \rho$, in which case $H(X/\rho + \cos(\vartheta)) - 1 = -1$ if $\vartheta \in \left[-\pi, -\arccos(-X/\rho)\right] \cup \left[\arccos(-X/\rho), \pi\right]$ and is zero otherwise.

For the case $n = 0$, i.e., the $\Phi$ dynamics, we therefore have

$$
\frac{1}{2\pi} \int_0^{2\pi} \tilde{p}_1 \, d\varphi = \begin{cases} 
H(X) - 1 & \text{if } \rho < |X|, \\
\frac{1}{\pi} \arccos(-X/\rho) - 1 & \text{if } \rho > |X|.
\end{cases}
$$

(3.17)

For the case $n = 1$, i.e., the $R$ dynamics, after some work one obtains

$$
\frac{1}{2\pi} \int_0^{2\pi} e^{-i\varphi} \tilde{p}_1 \, d\varphi = \begin{cases} 
0 & \text{if } \rho < |X|, \\
\frac{1}{\pi \rho} e^{-i\theta} \sqrt{\rho^2 - X^2} & \text{if } \rho > |X|.
\end{cases}
$$

(3.18)

Combining the results in Eqs. (3.17) and (3.18) with Eqs. (3.14) and (3.3a) we see that the dynamics for a spiral wave interacting with a step boundary are of the form

$$
\dot{\Phi} = \epsilon_s S_\Phi(X), \quad \dot{R} = \epsilon_s S(X),
$$

(3.19)
where

\[
S_\Phi(X) = \int_0^{2\pi} \int_0^{|X|} w_0(\rho, \theta)(H(X) - 1)\rho \, d\rho \, d\theta \\
+ \int_0^{2\pi} \int_{|X|}^{\infty} w_0(\rho, \theta) \left[ \frac{1}{\pi} \arccos(-X/\rho) - 1 \right] \rho \, d\rho \, d\theta
\]  
(3.20)

and

\[
S(X) = \frac{1}{\pi} \int_0^{2\pi} \int_{|X|}^{\infty} w_1(\rho, \theta) e^{-i\theta} \sqrt{\rho^2 - X^2} \, d\rho \, d\theta.
\]  
(3.21)

As argued in Sec. 3.2, the asymptotics for the forcing and step perturbations linearly superpose, providing the full picture of the dynamics of a resonantly forced spiral waves interacting with a step boundary. This is displayed in Eqs. (3.4a), (3.4b), and (3.4c) with the \( R \) dynamics separated into \( X \) and \( Y \) components: \( S_X := \text{Re}(S), S_Y := \text{Im}(S), F_X := \text{Re}(F), \) and \( F_Y := \text{Im}(F). \)
Chapter 4

Spiral waves in an elastic medium

4.1 Introduction

As discussed in previous chapters, spiral waves occur in nature in many diverse contexts. Real-world excitable media do not always conform to the ideal case of the flat homogeneous model used above. They may be curved, anisotropic, heterogeneous, or may deform in a nontrivial way. In this chapter, we shall be particularly interested in spiral waves in media undergoing deformation.

The most important example of spiral wave activity in a deforming medium is in the heart. They were first demonstrated in this context by Davidenko et al., who initiated spiral waves in small slices of tissue cut from dog and sheep hearts [DPS+92, PDS+93]. At the time of these experiments, it was already well understood that cardiac tissue functions as an excitable medium and the existence of rotating waves in the ventricles was thought to provide a good candidate explanation for certain kinds of irregular heartbeat rhythms [Win89]. Such pathologies can be serious enough to deprive the heart of its normal function, resulting in sudden cardiac death. Consequently, efforts to understand the behaviour of spiral waves in cardiac tissue are of great importance. Reviews of the theory and evidence for the role of spiral
waves in this setting are given in Refs. [Jal00, IR06].

In this system, the wave field consists of electrical potential differences across the membranes of muscle cells in the heart. These cells, called cardiomyocytes, contract and relax in response to changes in their potential difference, thereby enabling the organ to pump blood. This creates a coupling between the wave and the underlying mechanics of the medium in which it propagates. As the wave travels through the tissue, it causes local contraction wherever it passes. This is followed by mechanical relaxation after the wave has passed. Additionally, the dynamics of the wave are contingent on the geometry of the domain—as the medium deforms this affects the speed at which excited regions diffuse. This leads to a coupling from the mechanics back to the wave.

It is therefore interesting to study the interaction of spiral waves with a deforming body in which they propagate. In the biophysics literature, many models have been considered that attempt to capture the behaviour of cardiac tissue. These pair reaction-diffusion PDEs describing the dynamics of the electrical potential with nonlinear solid mechanics governing the tissue deformation. Serious efforts to model the heart attempt to replicate the biophysical complexity of a living organ to some extent. A review of current methods can be found in Ref. [Tra11]. State-of-the-art models involve detailed descriptions of the ionic currents at the cell level and anisotropic mechanics laws based on experimental results of tissue response. Cellular models can involve very many variables and represent a significant computational challenge [FC08]. Additionally, advances in medical imaging allow for anatomically realistic simulation domains depicting the geometry of a whole heart [VAP+10, GLC+11]. These can be integrated with fluid dynamics solvers to couple the system to blood flow simulations [NNN+11, NMK+11]. The primary purpose of this research is to better understand cardiac function (and dysfunction) across a range of physical scales, ultimately with a view towards clinical implications.

This will not be our aim. Our starting point is the contention that a spiral wave coupled to a deforming medium is an intrinsically interesting nonlinear system to study. Furthermore, we note that phenomenological approaches
have been instrumental in establishing much of our current understanding of spiral wave behaviour, for example in the study of initiation [PV91], shape and frequency selection [Kar91] and the problem of ‘meander’ [Bar95]. In this regard, there may be much to learn in the case of a generic spiral wave interacting with a deforming medium. Therefore, we pare down the description of cardiac tissue to its most basic components. We use the same generic excitable media PDEs as in the previous two chapters and couple these with a simple nonlinear elasticity model. The details will be explained later in Sec. 4.3.

A number of previous studies have proceeded in a similar vein. Most notable is the work of Nash and Panfilov [NP04], who conducted numerical simulations of the two-variable Aliev-Panfilov reaction-diffusion system coupled with isotropic nonlinear elasticity in two spatial dimensions. These were used to demonstrate the effect of deformation on the drift and rotation period. In later papers, an equation accounting for currents activated by mechanical stretching was added and in some cases the Aliev-Panfilov model was replaced with the three-variable Fenton-Karma system. This extended model was used to show that the existence of wave pulses with a steady ‘pacemaking’ rhythm can depend on whether or not the medium deforms [PKN05, KNP07] and to study drift and breakup of spiral waves [PKN07]. (The phenomenon of spiral wave ‘breakup’ shall be discussed later in Section 4.5.) Cherubini et al. used a similar model to simulate spiral waves propagating in slabs of active material undergoing large out-of-plane mechanical displacements due to external pressure [CFNT08]. An alternative approach for the mechanics modelling replaces the continuum description of elasticity with a discrete one based on a mass-spring system. Weise et al. showed that a system of this kind could reproduce many of the results of the continuum model [WNP11] and went on to study spiral wave initiation [WP11, WP12] and drift [WP13] in this setting. Finally, we observe that a handful of papers have considered the influence of prescribed deformations on spiral waves in simple excitable models. In particular, the effect of sinusoidal deformation on spiral wave breakup has received specific attention [ZRHO04, ZLS+06, CXYY09], while a recent paper by Chen et al. addresses various drift phenomena [CPZ+14].
Before we begin discussion of the details of our approach, it is important to note that we shall be considering waves in both two- and three-dimensional media. A spiral wave in three dimensions is more accurately called a ‘scroll wave’. Imagine extruding a two dimensional spiral in a third direction—the thickness of the cardiac tissue wall. Such an object looks like a roll of paper (hence ‘scroll’) or alternatively like many spiral waves stacked on top of each other. This system can behave differently to the two dimensional waves we have studied thus far, even in the absence of medium deformation. However, provided that the orientation of these stacked spirals remains roughly constant with respect to one another, i.e. the scroll does not twist too much, then the dynamics of the two- and three-dimensional cases are similar. This will almost always be the case if the thickness dimension of the medium is small compared with the characteristic length scales of the spiral. We implicitly assume this in the following and shall use the term ‘spiral wave’ for both two- and three-dimensional cases.

4.2 Nonlinear elasticity theory

Since we are already familiar with spiral waves in excitable media, let us discuss the part of this system that is new: elasticity. Both this chapter and the next consider systems in which it is necessary to use the nonlinear theory due to the presence of significant deformations. Therefore, we begin by developing nonlinear elasticity theory in a general setting, before adapting it for the cardiac problem. The theory presented here is well-established; fuller accounts may be found in the following books, which are themselves representative of a much larger body of literature on the subject [AF80, Gur81, Ogd84, Ant95, GS08].

Consider a material which adopts a shape $\Omega_0$, where $\Omega_0$ is a connected open subset of $\mathbb{R}^m$ and $m$ is any natural number (always 2 or 3 herein). This domain is called the reference configuration, for reasons that will become clear shortly. We can label any point in the reference configuration by its position in Euclidean space, $X = (X_1, \ldots, X_m) \in \mathbb{R}^m$. Now let us suppose
that $\Omega_0$ is subjected to a deformation over a particular time interval $[0, T]$. This can be described by a family of deformation maps, parametrised by the time $t$, $\chi_t : \Omega_0 \to \mathbb{R}^m$, which collectively form a motion of the body $\chi : \Omega_0 \times [0, T] \to \mathbb{R}^n$, where $\chi_t = \chi(\cdot, t)$. A diagram of the setup is shown in Fig. 4.1. Although it is not strictly necessary, one may assume, as suggested by

**Figure 4.1**: Diagram showing the deformation maps $\chi_t$ from the reference configuration $\Omega_0$ to the deformed configurations $\Omega_t$. The motion $\chi$ traces the paths of material points in space and determines the relationship between Lagrangian co-ordinates $X$ and Eulerian co-ordinates $x$.

our notation, that $\chi_0$ is the identity map and so $\Omega_0$ denotes the configuration of the deforming body at $t = 0$. At a particular time $t$ the image of the corresponding deformation map is called the deformed configuration and is denoted $\Omega_t$. This is the instantaneous geometry of the body undergoing deformation. We must require that each $\chi_t$ is injective so that no two material points may coalesce. Therefore every $x = (x_1, \ldots, x_m) \in \Omega_t$ is the image of a point $X$ in the reference configuration under its deformation map. Viewed in this way, the maps $\chi(X, \cdot)$ specify Lagrangian motions of the material points $X$, which is to say that they track the paths that points from $\Omega_0$ take in space as time elapses. Any material point at time $t$ can thus be referred
to by its so-called Lagrangian co-ordinate \( X \in \Omega_0 \) or its Eulerian co-ordinate \( x = \chi(X, t) \in \Omega_t \).

### 4.2.1 Strain

The motion of the body contains all the information necessary concerning its strain. This is a geometric property that measures the deformation of material elements relative to the reference configuration. There are many nonequivalent ways to define strain, which are useful in different circumstances. The most straightforward, is to take the spatial gradient of the motion with respect to the reference co-ordinates

\[
\mathbf{F} := \frac{\partial \chi}{\partial X}.
\]

To simplify notation, we have omitted both the time and space dependence in the above equation, which should be regarded as implicit in the following discussion. We refer to the tensor \( \mathbf{F} \) as the deformation gradient. It describes the local (linear) transformation of differential line elements, since \( dx = F \, dX \).

Likewise, the determinant of \( \mathbf{F} \), governs the transformation of areas and volumes. These transformation formulae are of particular importance in the theory because, as we shall see later, they describe the relationship between quantities in the Lagrangian and Eulerian frames.

We derive the formula for volumes first. Note that \( m \) linearly independent differential reference line elements \( dX^1, \ldots, dX^m \), define a parallelepiped whose volume \( dV \) is given by \( \det(dX^1, \ldots, dX^m) \). Therefore, the corresponding deformed volume element \( dv \) is \( \det(F \, dX^1, \ldots, F \, dX^m) = J \, dV \), where we have defined \( J := \det(F) \). If \( J = 1 \), then there is no local volume change. Bodies that satisfy this condition under any circumstances are referred to as incompressible. The physical interpretation of \( dv = J \, dV \) places two restrictions on the deformation gradient. Firstly, we must not allow any volume in the deformed body to shrink to zero. Therefore, \( J \neq 0 \), i.e. \( \mathbf{F} \) is nonsingular. Secondly, we regard as unphysical any motion which causes volume elements to change sign, since this would turn them inside-out. Hence we stipulate
that the deformation must be orientation-preserving, $J > 0$.

To derive the area transformation formula, let $dA$ be the reference area element with unit outward normal vector $\hat{N}$ and $da$ the corresponding deformed area with unit outward normal $\hat{n}$. Consider a differential line element $dX$ such that $dX \cdot \hat{N} \neq 0$. Then $dX \cdot \hat{N} dA$ is the volume of a cylinder with base $dA$ and axis $dX$, which deforms to $dx \cdot \hat{n} da$ under the motion. Therefore, by the rules derived above for transformation of lines and volumes:

$$J dX \cdot \hat{N} dA = dx \cdot \hat{n} da = F dX \cdot \hat{n} da. \quad (4.2)$$

Rearranging the leftmost and rightmost sides gives $dX \cdot (J \hat{N} dA) = dX \cdot (F^T \hat{n}) da$. Since this equation is true for any reference line element, we arrive at

$$\hat{n} da = JF^{-T} \hat{N} dA, \quad (4.3)$$

describing the transformation of areas. This is sometimes referred to as Nanson’s formula.

The ultimate motive behind quantifying the strain in a body lies in the fact that geometric deformations give rise to internal forces between material elements and vice versa. In this sense, $F$ is not a particularly good measure of strain, since it contains information about rigid movements which do not affect the forces inside the material. Specifically, consider any motion of the form

$$\chi(X, t) = Q(t)X + c(t), \quad (4.4)$$

where $Q(t)$ is an orthogonal matrix and $c(t)$ is a vector. This is a time-dependent rotation plus a translation and its deformation gradient is $F = Q$. Ideally, we would like to think of strain as being independent of such a transformation. For this reason, the Cauchy-Green strain tensor is defined as $C := F^T F$. It is easy to see that in the case of Eq. (4.4), $C$ is the identity tensor. If a more general motion is considered, the polar decomposition theorem may be used on the deformation gradient, yielding $F = RU$, where $R$ and $U$ are orthogonal and symmetric positive-definite tensors respectively. Therefore, since $C = (RU)^T RU = U^2$ in the general case, it does not depend on any
rigid motion of the material elements. The symmetric part $U$ is referred to as the stretch tensor for the following reason. Given differential line elements $dx$, $dX$, we can write $dx = \hat{n}||dx||$ and $dX = \hat{N}||dX||$, where $\hat{n}$ and $\hat{N}$ are unit vectors aligned with $dx$ and $dX$ respectively. Therefore, the transformation formula for differential line elements becomes $\hat{n}||dx|| = F\hat{N}||dX||$ and we see that

$$
||dx||^2 = (F\hat{N}||dX||) \cdot (F\hat{N}||dX||) = \hat{N} \cdot C\hat{N}||dX||^2, \quad (4.5)
$$

from which the ratio of the differential lengths may be computed

$$
\frac{||dx||}{||dX||} = \left[ (U\hat{N}) \cdot (U\hat{N}) \right]^{\frac{1}{2}} = ||U\hat{N}||. \quad (4.6)
$$

Hence, we see that the action of $U$ on a given direction $\hat{N}$ gives the change in length of a line element $dX$ aligned in that direction. Moreover, we may see that $C$ dictates the angles between deformed lines, since $dx \cdot dx' = dX \cdot C dX'$ for deformed elements $dx, dx'$ and reference elements $dX, dX'$. The Cauchy-Green strain thus provides a complete description of the local shape changes in the material.

### 4.2.2 Stress

We now turn our attention to the forces acting on the body. This will enable us to derive the equations of motion for an elastic medium. We shall begin our discussion by working with the Eulerian description of the body, since it is the more natural reference frame in which to understand the physics. Later on, a switch will be made to the Lagrangian frame, which is typically more convenient for numerical simulations since it is stationary. Consider a small chunk of material in the reference configuration, given by some connected subset $B_0$ of $\Omega_0$ and let $B_t$ denote the image of $B_0$ under $\chi_t$. As $B_t$ deforms under the motion of the body, its connectedness is preserved under the continuity of $\chi$. It has a mass density field $\rho$ which changes with time as the body expands, contracts and moves. Integration of the density field gives the mass of the chunk $\int_{B_t} \rho(x,t) dv$. By differentiating this expression
for mass, we can derive the associated conservation law. First, note that
\[
\frac{d}{dt} \int_{B_t} \rho(x,t) \, dv = 0, \tag{4.7}
\]
since \(B_t\) follows the chunk as it deforms and hence the total mass is constant. The time derivative can be brought inside the integration by applying the Reynolds transport theorem [Gur81, Sec. 10]. Let \(\nu(x,t)\) denote the velocity of the material at the point \(x\) at time \(t\). Then
\[
\frac{d}{dt} \int_{B_t} \rho \, dv = \int_{B_t} (\dot{\rho} + \rho \nabla \cdot \nu) \, dv = \int_{B_t} \left( \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \nu) \right) \, dv, \tag{4.8}
\]
where we have omitted dependence of the variables on space and time. Here and throughout we use a dot (\(\dot{}\)) as a shorthand denoting a total time derivative: in this case, \(\dot{\rho} := d\rho/dt\). At each material point \(x \in B_t\) this is the rate of change of density in the reference frame of an observer that follows the point. In the second equality this was decomposed into the rate of change at \(x\) in a stationary reference frame plus rate of change due to transport of the point, i.e. \(\dot{\rho} = \partial\rho/\partial t + \nu \cdot \nabla \rho\).

In Eq. (4.8) the domain of integration \(B_t\) is arbitrary, since it corresponds to an arbitrary choice of \(B_0\). Therefore, for any point \(x \in \Omega_t\):
\[
\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \nu) = 0. \tag{4.9}
\]
This is the conservation of mass equation for any continuous medium.

The chunk’s aggregate linear momentum may be computed by integrating \(\rho \nu\). Therefore, according to Newton’s second law, the total force \(f\) on the chunk is given by
\[
f(t) = \frac{d}{dt} \int_{B_t} \rho \nu \, dv = \int_{B_t} \left( \rho \dot{\nu} + \nu \left[ \frac{\partial \rho}{\partial t} + \nu \cdot \nabla \rho + \rho \nabla \cdot \nu \right] \right) \, dv = \int_{B_t} \rho \dot{\nu} \, dv, \tag{4.10}
\]
where the last equality follows due to conservation of mass.

There are two kinds of forces acting on \(B_t\). Surface tractions arise due to physical contact between the boundary \(\partial B_t\) and neighbouring bodies. These
may either be adjacent parts of the same medium $\Omega_t$, or (if $\partial B_t \cap \partial \Omega_t \neq \emptyset$) external bodies in contact with $\Omega_t$. Body forces act on material points in the interior of $B_t$. A typical example is gravity, which acts on the mass inside the body. Another is electric field, which acts on any charge present in the material. For each of these types of force, we can define force density vector fields. Let us define $\mathbf{r}$ to be the density field of surface tractions on $\partial B_t$ and $\mathbf{b}$ to be the body force density on $B_t$. Integrating these force densities gives an alternative expression for the total force on the chunk:

$$f(t) = \int_{B_t} \mathbf{b}(x,t) \, dv + \int_{\partial B_t} \mathbf{r}(x,t) \, da. \quad (4.11)$$

Let us consider the traction vector in more detail. The chunk $B_t$ was chosen to be arbitrary. At any point $x \in \Omega_t$, there are (infinitely) many such volumes whose boundaries touch $x$. Therefore, the point $x$ is subject to tractions from the contact between surfaces oriented in every possible direction. It is a fundamental assumption of continuum mechanics that the traction vector at $x$ associated with a surface $\mathbf{n}$, depends only on the (unit) normal vector $\hat{n}$ at $x$ and not on the curvature, or higher derivatives of $\hat{n}$. This assumption is usually referred to as Cauchy’s postulate, because it allowed Cauchy to prove that the contribution of all these tractions could be summarised by a linear map called the stress tensor. The relationship, in the Eulerian frame, is

$$\mathbf{r}(x,\hat{n},t) = \mathbf{T}(x,t)\hat{n}(x), \quad (4.12)$$

where $\mathbf{r}$ now depends on the normal vector $\hat{n}$ and is no longer tied to a specific choice of $B_t$. The tensor field $\mathbf{T}$ is called the Cauchy stress. Physically, it corresponds to a force per unit deformed area. Equation (4.12) says that in order to describe the tractions at $x$ completely, it is sufficient to know the traction vectors for $m$ surfaces oriented in linearly independent directions $\hat{n}_1, \ldots, \hat{n}_m$ at $x$. A traction exerted on a surface with normal $\hat{n}$ is simply a linear combination of the tractions corresponding to the surface normals $\hat{n}_1, \ldots, \hat{n}_m$. By way of example, let us consider the case where $m = 3$ and the basis of our co-ordinate systems is $\{e_1, e_2, e_3\}$. The components of $\mathbf{T}$ are
given by its action on the basis: $T e_j = (T_{1j} \ T_{2j} \ T_{3j})^T$. That is, $T_{ij}$ is the $i$-th component of stress exerted on a surface oriented in the direction $e_j$. (Some authors prefer a convention in which $i$ and $j$ in our definition are reversed.) A visualisation of the stress components is shown in Fig. 4.2. The three diagonal components of $T$ act out-of-plane and are called normal stresses, while the six in-plane off-diagonal components are called shear stresses.

### Figure 4.2

One way to visualise the components of the stress tensor at a point $x$ in three dimensions is to use the faces of a cube (or more generally, a parallelepiped, if the basis vectors $e_1, e_2, e_3$ are not mutually orthogonal). The cube should be thought of as infinitesimal, centred at the point $x$ and oriented such that its faces lie normal to the basis directions. Then the components of the stress tensor give the values of the tractions on the faces, as indicated in the diagram.

#### 4.2.3 Governing equations

The stress tensor and the body force density together fully describe the forces acting on any continuous medium. We are now in a position to derive the equations of motion for the body. Returning to the volume $B_t$, we can equate the different expressions for the total force on it given in Eqs. (4.10) and (4.11):

$$
\int_{B_t} \rho(x,t) \dot{\gamma}(x,t) \, dv = \int_{B_t} b(x,t) \, dv + \int_{\partial B_t} T(x,t) \hat{n}(x) \, da \quad (4.13)
$$
where we have used Eq. (4.12) to replace the traction vector with the stress tensor. Application of the divergence theorem yields:

\[
\int_{B_t} \rho(x,t) \dot{\nu}(x,t) \, du = \int_{B_t} \left[ \nabla \cdot T(x,t) + b(x,t) \right] \, du. \tag{4.14}
\]

Assuming all the physical variables in this equation are continuous throughout \( \Omega_0 \), then since \( B_0 \) is arbitrary, we can pass to the differential form

\[
\rho(x,t) \dot{\nu}(x,t) = \nabla \cdot T(x,t) + b(x,t). \tag{4.15}
\]

This is the *elastodynamics equation* for a continuous medium. An important special case is the *elastostatics equation*:

\[
\nabla \cdot T(x,t) + b(x,t) = 0, \tag{4.16}
\]

which determines the equilibria for a given body. Equations (4.15) and (4.16) relate the change in linear momentum to the resultant force at each point in the reference domain. In other words, they express the principle of conservation of linear momentum for the system.

Conservation of angular momentum places additional constraints on the body. In the case where the body is in equilibrium and \( b = 0 \), the net torque from surface tractions around each point \( x \) must be zero. By referring to Fig. 4.2, one can see that in order for torques to balance, it must be the case that

\[
T = T^T. \tag{4.17}
\]

This relation is the angular momentum conservation law for continuous materials. It can be proven in the general case by equating the net torque due to surface tractions and body forces, with the rate of change of angular momentum, just as we did in the case of linear momentum.

For our purposes, it will be convenient to formulate problems using the Lagrangian description of the body, since it allows numerical methods to take place on the fixed reference domain \( \Omega_0 \). In particular, we would like to write the equations of motion Eqs. (4.15) to (4.17) using the material co-ordinates
\( X \) in place of the spatial co-ordinates \( x \). Firstly, note that the density field of the reference domain, is constant in time. Let us call it \( \rho_0 \). For our chunk \( B_t \), the total mass is the same in the reference and deformed configurations:

\[
\int_{B_0} \rho_0(X) \, dV = \int_{B_t} \rho(x, t) \, dv = \int_{B_0} \rho(\chi(X, t), t) J(X, t) \, dV. \tag{4.18}
\]

The second equality follows by making a co-ordinate change. Therefore, (since \( B_t \) is arbitrary) the two density fields are related by \( \rho_0(X) = \rho(\chi(X, t), t) J(X, t) \).

For the same reason, the body force density field in the reference frame, \( b_0 \), is related to its counterpart in the deformed configuration by \( b_0(X, t) = b(\chi(X, t), t) J(X, t) \). The velocity on the other hand, does not depend explicitly on the deformed geometry—it is simply \( \dot{\chi}(X, t) \). To transform the Cauchy stress tensor, we can use Eq. (4.3) to yield:

\[
\int_{\partial B_0} T(x, t) \hat{n}(x) \, da = \int_{\partial B_0} J(X, t) T(X, t) F^{-T}(X, t) \hat{N}(X) \, dA. \tag{4.19}
\]

This motivates the definition of the so-called (first) Piola-Kirchhoff stress tensor field \( S := JTF^{-T} \), which describes surface forces on the body per unit undeformed area.

We may now write Equation (4.13) in the Lagrangian frame, as

\[
\int_{B_0} \rho_0(X) \ddot{\chi}(X, t) \, dV = \int_{B_0} b_0(X, t) \, dV + \int_{\partial B_0} S(X, t) \hat{N}(X) \, dA. \tag{4.20}
\]

After once again applying the divergence theorem and using the arbitrariness of \( B_0 \) to remove the integrals:

\[
\rho_0(X) \ddot{\chi}(X, t) = \nabla \cdot S(X, t) + b_0(X, t). \tag{4.21}
\]

This is the elastodynamics equation using the Lagrangian description of the body. The angular momentum conservation relation in this frame is given by

\[
SF^T = FS^T. \tag{4.22}
\]
Although we have suggestively referred to Equations (4.15) and (4.21) as the elastodynamics equations, our modelling assumptions up to this point have been somewhat modest. In fact, Eqs. (4.15) and (4.17) in the Eulerian frame and Eqs. (4.21) and (4.22) in the Lagrangian frame describe the motion of a wide range of continuous media. They are specialised to the problem at hand by specifying appropriate boundary conditions and constitutive laws. Boundary conditions place constraints on the body that are dictated by its interaction with the outer environment; constitutive laws define how the material responds mechanically to forces. We shall cover boundary conditions first.

4.2.4 Boundary conditions

There are two main kinds of boundary condition that are useful for elastic media. We shall describe them in the Lagrangian frame. Displacement boundary conditions specify the deformed positions of material points. Concretely, a function \( \chi = g \) on \( \partial \Omega_0 \times [0, T] \). In the broader context of PDE theory, this is called a Dirichlet condition. Note that it also dictates the velocity and acceleration at the boundary in the dynamic case. Traction boundary conditions specify the stresses acting at the boundary surface. In this case, a function \( h \) sets the traction via \( S \hat{N} = h \) on \( \partial \Omega_0 \times [0, T] \). Physically, this corresponds to external force (densities) impinging on the surface of the body. A special case is the traction-free condition \( S \hat{N} = 0 \), where there is no applied boundary force. These two types of boundary condition may be applied to Eq. (4.21) independently of one another, on disjoint subsets of \( \partial \Omega_0 \). They may also be applied separately to different co-ordinates of the boundary points. For example, consider an elastic cylinder surrounded by a rigid tube. The material points touching the edges of the tube cannot move radially—i.e. they are constrained to have zero (radial) displacement. However, the motion of the boundary points in the angular and azimuthal directions may not be known. In this case, traction components corresponding to surface shears in these directions must be prescribed.
4.2.5 Constitutive laws

Now we turn to discussion of constitutive laws. In order to solve Eqs. (4.21) and (4.22), we must first specify the way that the body strains under a particular stress field. This turns the generic equations for continuous media into a specific description of the material at hand. To solve the PDE system in three spatial dimensions, one must find three components of deformation and nine components of the Piola-Kirchhoff stress tensor, given six independent equations. Setting a constitutive law establishes the additional constraints required to close the system.

A straightforward way to do this is to write the Cauchy stress tensor as a some function \( g \) of the deformation gradient: \( T = g(F, X) \). The choice of \( g \) must lead to a symmetric \( T \) and it must determine the remaining independent components of stress. It must also be independent of the observer reference frame. A body that can be described in this way is called a Cauchy elastic material. The functional form in this case implies that the material stresses do not depend on the history of the deformation. Furthermore, the explicit dependence on \( X \) admits the possibility of heterogeneities, though we shall not consider such materials herein. More generally, constitutive laws may depend on any property of the physical system, e.g. higher derivatives of \( \chi \), temperature, electric field, time.

Our modelling work in the current and following chapters, assumes that the materials we are interested in are hyperelastic. (Some authors prefer to call them Green elastic.) This condition requires that the elastic potential energy stored in the material at each point depends on \( F \) only. This reasonable assumption is used widely in modelling elastic materials. Indeed, it can be argued from a thermodynamic perspective that every elastic solid should be considered hyperelastic (see e.g. [Gur81, Sec. 28] or [GS08, Sec. 7.2]). To see how hyperelasticity affects the stress tensor, let us take the scalar product of Eq. (4.15) with \( \nu \) to get an equation in units of power (density):

\[
\rho \nu \cdot \dot{\nu} = (\nabla \cdot T) \cdot \nu + b \cdot \nu, \tag{4.23}
\]
where once again, we omit the explicit space and time dependence of each field. By integrating over $B_t$ and applying the vector identity $\nabla \cdot (T\nu) = (\nabla \cdot T) \cdot \nu + T : \nabla \nu$, we obtain

\[
\int_{B_t} \rho \nu \cdot \dot{\nu} \, dv + \int_{B_t} T : \nabla \nu \, dv = \int_{B_t} \nabla \cdot (T\nu) \, dv + \int_{B_t} b \cdot \nu \, dv, \tag{4.24}
\]

where we use a colon to denote a Frobenius inner product between tensors.

This equation can be rewritten to show the power balance for the system:

\[
\frac{d}{dt} \int_{B_t} \frac{1}{2} \rho \nu \cdot \nu \, dv + \int_{B_t} T : \nabla \nu \, dv = \int_{\partial B_t} \tau \cdot \nu \, da + \int_{B_t} b \cdot \nu \, dv, \tag{4.25}
\]

where the first integral on the left-hand side was obtained using the Reynolds transport theorem and conservation of mass, as in Eq. (4.10). Physically, it is the rate of change of kinetic energy of the volume $B_t$. The two terms on the right-hand side give the work done per unit time by the traction and body forces. Therefore, the remaining term is the rate at which work is done by the stresses in the interior of $B_t$. This is referred to as the stress power. An analogous derivation may be performed in the Lagrangian frame, to obtain:

\[
\frac{d}{dt} \int_{B_0} \frac{1}{2} \rho_0 \dot{\chi} \cdot \dot{\chi} \, dV + \int_{B_0} S : \dot{\mathbf{F}} \, dV = \int_{\partial B_0} \tau_0 \cdot \dot{\chi} \, dA + \int_{B_0} b_0 \cdot \dot{\chi} \, dV. \tag{4.26}
\]

Let $W$ be the density field of the stored elastic potential energy in the material, written in the Lagrangian frame. It is commonly referred to as the strain energy density. Since an elastic medium does not dissipate energy, Equation (4.26) represents an energy conservation law and it must be the case that

\[
\frac{d}{dt} \int_{B_0} W \, dV = \int_{B_0} S : \dot{\mathbf{F}} \, dV, \tag{4.27}
\]

for any volume $B_0$. Hence, $\dot{W} = S : \dot{\mathbf{F}}$. The hyperelasticity assumption is that $W$ (at each point $X$) is a function of $\mathbf{F}$ only. Therefore, the chain rule implies

\[\text{For two arbitrary second-order tensors } A \text{ and } B, \text{ this is defined to be } A : B := \text{tr} \left( AB^T \right).\]
that
\[
\frac{d}{dt} W(F,X) = \frac{\partial W}{\partial F} : \dot{F} = S : \dot{F}.
\]  
(4.28)

Since \( W \) depends \textit{only} on the deformation gradient at each point, it is independent of \( \dot{F} \), as is the stress tensor \( S \). One can construct a motion with arbitrary \( F = F_0 \) and \( \dot{F} = G_0 \) using \( \chi(X,t) = (F_0 + G_0 t)X \). Hence, Eq. (4.28) must hold for any tensor \( \dot{F} \) and therefore

\[
S = \frac{\partial W}{\partial F}.
\]  
(4.29)

This equation gives the relationship between stress and strain in terms of the strain energy density \( W \). Furthermore, it implies that stress (for a hyperelastic body) is conservative, in the sense that the work done by stresses between two states of deformation does not depend on the motion.

In light of Equation (4.29), hyperelastic models specify constitutive laws by giving the form of \( W(F,X) \), thereby implicitly providing the stress-strain dependence. As we shall see shortly, even fairly simple formulas for \( W \) yield complex nonlinearities when incorporated into the elastodynamics equation.

Material properties place constraints on the allowed forms of the strain energy density. Henceforth, we shall assume that the materials of interest to us are both homogeneous and isotropic. The relevance of these assumptions shall be discussed later. Homogeneity implies that the constitutive law does not depend explicitly on the material co-ordinate \( X \). Isotropy means that the stress response of the material is independent of the direction of strain. It can be shown, perhaps unsurprisingly, that \( W \) is an isotropic function in this case, i.e. \( W(QFQ^T) = W(F) \) for all orthogonal tensors \( Q \). Moreover, this constraint implies that \( W \) depends only on the \textit{principal invariants} of the Cauchy-Green strain \( C \) (see e.g. [GS08, Sec. 7.1.5]). These are scalars which arise as the coefficients of the characteristic polynomial of a matrix and are consequently invariant to change of basis (i.e. change of co-ordinate frame). In three dimensions, they are \( I_1(C) = \text{tr}(C), I_2(C) = \frac{1}{2}(\text{tr}(C)^2 - \text{tr}(C^2)) \) and \( I_3(C) = \det(C) \). In view of this, the simplest way to define a general
constitutive law that makes physical sense, is

\[ W(C) = c_1(I_1 - 3) + c_2(I_2 - 3) + d(I_3 - 1)^2, \tag{4.30} \]

for arbitrary parameters \(c_1, c_2\) and \(d\). (In two spatial dimensions, \(I_1\) and \(I_2\) are dependent and so only one of these terms is used.) This is called the Mooney-Rivlin law, after its originator Melvin Mooney [Moo40] and Ronald Rivlin, who worked on it later [Riv48]. The special case \(c_2 = 0\) is the Neo-Hookean law. The form of \(W\) in Eq. (4.30) sets the reference configuration to be the state of zero strain energy (since \(W(I) = 0\), which is consistent with the convention that \(\Omega_0\) is unstrained. The final term, which contains the determinant of \(C\), is squared because volumetric compression and dilation both increase the stored energy in an elastic material. Equation (4.30) therefore represents the lowest-order expansion of \(W\) that contains all three principal invariants. Its construction is purely phenomenological. For a given material the model parameters \(c_1, c_2\) and \(d\) may be determined experimentally by fitting against known stress and strain data (see e.g. [Tre58]). The Mooney-Rivlin law may be readily generalised by including arbitrarily many higher-order terms. This gives the most general form of the strain energy density function for a homogeneous and isotropic material. Nevertheless, many alternatively constructed constitutive laws exist in this case, since they may capture material response better using fewer model parameters. Examples include the Ogden model [Ogd72] and the Gent model [Gen96].

A nice way to reformulate constitutive laws that are based on the principal invariants is to use the so-called deviatoric invariants, which are the principal invariants of the tensor rescaled to have unit determinant: \(\tilde{C} = \det(C)^{1/m}C\). These are \(\tilde{I}_1 = I_1 I_3^{-1/m}\) and \(\tilde{I}_2 = I_2 I_3^{-2/m}\). They are independent of volumetric changes in the material. For example, consider the rescaling \(F \mapsto \alpha F\). Then

\[ \tilde{I}_1 \mapsto \text{tr}(\alpha^2 C) \det(\alpha^2 C)^{-1/m} = \tilde{I}_1(C). \tag{4.31} \]

Therefore, we may write the strain energy density function in terms of two parts: a deviatoric part responsible for local shape changes to material ele-
ments and a volumetric part that controls volume change:

$$W(C) = \bar{c}_1 (\bar{I}_1 - 3) + \bar{c}_2 (\bar{I}_2 - 3) + \bar{d} (J - 1)^2, \quad (4.32)$$

where we recall that $J = I_3^{1/2}$. Note that in general, the phenomenological constants in this formulation will differ from those in Eq. (4.30). The first two terms on the right-hand side are the deviatoric components. They are bounded below by zero\(^2\) and increase as the shape of volume elements deviates from the zero-strain state\(^3\). The final term is the volumetric part.

The above constitutive law is useful because a material’s resistance to volume change—also called its compressibility—is a property that may be independently studied by experimentalists. One way to characterise compressibility is via Poisson’s ratio. Given a cube of material compressed on its top and bottom faces, the Poisson ratio is the (negative of the) ratio of transverse to axial strain measured, in the limit of small strains. Highly compressible materials reduce their volume significantly under compressive forces. Therefore, they experience small amounts of strain in the directions orthogonal to the compression and have low Poisson ratios. At the other extreme, incompressible materials preserve their volume by straining half as much in the transverse direction and thus have Poisson ratio 0.5.

Since $J = 1$ everywhere in an incompressible material, we note that the Mooney-Rivlin law in the case contains no volumetric term. Such a system must be solved alongside the condition $\det(F) = 1$, leaving the problem overdetermined. Let us write this constraint as the zero of the scalar function $\phi(F) := \det(F) - 1 = 0$. Note that

$$\frac{d}{dt} \phi(F) = \frac{\partial \phi}{\partial F} : \dot{F} = 0. \quad (4.33)$$

Therefore, the addition of any scalar multiple of $\partial \phi / \partial F$ to the Piola-Kirchhoff

\(^2\)To see this, one can write the deviatoric invariants in terms of principal stretches $\lambda_1, \lambda_2$ and $\lambda_3$, which are the eigenvalues of the deformation gradient. Analysis of the resulting functions determines a family of minima along the line $\lambda_1 = \lambda_2 = \lambda_3$, for which $\bar{I}_1 = \bar{I}_2 = 3$.

\(^3\)For example, in the case of a simple shear defined by $\chi(X, Y, Z) = (X + KY, Y, Z)$, for some arbitrary constant $K$, it can be shown that $\bar{I}_1 = \bar{I}_2 = 3 + K^2$ [Riv48]. Therefore, the deviatoric components are unbounded above.
stress leaves the stress power unaffected. Hence, we may introduce the Lagrange multiplier \( p \) to the stress at each point, as so

\[
S = \frac{\partial W}{\partial F} - p \frac{\partial \phi}{\partial F} = \frac{\partial W}{\partial F} - p F^{-T},
\]

(4.34)

where we used the standard formula for the derivative of the determinant function. In the Eulerian frame, the Lagrange multiplier term is \( -p I \). For \( p > 0 \) this is a compressive stress that is uniform in all directions. That is, \( p \) is a pressure field that opposes volume change at each point. It must be solved for alongside the stress and deformation maps.

The Mooney-Rivlin law shall be our constitutive model of choice. Later in this chapter, we use its compressible form, while in the following chapter, the material of interest is incompressible. We end this section by deriving the stress-strain relation for a compressible Mooney-Rivlin solid explicitly. To do this, we recall Eq. (4.28) and differentiate the strain energy density with respect to \( F \). We perform the following derivation by writing the relevant tensors in component form and using the Einstein summation convention wherever convenient.

Firstly, for a general strain energy density \( W \), we have

\[
S_{ij} = \frac{\partial W}{\partial F_{ij}} = \frac{\partial W}{\partial C_{kl}} \frac{\partial C_{kl}}{\partial F_{ij}}
\]

(4.35)

and we see that

\[
\frac{\partial C_{kl}}{\partial F_{ij}} = \frac{\partial}{\partial F_{ij}} (F_{mk} F_{ml}) = F_{mk} \frac{\partial F_{ml}}{\partial F_{ij}} + F_{ml} \frac{\partial F_{mk}}{\partial F_{ij}} = F_{ik} \delta_{jl} + F_{il} \delta_{jk},
\]

(4.36)

so

\[
S_{ij} = \frac{\partial W}{\partial C_{kl}} F_{ik} \delta_{lj} + \frac{\partial W}{\partial C_{kl}} F_{il} \delta_{jk} = \frac{\partial W}{\partial C_{kj}} F_{ik} + \frac{\partial W}{\partial C_{jl}} F_{il} = 2 F_{ik} \frac{\partial W}{\partial C_{kj}},
\]

(4.37)

where we used the symmetry of \( C \) to collect the terms. Therefore,

\[
S = 2 F \frac{\partial W}{\partial C}.
\]

(4.38)
We would like to differentiate $W$ with respect to the principal invariants that are used to specify isotropic material laws. Therefore, we shall apply the chain rule to Eq. (4.38), for which we need the derivatives of the invariants themselves. These are standard formulae. We have used the derivative of the determinant twice already. For any tensor $A$ (not necessarily symmetric), the results are:

$$\frac{\partial I_1(A)}{\partial A} = I, \quad \frac{\partial I_2(A)}{\partial A} = I_1(A)I - A^T, \quad \frac{\partial I_3(A)}{\partial A} = I_3(A)A^{-T}. \quad (4.39)$$

Therefore,

$$S = 2F \left[ \frac{\partial W}{\partial I_1} I + \frac{\partial W}{\partial I_2} (l_1I - C) + \frac{\partial W}{\partial I_3} I_3 C^{-1} \right]. \quad (4.40)$$

Turning to the specific case of the Mooney-Rivlin model, given earlier in terms of deviatoric and volumetric components [Eq. (4.32)], we get

$$\frac{\partial W}{\partial I_1} = c_1I_1^{-\frac{1}{m}}, \quad \frac{\partial W}{\partial I_2} = c_2I_2^{\frac{1}{m}}$$

and

$$\frac{\partial W}{\partial I_3} = -\frac{1}{m}I_3^{-\frac{1}{m}-1}\left(c_1I_1 + c_2I_2I_3^{-\frac{1}{m}}\right) + d(1 - J^{-1}). \quad (4.42)$$

Equations (4.41) and (4.42) can be substituted into Equation (4.40) to give the Piola-Kirchhoff stress in terms of gradients of the deformation maps. Despite the simple form of the strain energy density, this dependence is very complicated. The consequent nonlinearities in $\nabla \cdot S$ are responsible for the complex behaviour of such elastic materials.

### 4.3 Modelling

The starting point for our simulations of spiral waves in elastic media is the two-variable Barkley model of a generic excitable medium. This has already been discussed in Secs. 2.2 and 3.3. In this chapter, the wave field $u = (u, v)$ may have two or three spatial dimensions, depending on the domain geometry. Furthermore, we pose the reaction-diffusion model equations [Eqs. (3.6a) and (3.6b)] in a deforming domain $\Omega_t$. In the Eulerian reference frame, spatial
derivatives depend on the geometry, thereby providing a coupling mechanism from mechanical deformation to the wave. The numerical implementation of this dependence is indicated in Section 4.4. If we were modelling a deforming chemical reaction, such as a so-called 'BZ gel' [YB06, KYB08], it would be necessary to include an advective term in the Eulerian reaction-diffusion system, reflecting the fact that chemical concentrations change when regions contract or stretch. In cardiac tissue however, the wave-field variables are identified with potential differences across cells which, in principle, should be independent of their deformation. Therefore, we do not include an advective term. A thorough study of the influence of advection is beyond our scope and so we simply note that its exclusion is an inherent assumption of our approach.

In modelling the medium deformation, our approach follows in many respects, the seminal Nash and Panfilov paper [NP04] that was discussed in Sec. 4.1. We use a compressible Mooney-Rivlin constitutive law throughout [Eq. (4.32)], which confers a homogeneous and isotropic material response. Rather than solve the full elastodynamics equations, we make the simplifying assumption that the mechanics equilibrate on a faster time scale than the speed of the spiral wave fronts. Therefore, the mechanics may be treated quasi-statically. We solve the elastostatics equation with no body force. The time-dependence of the deformation is dictated through coupling to the wave field. This coupling of the mechanics to the wave is achieved by by adding a component $T_{\text{act}}$ to the stress tensor, called the active stress in the literature:

$$T(F, u) = T_{\text{pass}}(F) + T_{\text{act}}(u).$$  \hspace{1cm} (4.43)

Here, $T_{\text{pass}}$ represents the usual (Cauchy) stress of a hyperelastic medium, which depends on the strain energy density function. This is called the passive stress. The active stress $T_{\text{act}}$ is an additional constitutive assumption that depends on the wave field $u$. Like body forces and boundary conditions, this is an input to the nonlinear elastic system.

There are many ways that one might want to model an active stress. The simplest way is to make it spherical at each point. A spherical (active) stress
field in the Eulerian frame is of the form \( T_{\text{act}}(u) := p_{\text{act}}(u)I \). This has the form of a negative pressure. When \( p_{\text{act}} > 0 \), its corresponding internal surface tractions pull inwards with equal magnitude in every direction. Taking the divergence, one sees that it is equivalent to a body force equal to the gradient of \( p_{\text{act}} \) that pushes towards regions where \( p_{\text{act}} \) is higher. Written in the Lagrangian frame, Eq. (4.43) gives the formula for the first Piola-Kirchhoff stress as:

\[
S(F, u) = S_{\text{pass}}(F) + p_{\text{act}}(u)J^{-T}. \tag{4.44}
\]

To produce a model that compresses the medium wherever there is excitation, it would be enough at this point to set \( p_{\text{act}}(u) = u \). However, in our motivating example of cardiac tissue contraction there is a time delay between the front of the wave field and the onset of muscle cell contraction. This is because contraction is associated with concentration of calcium ions, the uptake of which happens on a slower time scale than for the ionic species responsible for excitation. In principle, this delay may have a considerable effect on the dynamics of the spiral wave. To mimic this we follow Nash and Panfilov [NP04] by introducing an ordinary differential equation for \( p_{\text{act}} \) that relaxes to \( ku > 0 \):

\[
\partial_t p_{\text{act}} = A(ku - p_{\text{act}}), \tag{4.45}
\]

where \( k > 0 \) is a model parameter that sets the strength of the active stresses and \( A > 0 \) controls the speed of relaxation, thereby determining the delay between excitation and the generation of active stresses in the material. This ODE is solved alongside the reaction-diffusion and elastostatics systems.

The model presented here is a caricature that includes only the most fundamental characteristics of real cardiac tissue. Simplified approaches have value because they can reduce the complexity of systems, rendering them accessible to our understanding. Moreover, as we have argued above, they may themselves be independently interesting. Nevertheless, we shall outline some of the limitations of this model with respect to the cardiac problem. Perhaps the most important of these is the assumption of isotropy.
The microstructure of the heart is organised into discrete sheets of muscle fibres whose mechanical response depends strongly on the orientation of applied stresses [NH00, DSYL02]. A number of different constitutive laws have been proposed that take this anisotropy into account [HO09]. Fibre orientation also leads to a directional dependence in ionic conduction, which can be accounted for via the diffusion component of the reaction-diffusion system [Rot92]. Moreover, this anisotropy in conductivity differs inside and outside the cell. The so-called bidomain model that addresses this is a standard tool in the cardiac modelling community [PDR+06, CBC+11]. This may be paired with a wide array of different kinetics models that accurately capture the excitation-recovery processes in a cell [FC08]. The coupling of the wave and mechanics may also be treated more realistically. For example, Ref. [HMTK98] describes a model of active stress informed by experimental tests on tissues at various length scales. Finally, we note that the physical properties of real tissue are nonuniform in space to some extent. Any traveling wave in this medium is consequently a less regular process, as observed in laboratory experiments [DPS+92, PDS+93]. This cannot be reflected by a simple model of the kind described here.

4.4 Methods

We solve both the elastostatics equations and the reaction-diffusion system using the finite element method. Computations are performed using the DUNE software libraries [BBD+16]. In particular, extensive use is made of the DUNE-FEM module, which provides a framework for the implementation of finite element methods. As such, all the low-level code required of a typical finite element program (grid management, basis construction, quadrature, etc.) is delegated to the DUNE libraries. The reader may refer to the documentation in Refs. [DKNO10, DUN16]. The design philosophy of these libraries permits very generic construction of finite element schemes, allowing for different implementations of elements, grids, nonlinear solvers and so forth, to be used interchangeably. The code written for this chapter inherits this
philosophy. In particular, it works in both two and three spatial dimensions and for elements of many polynomial orders.

In the remainder of this section, we provide an overview of how the PDEs are solved in our context. For the mechanics, our starting point is the elastostatics equation in Lagrangian form, with no body force

\[ \nabla \cdot \mathbf{S} = 0. \tag{4.46} \]

We note that this is the only equation that we have to solve. The angular momentum condition [Eq. (4.22)] is automatically satisfied for any isotropic constitutive law. This can be seen by multiplying Eq. (4.38) on the left by \( \mathbf{F}^T \), from which, the symmetry of \( \mathbf{S} \) is immediately apparent.

Since we are in the static case, the solution to Eq. (4.46) is a single deformation map \( \mathbf{\chi} \), whose image is the deformed configuration at equilibrium.\(^4\) Let \( \mathbf{w} \) be an arbitrary deformation map. Taking the dot product of Eq. (4.46) with \( \mathbf{w} \) and integrating over the reference domain, we obtain

\[
\int_{\Omega_0} (\nabla \cdot \mathbf{S}) \cdot \mathbf{w} \, dV = \int_{\partial \Omega_0} \mathbf{S} \mathbf{n} \cdot \mathbf{w} \, dA - \int_{\Omega_0} \mathbf{S} : \nabla \mathbf{w} \, dV = 0 \tag{4.47}
\]

using integration by parts. The integral over the boundary of \( \Omega_0 \) is the work done by external tractions under the deformation \( \mathbf{w} \). Therefore, the solution \( \mathbf{\chi} \) satisfies

\[
\int_{\Omega_0} \mathbf{S}(\mathbf{\chi}) : \nabla \mathbf{w} \, dV - \int_{\partial \Omega_0} \mathbf{\tau}_0 \cdot \mathbf{w} \, dA = 0, \tag{4.48}
\]

for any deformation map \( \mathbf{w} \). This is the weak form of Eq. (4.46). To solve this equation, the reference domain is approximated by a discrete mesh \( \Omega_{0,h} \), on which a finite dimensional function space \( W_h \) can be defined that specifies a set of admissible candidate solutions. (In particular, \( W_h \) must be defined such that its elements satisfy any Dirichlet conditions placed on \( \partial \Omega_{0,h} \).) Here, \( h \) denotes the characteristic length scale of the discretisation. The discretised

\(^4\)This is a slight abuse of notation, since \( \mathbf{\chi} \) no longer refers to a motion, in the strict sense defined above in Section 4.2.
The numerical problem is therefore to find $\chi_h \in W_h$, such that
\[
\int_{\Omega_{0,h}} S(\chi_h) : \nabla w_h \, dV_h - \int_{\partial \Omega_{0,h}} \tau_0 \cdot w_h \, dA_h = 0,
\] (4.49)
for all $w_h \in W_h$. Let $N$ be the dimension of $W_h$ and $\{\xi_1, \ldots, \xi_N\}$ a basis. Equation (4.49) is linear with respect to $w_h$. Therefore, an equivalent problem is to find scalars $a_1, \ldots, a_N$ such that
\[
L_i(a_1, \ldots, a_N) := \int_{\Omega_{0,h}} S(a_j \xi_j) : \nabla \xi_i \, dV_h - \int_{\partial \Omega_{0,h}} \tau_0 \cdot \xi_i \, dA_h = 0,
\] (4.50)
for all $i \in \{1, \ldots, N\}$. The functions $L_i$ define $N$ equations with $N$ unknowns, on which a suitable nonlinear solver can be used. In our case, we use Newton’s method, for which we require the system Jacobian, which has components
\[
\frac{\partial L_i}{\partial a_j} = \int_{\Omega_{0,h}} \left( \frac{\partial S}{\partial F}(a_k \xi_k) \nabla \xi_j \right) : \nabla \xi_i \, dV_h.
\] (4.51)
The term $\partial S / \partial F$ is a rather complicated fourth-order tensor. A formula for it can be derived by adding the active stress to Eq. (4.40) and differentiating with respect to $F$. To avoid the need to apply the product rule directly on the large expression in this equation, it is convenient to define $P := F^{-1} S$. In the literature, this is called the second Piola-Kirchhoff stress tensor. Then, writing the relevant tensors in components and differentiating, gives
\[
\frac{\partial S_{ij}}{\partial F_{kl}} = \frac{\partial}{\partial F_{kl}} (F_{iq} P_{ij}) = \frac{\partial F_{iq}}{\partial F_{kl}} P_{ij} + F_{iq} \frac{\partial P_{ij}}{\partial C_{rs}} \frac{\partial C_{rs}}{\partial F_{kl}}
\] (4.52a)
\[
= \delta_{ik} P_{lj} + F_{iq} \frac{\partial P_{lj}}{\partial C_{rs}} (F_{kr} \delta_{ls} + F_{ks} \delta_{lr})
\] (4.52b)
\[
= \delta_{ik} P_{lj} + F_{iq} F_{kr} \left( \frac{\partial P_{lj}}{\partial C_{rl}} + \frac{\partial P_{lj}}{\partial C_{lr}} \right),
\] (4.52c)
where we used the result from Eq. (4.36) in the second line. Writing $\partial S / \partial F$ in terms of derivatives with respect to the Cauchy-Green strain anticipates...
that $P$ depends on $C$ only. We have

$$P = 2 \sum_{i=1}^{3} \frac{\partial W}{\partial l_i} \frac{\partial l_i}{\partial C} + p_{\text{act}} J C^{-1}$$

(4.53)

and therefore, after some work

$$\frac{1}{2} \frac{\partial P_{qj}}{\partial C_{rs}} = \sum_{k, l=1}^{3} \frac{\partial^2 W}{\partial l_k \partial l_l} \frac{\partial l_k}{\partial C_{qj}} \frac{\partial l_l}{\partial C_{rs}} + \frac{\partial W}{\partial l_2} (\delta_{qj} \delta_{rs} - \delta_{qj} \delta_{js})$$

$$+ I_3 \frac{\partial W}{\partial I_3} \left( C_{rs}^{-1} C_{rs}^{-1} - C_{jr}^{-1} C_{qs}^{-1} \right) + \frac{1}{2} p_{\text{act}} J \left( \frac{1}{2} C_{rs}^{-1} C_{rs}^{-1} - C_{jr}^{-1} C_{qs}^{-1} \right).$$

(4.54)

Finally, for completeness, we give the second derivatives of $W$. For the (compressible) Mooney-Rivlin model, only three of these derivatives are nonzero. These are

$$\frac{\partial^2 W}{\partial I_1 \partial I_3} = -\frac{1}{m} \bar{c}_1 I_3^{\frac{1}{m}-1}, \quad \frac{\partial^2 W}{\partial I_2 \partial I_3} = -\frac{2}{m} \bar{c}_2 I_3^{\frac{2}{m}-1}$$

(4.55)

and

$$\frac{\partial^2 W}{\partial I_3^2} = \frac{1}{m} \left( \frac{1}{m} + 1 \right) \left( \bar{c}_1 I_1 + \bar{c}_2 I_2 I_3^{\frac{1}{m}} \right) I_3^{\frac{1}{m}-2} + \frac{1}{m} \bar{c}_2 I_2 I_3^{\frac{2}{m}-2} + \frac{1}{2} \bar{d} J^{-3}. \quad (4.56)$$

Equation (4.52c), together with Eqs. (4.53) to (4.56), gives the full expression for $\partial S/\partial F$, needed for the linearisation of the weak form. This information, along with the expression for the weak form itself [Eq. (4.49)] and the formula for $S$ in Eqs. (4.40) to (4.42), is sufficient to implement compressible Mooney-Rivlin elasticity in a generic finite element library such as DUNE–FEM.

Both kinds of boundary condition mentioned in Section 4.2 may be used by the mechanics solver. These are implemented quite generally. The specific boundary conditions used to compute the results in the next section are as follows. The Dirichlet boundary fixes the edges of the medium by specifying zero displacement there. When traction boundary conditions are used, we
apply a traction that points along the outward normal to the domain boundary, with constant magnitude \( p_{\text{trac}} \) at every point \( X \in \partial \Omega_0 \). If no part of the boundary is Dirichlet, then the elastostatics equations possess infinitely many equivalent solutions related by translational and rotational symmetries. Therefore, for domains with two spatial dimensions, we fix three degrees of freedom on the boundary of the domain. In three spatial dimensions, we define one zero displacement Dirichlet point and two additional points—one with its displacement fixed in the \( z \)-direction and the other in the \( y \) and \( z \) directions. These two points remove solutions corresponding to the three rotational symmetries and are chosen such that the three points together are not collinear.

The Barkley model reaction-diffusion PDEs [Eqs. (3.6a) and (3.6b)] are implemented in a similar way to the mechanics, via numerical solution of their weak form with finite elements. Since the equations are time dependent, it is necessary to use a time stepping scheme. We use the explicit Euler method with time step \( \Delta t = 0.02 \). Aside from time-dependence, the main difference is that we solve the spiral wave equations in the Eulerian reference frame using a grid that is a discretisation of \( \Omega_t \). It is constructed by transforming the finite element mesh of \( \Omega_0 \) with the deformation map (which we obtain from the solution to the mechanics problem) at each time step. Manipulation of the mesh in this way is a feature available in DUNE that we make use of. The consequence for the finite element scheme is that the spatial derivatives in the reaction-diffusion PDEs take place in the deformed configuration and therefore they automatically observe the local geometry changes that provide the coupling from the mechanics to the wave.

An overview of the full simulation procedure is as follows. Time is discretised into \( N \) steps \( t_1, \ldots, t_N \), with increment \( \Delta t \). At each \( t_n \), we first solve the mechanics system

\[
\nabla \cdot \mathbf{S} = 0 \quad \text{on } \Omega_0, \\
\chi(X, t_n) = X \quad \text{on } \partial \Omega_0^{\text{disp}}, \\
\mathbf{S} \dot{\mathbf{N}} = p_{\text{trac}} \dot{\mathbf{N}} \quad \text{on } \partial \Omega_0^{\text{trac}},
\]

(4.57a) \quad (4.57b) \quad (4.57c)
where $\partial \Omega_0^{\text{disp}}$ and $\partial \Omega_0^{\text{trac}}$ are disjoint subsets of $\partial \Omega_0$ on which the displacement and traction boundary conditions are applied respectively. At $t_1$, the initial guess for the mechanics solver is the reference configuration. For each subsequent time step $t_n$, the (previously computed) deformed configuration at $t_{n-1}$ is used. Next, the Barkley model PDEs [Eqs. (3.6a) and (3.6b)] are time-stepped on the deformed geometry $\Omega_{t_n}$. The initial condition used for the reaction-diffusion system at $t_1$ is either a converged spiral wave from a previous simulation, or the cross-field stimulation

$$
\begin{align*}
  u(x, y, 0) &= \begin{cases} 
    1 & \text{if } x < L/2, \\
    0 & \text{otherwise,}
  \end{cases} \\
  v(x, y, 0) &= \begin{cases} 
    0 & \text{if } y < L/2, \\
    0.5 & \text{otherwise,}
  \end{cases}
\end{align*}
$$

(4.58)

which initiates a spiral wave. Here, $L$ is the length of the reference domain in the horizontal direction. The reaction-diffusion scheme is time-stepped $M$ times, with an increment of $\Delta t/M$, for some $M \in \mathbb{N}$. After each reaction-diffusion time step, the active stress ODE [Eq. (4.45)] is also time-stepped using the explicit Euler method, with the spiral wave field as input. The resulting active stress field is an input to the mechanics scheme at $t_{n+1}$.

We use $M = 3$ in simulations of two-dimensional media and $M = 2$ for three-dimensional media. In all situations, the nonlinear solver is Newton’s method, to which we apply damping in the case of three-dimensional domains. The linear problem is solved using the PETSc software library [BAA+14] (via DUNE) with the biconjugate gradient stabilised method and an additive Schwarz preconditioner for the linear problem.

### 4.5 Results

In this section, we provide some pictures that illustrate the sort of results that can be obtained with the implementation discussed above.

For all of the following results in two spatial dimensions, we set the Mooney-Rivlin parameters to $c_1 = 1$, $d = 1$. (In two dimensions the remaining mechanical parameter $\tilde{c}_2$ is redundant and is set to zero.) The active stress relaxation speed parameter $A$ [defined in Eq. (4.45)] is set to 1 throughout.
For the geometry of the reference domain we use a square with sides of 30 space units, meshed with $128 \times 128$ linear square finite elements. All other modelling parameters are given in the figure captions. As indicated in Section 4.4, for simulations involving traction conditions over the whole boundary, it is necessary to restrict some degrees of freedom to select a particular solution from the infinitely many that are available under the symmetries of the problem. In two spatial dimensions, we fix all co-ordinates of the origin and the second co-ordinate of the point $(1, 0)$. When simulating three dimensional media we fix the origin, the second and third co-ordinate of $(1, 0, 0)$ and the third co-ordinate of $(0, 1, 0)$.

Figure 4.3 shows a spiral deforming in a two-dimensional medium with a Dirichlet boundary condition. All the spiral waves in this section rotate anticlockwise. The deformation of material lines is shown on the figure. One clearly sees that areas near to the front of the wave are stretched. As the

![Figure 4.3: Snapshot of a spiral wave in a deforming domain with Dirichlet boundary conditions. Some representative material lines (black dashes) have been overlaid. These are spaced 16 elements apart. The active stress parameter is $k = 5$. Barkley model parameters $a = 0.8$, $b = 0.05$, $\varepsilon = 0.02$.](image-url)
wave passes a region, active stresses are generated causing area elements towards back of the wave to be compressed.

In Figure 4.4 we show multiple frames of a typical simulation to demonstrate the behaviour as time elapses. This time, traction boundary conditions are used. The spiral wave starts from an initial condition with a zero active stress field. As the simulation progresses, the medium contracts globally as the rising active stresses pull against the external traction. The sides of the domain warp, moving inwards wherever the active stresses are higher. Eventually, the global contraction of the medium ceases. The total active stress, which is proportional to the area that is excited at any given moment,
stays roughly constant. As the spiral wave rotates, the medium continues to deform locally.

Under certain conditions, the rigid rotation of a spiral wave may be disturbed, breaking the spiral into turbulent patterns [IG91, PH93, BE93]. A modification of the Barkley model commonly used to simulate this situation is due to Bär and Eisworth [BE93]:

\[
\frac{\partial u}{\partial t} = \nabla^2 u + \frac{1}{\varepsilon} u(1-u) \left( u - \frac{v + b}{a} \right), \tag{4.59a}
\]

\[
\frac{\partial v}{\partial t} = g(u) - v. \tag{4.59b}
\]

The standard Barkley model uses \( g(u) = u \), while Bär and Eisworth’s kinetics use a continuous, piecewise smooth function of the form

\[
g(u) = \begin{cases} 
0 & \text{if } u < \frac{1}{3}, \\
1 - 6.75u(u - 1)^2 & \text{if } \frac{1}{3} \leq u \leq 1, \\
1 & \text{if } u > 1.
\end{cases} \tag{4.60}
\]

Note that \( v \) is no longer activated below the threshold \( u = \frac{1}{3} \). As the excitation parameter \( \varepsilon \) increases, the recovery wave (\( v \)-field), which inhibits \( u \), shrinks and the rotation period of the spiral increases. For high enough values of \( \varepsilon \) (around 0.08 for typical model parameters) the system no longer supports rigid or meandering spiral waves, which break onto themselves creating turbulent patterns. The mechanism behind the loss of stability is explained in the original paper [BE93].

A simulation using the Bär-Eisworth kinetics is shown in Figure 4.5. The initial spiral wave grows rapidly before annihilating itself and breaking apart into patches. Disordered dynamics ensue. The proportion of active stress changes greatly and there is significant global deformation.

Finally, in Figure 4.6 we show a simulation of a Barkley model spiral in three spatial dimensions. The reference geometry is a cuboid with dimensions \( 30 \times 30 \times 2 \) space units, meshed with 60 linear cuboid elements in each lateral direction and 2 elements in the thickness direction. The Mooney-
Figure 4.5: Frames of a spiral wave breaking up in a deforming medium with traction boundary conditions. Some representative material lines (black dashes) have been overlaid. These are spaced 32 elements apart. The mechanics parameters are $k = 4$ and $p_{\text{trac}} = 5$. Breakup is induced by using Bär-Eisworth reaction kinetics with model parameters $a = 0.84$, $b = 0.07$ and $c = 0.08$. The first six frames are spaced at equal intervals of 1 time unit apart and show the spiral wave breaking apart. The later frames are spaced 2.9 time units apart and are representative of the resulting disordered dynamics.
Figure 4.6: Snapshot of a spiral wave propagating in a three-dimensional slab with traction boundary conditions. Some representative material lines (black dashes) have been overlaid. These are spaced 10 elements apart. The mechanics parameters are $c_1 = 1$, $c_2 = 2$, $d = 20$, $A = 1$, $k = 25$, $p_{trac} = 4$ and the Barkley model parameters are $a = 0.8$, $b = 0.05$ and $\epsilon = 0.02$.

Rivlin parameters used in this case are $\bar{c}_1 = 1$, $\bar{c}_2 = 2$, $\bar{d} = 20$. All other model parameters are given in the figure caption. In addition to in-plane compression and stretching of area elements, out-of-plane deformation is possible in the three-dimensional case and one sees that regions towards the back of the spiral wave are depressed.

4.6 Discussion

In this chapter, we have presented a phenomenological model of a spiral wave coupled to a deforming elastic medium. Our model can be viewed as one of the simplest possible nonlinear depictions of excitable wave propagation in cardiac tissue and follows in the footsteps of Nash and Panfilov who described a similar approach [NP04].

Nevertheless, the underlying theory governing general mechanical deformations is quite involved and may be unfamiliar to nonspecialists. To this end, we detailed the essentials of continuum mechanics theory and explained its application to our particular context. In particular, we showed how the equations of motion for a continuous deforming medium are constructed and discussed the boundary conditions and constitutive laws needed to close the system. For the case of a Mooney-Rivlin law—one of the simplest possible phenomenological laws—we derived the form of the stress tensor and ob-
served that even in this case, the dependence on the deformation gradient is nonlinear and highly complicated.

We subsequently explained how to combine a Mooney-Rivlin material with a generic excitable medium via an intermediate coupling variable known as the ‘active stress’ in the literature. Implementing a numerical method for solution of the resulting system is nontrivial. Despite this, discussion of how to do so must typically be omitted from research papers for the sake of brevity. Nash and Panfilov’s original study [NP04] contained some details on solution of their model using finite differences. In Sec. 4.4, we covered solution via the finite element method and gave some useful expressions that are not normally written explicitly in the literature, including a derivation of the system Jacobian. Together with Secs. 4.2 and 4.3, this constitutes a relatively full account of the theory and implementation of coupled reaction-diffusion-mechanics systems and may act as useful reference for future modelling efforts.

Our particular implementation, which uses the DUNE software libraries, exhibits some advantages over the finite difference approaches predominantly used in the prior studies of similar models. The finite element scheme may be solved on almost any sensibly constructed mesh. Consequently, different reference geometries may be considered in future work, including those that, like cardiac tissue, possess intrinsic curvature. Furthermore, our implementation is highly generic, in the sense that we may pass easily between simulating two and three spatial dimensions, using different linear and nonlinear solvers and elements of higher polynomial order. These features are in a large part due to DUNE, which was designed with this generic philosophy in mind.

Finally, to conclude this chapter, we presented some example results using our implementation in both two and three spatial dimensions, with different boundary conditions. An interesting example of spiral breakup with Bär-Eisworth kinetics was also presented. These examples give a good flavour of the sort of results that may be gathered using our approach. Detailed study of the dynamics and behaviour of the system is left as a topic for future research.
Chapter 5

Shape selection of dielectric elastomers

5.1 Introduction

Dielectric elastomers (DEs) are soft, flexible materials that deform mechanically when subjected to forces from electric charges. A flurry of interest in these systems was generated when Pelrine et al. [PKPJ00] demonstrated that very high strains could be obtained if the DE is stretched before use—a procedure known in the literature as ‘prestretching’. When actuated, DE films may strain to double their original area and in some cases even further [HLF+12, KLB+12]. Moreover, they may be designed to perform useful out-of-plane deformations [KPB06, KWPB07, RNDS09, AGS+15]. The impressive performance of these devices has led to many potential applications across the fields of engineering [PSLK+01, OOM08, BP10], medicine [GFM04, BC08, PBC+13] and even art [MGK12].

Typical setups in experiments and applications involve a thin elastomer film coated on opposite faces with areas of conducting material onto which charge can be deposited. We shall refer to these conducting areas as ‘active regions’. They are fabricated in such a way that they are free to bend and stretch with the elastomer and do not constrain its movement. These two conductors are connected to a circuit which produces a potential difference
between them, turning them into oppositely charged electrodes. This creates a capacitor in which the intervening dielectric (the elastomer) is apt to deform under the influence of electrostatic forces between the charge distributions.

Figure 5.1 contains some diagrams of this situation. Part (a) shows an example

![Figure 5.1](image)

**Figure 5.1:** Diagrams of a dielectric elastomer in different situations. In parts (a)–(c) the top electrode is shaded in grey. (a) DE with no applied electric field ($E = 0$). (b) DE with a strong applied electric field (large $|E|$), which causes the elastomer to deform. If the boundaries of the medium are free to move, the material compresses in the thickness direction and extends laterally. (c) If the edges of the medium are instead held fixed, the electrostatic forces between the electrodes force the elastomer to buckle out-of-plane. (d) Two dimensional cross-section through the middle of the elastomer showing the fringing of the electric field $E$. The top and bottom electrodes are represented by thick grey lines. Dashed lines with arrows indicate the direction of the electric field. At the centre the field lines are uniformly spaced and normal to the electrodes. At the edges they warp, leaving a small nonzero component of electric field tangent to the medium surface.

DE geometry in its zero strain configuration, before any electric field, $E$, has been applied. The medium is a thin cuboid, on which the top electrode can be seen, shaded in grey. Typical materials used in applications are isotropic and incompressible. They produce significant strains in response to high voltages on the order of kilovolts. Part (b) shows the DE in a simple situation in which the lateral sides of the medium are unconstrained. When the voltage is turned on, attractive forces arising from the charge imbalance on the two electrodes push the top and bottom faces of the elastomer together. This
compression is coupled to lateral stretching of the film via incompressibility, causing the area of the top and bottom faces to expand. If instead, the edges of the elastomer are held fixed in space, the active region and surrounding area will buckle out-of-plane as shown in Fig. 5.1(c). This is the only way that the incompressible material can preserve its volume under the compression of the electrodes. Such out-of-plane deformations are important for many DE applications in engineering such as pumps [PSLK+01], loudspeakers [HKEP06] and refreshable tactile displays [CBDR10]. The equilibrium shape adopted by a deformed elastomer is frequently nontrivial and can contain waves or wrinkles [PKPJ00, PD06, KKAB08, KZSK12].

The aim of this chapter will be to numerically model buckled DE shapes and make direct comparisons with experimental deformation profiles and images. We propose a straightforward approach to DE modelling that highlights some of the underlying physics at play and yields predictive results.

Before going into detail, we give a brief preview of the main result. Figure 5.1(d) shows a simplified diagram of the electric field between the electrodes, without any deformation indicated. The picture is the same as one sees in physics textbooks for the classical parallel plate capacitor. At the centre of the active regions, the electric field lines are uniform, parallel and extend perpendicularly from one electrode to the other. This produces a constant pressure, normal to the top and bottom surfaces. It is the primary effect responsible for actuation of the elastomer film and will be discussed in greater detail in Sec. 5.3. At the boundaries of the active regions, there is a fringing of the electric field and the field lines are slightly curved. This gives rise to forces at the electrode boundaries that are tangent to the top and bottom surfaces. We refer to these as ‘tangential tractions’. Although this effect is very small compared with the normal pressure, we find that it plays a significant role in shape selection. Indeed, we claim that consideration of both these normal and tangential forces, which are displayed in Fig. 5.1(b), is necessary and sufficient to effectively capture the shapes of the DEs considered herein.

Figure 5.2 shows an example of the importance of the tangential force. It shows points of vertical deflection from a DE experiment plotted alongside corresponding profiles from our numerical simulations. The solid yellow line
Figure 5.2: An example of the effect of tangent forces. The vertical axis has been scaled by a factor of 3, relative to the horizontal axis, to make the deformation profiles clearer. Red crosses are experimentally measured points of vertical deflection for an DE that has buckled out-of-plane. The experimental geometry used was a thin disc, with diameter 100 mm and thickness 0.15 mm. The electrodes were circular with diameter 40 mm and were subject to a potential difference of 3 kV. Lines display profiles from simulations that match the experimental geometry. The solid yellow line ($\kappa = 0.6$) is from a simulation that accounts for the tangent forces, while the dashed green line ($\kappa = 0$) is from a simulation with only normal pressure included. Details concerning the experimental setup, model simulations and the definition of $\kappa$ are given later in the text.

is from a simulation which takes into account tangential forces and the dashed green line is from a simulation which only models the normal pressure on the electrodes. Saving the details of the experiment and model for later in the text, we immediately see from this figure the dependence of the elastomer shape on the presence of the tangent force. The small amount of tension at the electrode edges has the effect of stretching and smoothing the deflection profile slightly. In particular, it causes the active region to flatten as we shall see below. We find that this shape cannot be replicated by considering normal pressure alone. Moreover, the indispensability of the tangent force is a feature common to all the results presented in Sec. 5.5.

5.2 Experiment

The experiments are conducted by our colleagues at laboratoire de Physique et Mécanique des Milieux Hétérogènes (PMMH) ESPCI, Paris. They use a circular disc of elastomer film which is held in a frame, keeping the edges
fixed in place. There is no prestretching of the elastomer before attaching it to the frame. The thickness of the film is typically around 100 to 200 µm and its diameter is 10 cm. A top-down photograph of the setup is shown in Figure 5.3. The electrodes in this case are circular. They are made from a thin layer of material called ‘carbon black’ which is painted directly onto the elastomer in the desired shape. Extending from the edges of the active region, one can see thin strips connected to crocodile clips, which join the electrodes to a voltage source. Underneath the elastomer, an optional air cushion may be applied which is adjusted so that the DE lies flat when no voltage is present, thereby counterbalancing the effect of gravity. We typically include the effect of gravity in our model. In particular, all our comparisons with experimental data use the case in which this air cushion is not present.

Figure 5.4 shows the experiment in use. In the centre, the circular electrodes are just visible against the unlit background. A red laser is directed across the surface of the film. Its path follows the deformation of the material and enables the experimentalists to measure the deflection along the diameter of the system. Whenever we fit our model results to the experiment, we use data points that correspond to these measurements of the laser trajectories.

Figure 5.3: Photograph of the experiment from above, courtesy of Hadrien Bense.
5.3 Modelling

Like any elastic object, when in equilibrium, elastomers obey the elastostatics equation introduced in the previous chapter (see Sec. 4.2):

$$\nabla \cdot T(x) + b(x) = 0.$$  \hfill (5.1)

For completeness, we recap the relevant details. The equation is posed over a domain $\Omega$ comprising of all the material points $x$ in the deformed body. Here, $T$ is the standard Cauchy stress tensor field, which dictates internal surface forces arising from material strains. Additional forces such as gravity are specified via the body force density field $b$. By defining an appropriate constitutive law relating stress to strain, Eq. (5.1) can be solved for the deformation of the elastic body, whose equilibrium shape depends on the boundary conditions. For each point on the domain boundary, one may prescribe either its position, or an external traction $\tau$ impinging on the material’s surface. The former case refers to a standard Dirichlet condition. In the latter case, $\tau$ dictates the component of the stress acting on the surface with outward unit
normal $\hat{n}$. This is given by $\tau = T\hat{n}$. It is also often desirable to define hybrid conditions where the boundary point is fixed in some directions and free to move (subject to prescribed stresses) in others. Finally, while material stresses arise physically between deforming material regions, we note that in practice the governing equations are solved in a fixed, unstrained, reference domain $\hat{\Omega}$ using appropriate transformation formulae. For details, see Chapter 4, especially Sects. 4.2 and 4.4.

In the case of DEs, the elastic stresses that we are primarily concerned with originate due to the presence of electric charges on the top and bottom active regions. The oppositely charged electrodes attract one another, giving rise to a compressive pressure. Towards the edges of the electrodes, the force on the charge distribution has a nonzero component that is tangent to the surface of the elastomer film. This arises because the repulsive forces between like charges there are not balanced, as they are in the centre. The effect is slight, compared with the normal pressure. Nevertheless, it causes the compliant electrodes to stretch and pull on the material to a certain extent, giving rise to additional surface traction.

The full picture of DE physics is complicated, since it must describe the coupling between the electric field, surface charges and the mechanical deformation in a general setting. Moreover, complex interactions arising from polarisation of the dielectric and strain-dependent permittivity should be accounted for in a comprehensive treatment [ZHS07, ZS08a]. Papers that formulate general theories of deformable dielectrics cover these issues in depth [Tou63, DO05, ML05, Suo10] and it is not our intention to consider such details here. Rather, our approach will be to try to treat DEs as simply as possible in order to understand some of the basic mechanisms governing the selection of the shapes that they adopt after buckling.

### 5.3.1 Effective pressure

A surprising fact about dielectric elastomers is that the pressure acting on each electrode due to the electric field is exactly twice that of a ‘classical’ capacitor with rigid plates and the same strength electric field. This subtlety was
understood by Pelrine et al. [PKJ98] who derived an expression for what they termed the ‘effective actuation pressure’—‘effective’ because the additional pressure arises indirectly due to the coupling of transverse and longitudinal deformation in an incompressible material. Their expression for the pressure (which we derive below) is used throughout the literature (e.g. [SLKS+02, WM05b, YYYR+05, KKL+11, GH15]) and has been verified for commonly used DE materials, in experiments [PKK00, KSLKP03, KSL05, SSA+16] and model simulations [WM05a].

Consider a DE of permittivity $\varepsilon$, whose electrodes are held at potential difference $V$, that deforms to thickness $D$. The magnitude of the electric field between the electrodes is $E = V/D$ and the formula for the effective pressure $p_{\text{eff}}$ is

$$p_{\text{eff}} = -\varepsilon (V/D)^2 = -\varepsilon E^2,$$

(5.2)

where the minus sign denotes that the pressure is compressive. In Appendix 5.A we provide a derivation in the case of a dielectric with general Poisson ratio $\nu$ that highlights the difference between Eq. (5.2) and the classical rigid capacitor. For now, it is sufficient to note the dependence on the physical observables, in particular the dependence on $1/D^2$, which is important later on.

### 5.3.2 Our model

Two sensible assumptions to make in a DE model are that the electric field between the electrodes is constant and that its effect on the elastomer can be summarised by a normal effective pressure only. The majority of early modelling studies tended to focus on simple deformations of interest in engineering applications (e.g. uni- or biaxial compression) and in general these assumptions worked well in describing experiments [KPJ+99, CDR04, YYYR+05, WM05a, WM05b, KSL05, WM07a]. As a result, improvements in this direction largely focussed on more accurate modelling of the strain response through higher order constitutive laws and the incorporation of viscoelasticity.

A simple way to apply this approach to the problem of buckling DEs is to
numerically solve Eq. (5.1) with a constant compressive traction across the top and bottom electrodes, fixing the outer edges of the medium so that it is forced out-of-plane as in Fig. 5.1(c). In Refs. [ZH+08, ZS08b, OMC+09], the assumption of constant prescribed electric field was used to model nontrivial out-of-plane DE shapes, though the electrostatic effects were accounted for via the material law, rather than as a boundary condition. As discussed in Sec. 5.5, this level of detail does not closely reproduce observed buckled shapes.

At the other extreme, the system may be modelled by solving Maxwell’s equations for the dielectric in concert with the elastostatics equation. In principle, this allows the electrostatic forces to be treated realistically, though in some cases the fringing effect is neglected to simplify the problem [GTC11, HCB13, WDHC16]. While implementations differ, several papers in recent years have used this general approach to model interesting out-of-plane deformations [PSZK12, KZM13, PWZK13, KWES13, VGSK14, SP16]. In particular, such a system was numerically solved by Vertechy et al., who also used an analytical model to accurately capture the deformation of buckled circular DEs [VFB+12].

In our study, we postulate that the effect of the electric fringing field can be summarised by a small constant surface traction that acts tangent to the electrode surface and normal to the boundary between the active and inactive regions. This is added to the usual effective pressure which acts normal to the active regions. The key advantage of this over alternative approaches is that it enables us to easily judge the importance of the fringing effect. We shall return to this point after explaining the model in detail.

Our setup is depicted in Figure 5.5. The normal and tangential tractions, which we call \( \tau_n \) and \( \tau_t \) respectively, enter as boundary conditions to Eq. (5.1). Both are applied to the zero-strain reference domain \( \hat{\Omega} \)—i.e. they do not depend on deformation of the medium. The normal pressure \( \tau_n \) is applied as a constant traction across the top and bottom electrodes, while \( \tau_t \) is applied in a small neighbourhood of the active region boundary of thickness \( s_0 \). The magnitude of the tangential traction is constant and its direction is given by the outward normal to the corresponding point along the electrode boundary.
In addition to \( s_0 \), there are two important length scales present in the model: the thickness of the unstrained domain \( D_0 \) and the characteristic length of the active region \( l_0 \). The exact definition of \( l_0 \) depends on the shape of the particular active region. In the case of the circle, it refers to the diameter. Finally, we account for the effect of gravity with a constant body force density that acts perpendicular to the elastomer film. Physically, it has magnitude \( \rho g \), where \( \rho \) denotes the material density (which we assume to be constant) and \( g \) is gravitational acceleration.

The idea behind \( \tau_t \) is to account for the shearing stresses that arise due to repulsion of like charges. This fringing effect occurs over a small neighbourhood of the electrode edges where the electric field has a nonzero tangential component. The application of \( \tau_t \) over the width \( s_0 \) mimics this. While the effect that it models is small, compared with the normal pressure, we shall nevertheless see that it can have a significant impact on the shapes of solutions.

For any particular DE, the relative magnitudes of the electrostatic forces do not change if the potential difference across the plates is altered, due to the superposition principal for electric fields. Therefore, the ratio \( \tau_t / \tau_n \) seems like a natural candidate for a dimensionless parameter that determines the relative strength of the tangential traction applied in the model. However, a
better choice is
\[ \kappa := \frac{s_0 \tau_t}{D_0 \tau_n}, \]  
which takes into account the length scales of the problem. To see why this is necessary, let us first note that unlike the other geometric parameters, the width \( s_0 \) is explicitly a modelling choice. If we change \( s_0 \), without altering \( \tau_t \), then the total applied tangential force changes with it. Therefore, it makes sense to combine the two parameters as \( f_t := s_0 \tau_t \). This is a linear force density along each active region perimeter, which can be adjusted to give the correct total tangential force (relative to the normal pressure), independently of the choice of \( s_0 \). The reason why the factor of \( 1/D_0 \) must be included is due to the relative scaling of \( \tau_n \) and \( f_t \). We know from Eq. (5.2) that the normal pressure (and therefore \( \tau_n \)) scales in proportion with \((V/D_0)^2\). The tangential traction on the other hand scales with \( V^2/D_0 \). Therefore, \( s_0 \tau_t/D_0 \tau_n \) is constant with respect to changes in \( D_0 \). To see why \( f_t \) scales the way it does, we can use dimensional analysis. The relevant physical parameters for the problem are voltage \( V \), \( D_0 \) and \( l_0 \). As long as \( D_0 \) is much smaller than \( l_0 \), a change in \( l_0 \) does not affect the amount of tangential traction at the electrode edge, whereas a change in \( D_0 \) does, since it affects the electric field strength (as does \( V \)). The only dimensionless combination of \( f_t \), \( V \) and \( D_0 \) is \( f_t D_0/V^2 \). Therefore \( f_t \) scales as \( V^2/D_0 \).

Hence, Eq. (5.3) defines a dimensionless parameter that dictates the magnitude of the tangent force applied in the model in a geometry-independent way. Note that \( \kappa = 0 \) means zero tangential force and that as \( \kappa \) increases, so does the relative strength of \( \tau_t \). An implicit assumption in adopting \( \kappa \) is that solutions to the model system are not significantly affected by the width \( s_0 \). This was verified for the example result, pictured in Fig. 5.6(a). We believe this to be true in general, provided that \( s_0 \) is sufficiently small.

Assuming hyperelasticity, Equation (5.1) is closed by specifying a constitutive law in terms of a strain energy density function \( W \). We shall use the isotropic Mooney-Rivlin law, which in its incompressible formulation, is
\[ W = c_1(I_1 - 3) + c_2(I_2 - 3) \]  
(5.4)
for phenomenological model parameters $c_1, c_2$, where $I_1$ and $I_2$ denote the first and second principal invariants of the Cauchy-Green strain tensor. A variety of more sophisticated laws, including the Ogden, Gent, Yeoh and Arruda-Boyce models, are often used in DE modelling studies, which can capture elastomer strain responses with greater accuracy in various circumstances [GMF05, WM05b, WM07b, XMKG10, LKB$^+$13]. However, since there is no prestretch applied in the experiments in this study (see Sec. 5.2), we work in a moderate strain regime in which we will see that the Mooney-Rivlin law is adequate. Moreover, a key advantage to using a Mooney-Rivlin solid is that we need only fit two model parameters, $c_1$ and $c_2$. It is also known that elastomers exhibit viscoelastic properties [SLL04, PD07]. We shall work only in the static setting only and do not consider viscoelasticity here.

By making the various simplifications detailed above, we sacrifice a certain degree of precision in favour of a more conceptually straightforward model. We argue that there are only two electrostatic effects of principle importance: the normal pressure and the fringing traction. The chief advantage of our model is that these effects are decoupled from one another. By varying $\kappa$, we can alter the relative strength of the fringing effect and thereby judge the degree to which it influences the equilibrium shapes of DEs. When applying our model, we use a nondimensional approach, explained in Sec. 5.5. This means that we need not worry about matching the effective pressure with the exact voltage and deformed material thickness. Instead, model parameters are fitted such that the applied tractions scale in a manner consistent with Eq. (5.2).

5.4 Methods

We perform nonlinear elasticity simulations using the finite element continuum mechanics solvers from the Chaste software libraries [MAB$^+$13], which provide an incompressible nonlinear elasticity implementation\(^1\), that was modified for our purposes. The nonlinear solver is a damped Newton’s method and the linear solver in each case is the generalised minimal residual

\[^1\]Note that incompressibility was not implemented in the DUNE code used in the previous chapter.
method (GMRES) with PETSc’s additive Schwarz preconditioner using LU factorization blocks [SBG04, BAA+14]. The deformation map is solved on a zero-strain reference domain $\hat{\Omega}$, as depicted in Fig. 5.5, using tetrahedral quadratic elements. Meshes are constructed using Gmsh [GR09], with a minimum of two layers of tetrahedra in the thickness direction. To reduce the number of degrees of freedom these are refined more at the active region and towards the centre where most of the strain occurs. Furthermore, we allow elements in the reference domain to be longer in the transverse direction than they are in their thickness. The ratio of these respective dimensions is approximately $1.5 : 1$ near the active regions and $10 : 1$ by the outer Dirichlet boundaries where there is very little deformation. In spite of these optimisations, the aspect ratios of the physical system dictate that even the coarsest possible meshes have many elements—typically our simulations use $\sim 10^5$ degrees of freedom.

### 5.4.1 Deflation

The elastostatics equation [Eq. (5.1)] can have multiple solutions. Consequently, there may be many different shapes that an elastomer can adopt in which the material is in equilibrium with the external forces imposed on it. This presents us with a problem when attempting to predict the shape of a DE: the solution that nature selects may not be the one that we find using our nonlinear solver. To address this issue, we have implemented an algorithm called ‘deflation’, whose use in the context of numerical PDE solving is due to Farrell et al. [FBF15]. It has been applied previously in the context of nonlinear hyperelasticity in [FBB16].

The basic idea behind deflation is as follows. Suppose we are trying to find the zeros of some nonlinear system $\mathcal{F}$. In our case, this means deformations for which Eq. (5.1) is satisfied. Every time we find a solution $u^*$ of $\mathcal{F}$, we augment $\mathcal{F}$ via a ‘deflation operator’ $\eta$ to obtain a new system $\mathcal{F}' = \eta \mathcal{F}$, which has the same solution set as $\mathcal{F}$, except with $u^*$ removed. We then solve $\mathcal{F}'$ to find a distinct solution. This process can be repeated to locate many additional solutions, until the nonlinear solver can no longer find any more.
There is more than one way that one might want to choose a deflation operator. A straightforward choice is

\[ \eta(u) := \frac{1}{\|u - u^*\|}, \quad (5.5) \]

which causes \( F' \) to blow up if \( u \) gets close to \( u^* \), but does not disturb any other solutions. To find multiple DE shapes in this work we use a slightly more sophisticated version. Suppose we have found solutions \( u^1, \ldots, u^n \) to \( F(u) = 0 \). Then we solve

\[ F'(u; u^1, \ldots, u^n) = \left( \alpha + \sum_{i=1}^{n} \frac{1}{\|u - u^i\|^p} \right) F(u) = 0 \quad (5.6) \]

for some \( \alpha, p > 0 \). Clearly any solution to \( F'(u) = 0 \) that is distinct from \( u^1, \ldots, u^n \) also solves \( F(u) = 0 \). The shift by \( \alpha \) is vital, since otherwise the system can be ‘solved’ numerically by pushing the intermediate guesses further and further from \( u^1, \ldots, u^n \) until the solver tolerance is met. The parameter \( p \) allows one to adjust the form of the singularity at each solution \( u_i \).

The procedure to solve Eq. (5.6) was implemented in PETSc. The augmentation of the nonlinear system by the deflation operator results in a rank-one update to Jacobian matrix, causing it to lose its sparsity. Consequently, whenever it is needed its application is implemented in terms of the Jacobian of the original system via matrix-free methods. Similarly, the preconditioner is implemented matrix-free and is computed via the original preconditioner using the Sherman-Morrison formula as suggested in [FBF15]. Below, we give some practical details concerning how deflation was used to find multiple DE shapes.

Controlling the order of the singularities in the deflation operator with \( p \) affects how close any additional candidate solutions can get to \( u^1, \ldots, u^n \), as does varying \( \alpha \). Selection of these parameters can greatly alter which solutions can be found by the nonlinear solver. Unfortunately, it is not currently possible to know \textit{a priori} what good choices of \( \alpha \) and \( p \) will be. In
the situations where deflation was used we have tried to maximise the number of solutions obtained by scanning through the $\alpha, p$-parameter space. To do this, whenever deflation is used in this work, we fix $p = 1.5$ and try many different $\alpha$ values in the range $(0, 1]$. (Whilst it would be more comprehensive to scan through a range of exponents as well, this is much more time consuming and was found to be a comparatively less effective way to locate additional solutions.) The exact values of the shifts used are not as important as the need to cover a range encompassing different orders of magnitude. We begin deflation with an initial $\alpha_0$, typically in the range $[0.5, 1]$ and find successive solutions until the nonlinear solver fails (e.g. due to exceeding the maximum allowed iterations). Each time a new solution is found, it is used as the new initial condition for the solver, after applying a small perturbation to ensure that the deflation operator is finite. After exhausting the solutions we can find with the initial $\alpha_0$, we continue, scanning through a geometric progression of shifts $\alpha_n := r \alpha_{n-1}$, until $\alpha_n < \alpha_{\text{min}}$, whereupon deflation is halted. For the systems considered below, $r = 2/3$ and $\alpha_{\text{min}} = 5 \times 10^{-3}$ were used. At higher values of $\alpha$ the nonlinear solver stays near to the previously deflated solutions since the non-deflated part of the system Jacobian is more significant with respect to the deflated part. As $\alpha$ decreases, more remote solutions become accessible, often at the expense of those with shapes that are structurally close to the deflated ones. For small values of $\alpha$, Newton’s method may take very large steps that decrease the residual of deflation operator, but not the residual of the original system. This can cause numerical instabilities if it produces an intermediate guess which is highly strained. To avoid this, we set an upper limit on the original system residual which, if reached, causes the algorithm to reset the initial condition and move on to the next $\alpha_n$. Finally, we note that after a solution has been deflated, this does not prevent Newton’s method from taking steps towards it. In general, the solver is not guaranteed to find a region where it will converge quadratically to a new solution and can spend a long time approaching already-deflated results. It is not uncommon for the method to take more than 100 iterations to converge. To catch most of the solutions, we allow for a maximum of 300 iterations.

In addition to finding solutions with deflation, we were able to find further
deformations using parameter continuation. Starting from a given initial solution with an initial $\kappa = \kappa_0$, one can successively increment or decrement $\kappa$ until a qualitatively different deformation shape is adopted by the system. Then continuing $\kappa$ gradually in the reverse direction back to $\kappa_0$ may produce a new solution. The interpretation of this procedure is that by continuing $\kappa$ we have passed a bifurcation point and uncovered a new solution branch, which we then trace in the opposite direction.

Given a set of distinct solutions, it is necessary to determine which will be preferred by the physical system. The potential energy $\Pi$ of the DE is given by integrating the strain energy density over the whole body, minus the work done by the body forces and tractions. This is

$$
\Pi(u) = \int_{\tilde{\Omega}} W(u) \, dV - \int_{\tilde{\Omega}} b \cdot u \, dV - \int_{\partial\tilde{\Omega}} \tau \cdot u \, dA,
$$

(5.7)

where $u$ is a function that gives the displacement of a material point, relative to its position in the undeformed configuration $\tilde{\Omega}$, and $\tau$ is the field of tractions on the domain boundary. We perform these integrations numerically over the discretisation mesh that we use to solve Eq. (5.1). This allows us to calculate the minimum energy solution from the shapes found.

## 5.5 Results

### 5.5.1 Circular active region

Before delving into the details of matching simulations with experiment, we present a representative simulation of an elastomer with applied normal and tangential tractions. Fig. 5.6(a) shows a solution for a circular disc-shaped elastomer with a circular active region at the centre. The plot is an oblique view of the deformed configuration, with the active region indicated. To save computational effort, we solve the elastostatics equation for only a quarter of the axisymmetric geometry. Consequently, we see a cross-section of the elastomer in the figure and may easily inspect the solution’s out-of-plane deflection. Starting from the outer extent, the profile slopes gently
Figure 5.6: (a) Example of a typical deformed configuration for a thin disc with circular active region. The inactive part is shaded in light grey and the active part is shaded in cyan. Only a quarter of the geometry is simulated—the rest is accounted for via boundary conditions which preserve fourfold rotational symmetry. The diameter of the full geometry is \( L_0 = 666^{2/3} D_0 \) and the diameter of the full active region is \( l_0 = 266^{2/3} D_0 \). Other parameters are: \( \kappa = 0.6, \tau_n = 0.5, \tau_t = 0.0225 \) and \( \rho g = 0 \). (b) Schematic showing the boundary conditions for the circular disc.

downwards, before an abrupt transition at the edge of the active region where the gradient becomes much steeper. In the bulk of the active region however, the profile levels out and is close to flat. Figure 5.6(b) indicates the boundary conditions used. The outer arc of the disc is fixed in place with a Dirichlet condition. The other two edges are free to move both in the radial direction and out-of-plane (z-direction), while their remaining degree of freedom is fixed. Solutions for these boundary conditions correspond to solutions to the full problem with at least fourfold rotational symmetry.

In each of the following cases the geometry of the simulation is set such that the aspect ratio of the finite element mesh exactly equals that of the experiment. We set our model parameters using a nondimensional approach, taking the undeformed material thickness \( D_0 \) to be the natural length unit for the system. We choose \( s_0 \) such that the tangent force is applied over a width of at least two (quadratic) finite elements. In all results, \( 10 \leq s_0/D_0 \leq 20 \). After fixing the geometry, there are five free parameters in the model: the Mooney-Rivlin constants \( c_1 \) and \( c_2 \), density \( \rho \) and the tractions \( \tau_n \) and \( \tau_t \). We fit these parameters to match equilibrium shapes of the elastomers from the experiment described in Sec. 5.2 and consequently are free to choose \( c_1 = 1 \). A different choice for \( c_1 \) leads to a corresponding rescaling of the remaining parameters in order to match experiment with simulation, but
does not change the resulting solution. Moreover, we found that varying the ratio $c_1/c_2$ had no noticeable effect on the shape of our solutions in any of the contexts studied herein. Hence, we set $c_1 = c_2 = 1$ throughout. The redundancy of the $c_2$ parameter suggests that, at least for the range and type of strains that we consider, a Neo-Hookean constitutive law ($c_2 = 0$) may be sufficient to model the elastomer well.

Figure 5.7 shows comparison between simulation and experiment for six different applied voltages. Each plot shows the midline of a numerical solution restricted to the $y = 0$ plane, together with points of experimentally measured deflection. The experimental data covers the full diameter of the elastomer so the simulation midline in this case is mirrored across the axis of symmetry in the plots. The procedure for fitting the model parameters is as follows. First, the profile of the elastomer with no applied voltage is measured. In this case, there is only one free model parameter—the material density—which is incrementally adjusted in the simulations until the amplitude at the centre matches the experiment. Next, voltage is applied in the experiment to produce significant additional strain in the elastomer and the profile is measured again, in this case at 3 kV. The nontrivial shape adopted by the data points allows us to fit both $\tau_n$ and $\tau_t$ concurrently and thereby determine $\kappa$ [Eq. (5.3)]. This is because the amplitude of the active region deflection and the shape of the profile at the electrode boundary are effectively independent of one another. These roughly correspond to the total applied traction and the ratio $\tau_t/\tau_n$ of tangential to normal traction respectively. We will return to this point shortly. After making an initial guess of their approximate magnitudes and ratio, $\tau_n$ and $\tau_t$ are incrementally increased or decreased until the solution amplitude matches the experiment. Next, to match the profile shape, $\tau_t$ is incremented or decremented. Since this changes the deflection height, it is typically necessary to readjust the total applied force in turn. Small discretionary adjustments to the tractions are then made to improve agreement further. From this point on, both $\rho$ and $\tau_t/\tau_n$ are considered to be fixed.

We know from Eq. (5.2) that the effective pressure is proportional to $(V/D)^2$. The two fitted results at 0 kV and 3 kV uniquely determine the co-
Figure 5.7: Comparison of experimentally measured deflections with simulation profiles for a succession of increasing voltages: 0 kV, 1 kV, 2 kV, 3 kV, 4 kV and 5 kV. The aspect ratio is 1 : 1. Red crosses indicate experimental data points. The experimental geometry used was a thin disc, $L_0 = 100$ mm, $D_0 = 0.15$ mm, with centred circular electrodes, $l_0 = 40$ mm. Yellow lines are midlines through a model simulation with corresponding geometric parameters and $\kappa = 0.6$ in each case. The gravitational body force is $\rho g = 3.6 \times 10^{-4}$ throughout. Applied tractions across the different voltages are (to 4 significant figures): $\tau_n = \tau_t = 0$ for 0 kV; $\tau_n = 0.01857$, $\tau_t = 8.538 \times 10^{-4}$ for 1 kV; $\tau_n = 0.07633$, $\tau_t = 3.435 \times 10^{-3}$ for 2 kV; $\tau_n = 0.18$, $\tau_t = 8.1 \times 10^{-3}$ for 3 kV; $\tau_n = 0.347$, $\tau_t = 0.01561$ for 4 kV; $\tau_n = 0.6176$, $\tau_t = 0.02779$ for 5 kV.
efficient of proportionality. However, since $p_{\text{eff}}$ depends on the deformed thickness $D$, the correct amount of normal pressure for a given voltage is coupled to the solution. Furthermore, since we apply $\tau_n$ in the reference domain, the effective pressure that this corresponds to in the deformed body also depends on $D$. Therefore, for the remaining plots, we select $\tau_n$ iteratively, each time using the computed thickness from the previous solution. This approach is also used in [WM07b]. In this way, iterations converge to produce effective pressures that scale correctly with electric field. The applied tractions used in Fig. 5.7 all obey Eq. (5.2) to within 1\% relative error. Throughout this procedure, $\tau_t$ is chosen such that $\tau_t/\tau_n$ (and thus $\kappa$) stays the same.

The profile shapes obtained this way agree extremely well across all the plots, even though the parameters were only fitted using the 0 kV and 3 kV cases. For voltages greater than or equal to 4 kV there are very small discrepancies which may, for instance, be due to the constitutive law used, or the simplified treatment of the forces acting on the elastomer in our model. Nevertheless, even at these higher strains agreement between the model and experiment is good.

For the tractions used in this particular case with a circular active region centred inside a disc, $\tau_t/\tau_n = 0.045$. Taking into account the geometric parameters, this corresponds to $\kappa = 0.6$. Let $f_t$ denote the magnitude of applied tangent force along the edge of one active region. This is given by $f_t = \pi l_0 s_0 \tau_t$. Let $f_n = \pi (l_0/2)^2 \tau_n$ denote the corresponding amount of total applied normal force. A straightforward calculation shows that $f_t/f_n = 4 s_0 \tau_t/l_0 \tau_n = 0.9\%$, so the tangential traction represents a small effect when compared with the normal compression. Due to this fact, one may wonder whether the tangential forces may simply be neglected. However, despite its magnitude, slight changes in $\tau_t$ can have a marked effect on solutions. Indeed, we find that $\kappa = 0.6$ fits the experimental data better than either $\kappa = 0.58$ or $\kappa = 0.62$, though the differences between model profiles are subtle at this level. Figure 5.8 demonstrates the much more significant effect of changing $\kappa$ by $\pm 0.2$. Here, the 3 kV experimental data from Fig. 5.7 are replotted alongside three model deflection profiles with $\kappa = 0.4, 0.6$ and 0.8. As $\kappa$ increases, the proportion of tangential force increases. This has two main effects. Increased
Figure 5.8: Effect of tangent force on the shape of model profiles. The vertical axis has been scaled by a factor of 2 to show the variation between the profiles more clearly. Red crosses are data points from the 3 kV experiment used in Figs. 5.2 and 5.7. Green, yellow and blue lines are model results with $\kappa = 0.4$, 0.6 and 0.8 respectively. In each case, the total traction was chosen so that the model profile matched the experimental deflection in the centre, at $x = 50$. The tractions were as follows: $\kappa = 0.4$ used $\tau_n = 0.156$, $\tau_t = 4.68 \times 10^{-3}$, $\kappa = 0.6$ used $\tau_n = 0.18$, $\tau_t = 8.1 \times 10^{-3}$ and $\kappa = 0.8$ used $\tau_n = 0.22$, $\tau_t = 0.0132$. All other model parameters match those from Fig. 5.7.

tension at the edges causes the active region to flatten out and stretch. This in turn modifies the shape at the electrode boundary. Both the $\kappa = 0.4$ and $\kappa = 0.8$ profiles feature an abrupt change of gradient near the active region edge. Only $\kappa = 0.6$ features the smooth transition from inactive to active region that matches the experiment. Thus, in the proceeding results, we use $\kappa = 0.6$ unless otherwise stated. In Sec. 5.5.4 we check that it is a physically reasonable value.

5.5.2 Annular active region

Another system of experimental interest is shown in Figure 5.9. In this case, the active region is annular. For sufficiently high applied voltage, this DE readily buckles to produce azimuthal waves in the active region. Wavelengths measured from the experiment are robust over a range of voltage (3 kV to 5 kV) and depend principally on the width of the annulus. These ripples in the active region are distinct from the much smaller wavelength wrinkles that result from a pull-in instability [PD06].

In this case, the active region has two edges. Consequently, there is an additional fringe effect pointing toward the centre. A diagram of the simula-
Figure 5.9: Overhead view of an elastomer experiment with an annular active region, whose geometry corresponds to the simulations in Figs. 5.10(b) and 5.10(c). The inner radius of the annulus is $r_0 = 17$ mm, the outer radius is $R_0 = 25$ mm and $D_0 = 0.15$ mm. The applied voltage is 4 kV. Azimuthal ripples are visible on the electrode; their undulation is highlighted by directing a laser across the surface. Photograph courtesy of Hadrien Bense.

tion domain is shown in Fig. 5.10(a). We use the same symmetry conditions as for the circular disc [see Fig. 5.6(b)], simulating only a quarter segment of the whole system in order to save computational cost. However, in this case solutions do not possess continuous rotational symmetry. Therefore it is important to note that the boundary conditions used place constraints on the range of admissible wavelengths. Nevertheless, since the experimentally observed wavelengths are always small enough to fit multiple waves in a quarter of the disc, we do not believe that this alters our conclusions here significantly.

Figures 5.10(b) and 5.10(c) show overhead and oblique views of a simulated result for an annulus of thickness $53^{1/3}D_0$. The dimensions of this simulation correspond to the experiment photograph in Fig. 5.9. From visual inspection one sees a qualitative agreement between the experiment and simulation, both in the overall deformation profile and the character of the waves.

As mentioned earlier, there can be many distinct solutions to the elastostatics equation [Eq. (5.1)] that are not related by symmetry. Indeed, for this system it is possible to find solutions with different azimuthal wavelengths.
Figure 5.10: (a) Diagram showing the top/bottom surface of the model setup for a circular disc with annular active region. Compressive normal pressure $\tau_n$ is applied into the page across the shaded purple area. Tangential surface tractions $\tau_t$ are applied at both boundaries of the active region in the two orange areas shown. Important length scales are labelled: the inner radius $r_0$ and outer radius $R_0$ of the annulus, the width of the active region $l_0$ and the width $s_0$ over which the tangential traction is applied. (b) Example result from the setup depicted in part (a). The blue colouration indicates deformation in the negative $z$-direction. Deeper blue means that a point is displaced further below its original position in the flat reference configuration. The active region is indicated as an area of darker shading. The geometry is set to match an experiment with $r_0 = 17$ mm, $R_0 = 25$ mm, $D_0 = 0.15$ mm and diameter $L_0 = 100$ mm. Other model parameters are: $\kappa = 0.6$ and $\rho g = 3.6 \times 10^{-4}$. (c) Oblique view of the result in part (b) showing the azimuthal ripples in the active region.
The wavelength selected by the physical system would typically be the one which minimises the potential energy, given in Eq. (5.7). This is not generally the solution first discovered by our nonlinear solver. To overcome this problem, we use the deflation method, described in Sec. 5.4.1, to find as many different solutions as we can. The result pictured in Figs. 5.10(b) and 5.10(c) is the minimum energy solution of four different equilibrium configurations computed by this technique. Likewise, the annular active region results below are minima from sets of deflated solutions. However, deflation does not guarantee that every solution will be found. To increase our confidence that these results are close the global minima, we can compare their azimuthal wavelengths with measurements from the experiment.

Figure 5.11 shows simulations with various annular active region widths. One sees that as \( l_0 \) increases, the wavenumber observed across the quarter segment decreases. This is observed in experiment: in Figure 5.12 we plot experimental and simulated ripple wavelengths against \( l_0 \) and see that both datasets follow an upward trend. There is a degree of uncertainty associated with measuring these data points experimentally. Moreover, obtaining good model results for \( l_0 > 12 \) mm is not feasible when simulating only a quarter of the full experimental domain, since the ripple wavelengths become too large. Nevertheless, the model does a good job of matching the smaller reported wavelengths in the physical system.

Figure 5.13 shows results that demonstrate the effect of tangent force for the annular active region. As in the case of a circular active region, we find
Figure 5.12: Width $l_0$ of annular active region versus observed ripple wavelength $\lambda$ for the experiment and model simulations. The experiment parameters were $D_0 = 0.15 \text{mm}$ with $R_0$ fixed at 25 mm and various $r_0$ between 7 and 22 mm. The applied voltage was 3 kV.

Figure 5.13: Deformed configurations for a circular disc with annular active region $l_0 = 40D_0$ using different amounts of tangential force. In these simulations, $\tau_n$ was fixed at 0.32 and $\tau_t$ was varied to give different $\kappa$ values as indicated. Increasing $\kappa$ means an increasing amount of tangential force. The other model parameters match those used in Fig. 5.10.
that $\kappa = 0.6$ best represents the experimental results. Decreasing $\kappa$ from 0.6 to 0.4 creases the edges of the electrodes and the spacing between ripples becomes uneven. Removing the tangent force altogether exacerbates these characteristics. Additionally, the inactive part becomes almost completely flat in spite of the action of gravity. Increasing $\kappa$ from 0.6 to 0.8 flattens the active region, just as it did for the case of circular electrodes. With this amount of tangent force, the system does not appear to support azimuthal waves.

### 5.5.3 Rectangular active region

An active region similar to that of the annulus is the case of a long rectangular strip. Provided that the length of the rectangle is sufficiently greater than its width, this system also readily buckles to produce ripples in the direction of its length. This was previously noted by Pelrine et al. [PKPJ00].

By imposing symmetries the simulation domain can again be reduced in size. We mimic a long rectangle by considering the limiting case where a thin rectilinear elastomer with rectangular electrodes is extended indefinitely in its lengthwise extent. To do this we consider a thin cuboid elastomer with its two largest faces oriented parallel to the $xy$-plane. These two faces are split along their longest direction into into active and inactive regions. This means that the active region touches the edges of the domain on three sides. We align the shorter sides with the $x$-axis and the longer sides with the $y$-axis.

The boundary conditions are as follows. The two shorter ends that are part active and part inactive are free to move in the $x$ and $z$ directions only. Fixing them in $y$ enforces periodic symmetry.\(^2\) No tangential force is applied in this direction. The two remaining longer ends enforce reflective symmetry along the axis following the side in contact with the active region, which is free to move in the $y$ and $z$-directions. The other side is held fixed. A schematic of the setup is shown in Fig. 5.14, alongside a representative result. Similar to the case of an annular active region, we must note that the finite extent of the domain means that some wavelengths are inaccessible. Guided by the

---

\(^2\)Note that since we do not require each end to deform to the same height in the $z$-direction, the period is equal to twice the length of the simulation domain. Fully periodic solutions are obtained by a reflection at either end.
Figure 5.14: (a) Diagram of the boundary conditions for simulation of a thin infinite strip. Periodic symmetry is enforced at the top and bottom edges. The other two edges implement reflective symmetry in the axis of the right-hand side. See the text for details. Surface tractions $\tau_n$ and $\tau_t$ are applied in the purple and orange regions respectively, as indicated. The definition of the characteristic width $l_0$ for this active region is as labelled. It covers half the simulated domain. (b) Overhead view of an example deformed configuration. The active region is indicated with an area of darker shading. As in the annular case, waves are present. These follow the direction of the strip’s longer dimension. The model parameters are: $\kappa = 0.6$, $\tau_n = 0.37$, $\tau_t = 0.0222$, $\rho g = 0$ and the dimensions are height $H_0 = 250D_0$ and width $L_0/2 = 92^{1/2}D_0$. 

\[ l_0 / 2 \]
results in the annular case and intuition from experiments with rectangular electrodes, we believe that the domain length chosen is sufficient to capture any important solutions.

We were able to obtain many solutions via deflation for this geometry. These are shown in Figure 5.15 with their corresponding elastic potential energies printed underneath. In this case, a variety of interesting different solutions can be found and to this end, we omitted the gravitational body force from the model. This encourages the DE to buckle up, as well as down and enables us to find more solutions. In the centre of the figure, are solutions with regular waves, analogous to those seen in the annular active region. To the left, there are four solutions composed of a large wavelength mode and smaller ripples. To the right are solutions with higher frequency ripples: one with regular ripples, another with irregular ripples and one with a smooth, mostly flat active region. For each shape shown, reflections in the planes $y = H_0/2$ and $z = D_0/2$ give solutions that are equivalent under the symmetries of the problem. These have been omitted from Figs. 5.15 and 5.16. The final two solutions to the right were found using the parameter continuation method described in Sec. 5.4. The highlighted entry is the minimum energy solution. It is worth noting that this was not the first solution to be found by the nonlinear solver. In this case it was essential to use deflation (or some alternative method) to find the multiple solutions, in order to identify the

Figure 5.15: Deflated solutions for an infinite rectangular strip. $\kappa = 0.6$. The blue and red colouration indicate deformation in the $z$-plane. Darker red (blue) means that a point is displaced further above (below) its original position in the flat reference state. Beneath each solution the potential energy computed from Eq. (5.7) is printed. Solutions that are equivalent under symmetry to the ones shown are omitted.
correct equilibrium shape of the elastomer.

The difference between the solutions in Figure 5.15 and corresponding solutions with no applied tangential force is particularly striking. Figure 5.16 shows the various solutions we were able to obtain after setting $\tau_t = 0$. The leftmost solution is highly frustrated and likely unstable. It absorbs the applied traction force via high frequency strains that are of such low amplitude as to be invisible to our colour scheme. The next two along are buckled upwards. Both have a large amplitude deflection in the widthwise direction and the rightmost features additional low wavelength ripples. These ripples give way to larger waves and folds in the following five solutions, the first few of which resemble the $\kappa = 0$ solution in the annular case (Fig. 5.13). The final solution admits a large downward crease that breaks translational symmetry.

Whilst we do not believe the solutions in Fig. 5.16 are a faithful description of the physical experiment, they are undoubtedly curious from a pattern forming perspective. Moreover, the stark contrast between Figs. 5.15 and 5.16 demonstrates graphically the important role played by the tangential forces on the elastomer electrodes.

**Figure 5.16:** Deflated solutions for an infinite rectangular strip with no applied tangential force, $\kappa = 0$. The blue and red colouration indicate deformation in the $z$-plane. Darker red (blue) means that a point is displaced further above (below) its original position in the flat reference state. Beneath each solution the potential energy computed from Eq. (5.7) is printed; the lowest value is red. Solutions that are equivalent under symmetry to the ones shown are omitted.
5.5.4 Computing $\kappa$

By tuning the tractions $\tau_n$ and $\tau_t$, we have obtained numerical results that closely match phenomena observed in experiments on DEs. Moreover, we claim that $\kappa = s_0\tau_t/D_0\tau_n$ is the critical dimensionless parameter that determines elastomer shape selection. If we are to believe that our simple way of modelling the DE is faithful to the physics of the experiment, we should independently verify the forces imposed on the elastomer by the electric field, i.e. the ratio of the effective pressure $\tau_n$ to the magnitude of the fringing effect $\tau_t$ at the boundaries of the active regions.

Both electrodes are held at fixed potential in the experiment. The edges are interfaces between finite charge density inside the active region and zero charge density outside, leaving a singularity in the electric field. We are unaware of any method for calculating the forces at the electrode boundaries analytically, although the fundamental form of the singularity for the edge of a flat plate is a classical result [Jac75, Sec. 1.6].

One way to deal with the singularity is to compute instead the aggregate force $F$ on a small volume $B$ that contains part of the electrode edge. This can be computed by integrating the electromagnetic body force density $f$ that is associated with the electric field. Electrostatics tells us that a simple formula for this is $f = \varrho E$, where $\varrho$ is the charge density. However, a more convenient way to write $f$ for our purposes, is as the divergence of the classical Maxwell stress tensor $S^M$, the derivation of which may be found in standard textbooks (see e.g. [Str41, Sec. 2.5] or [PP62, Sec. 6.5]). It is defined both in and outside the elastomer, by:

$$S^M := \varepsilon (E \otimes E - \frac{1}{2}(E \cdot E)I). \quad (5.8)$$

The advantage of this is that it allows us to use the divergence theorem to integrate around the boundary $\partial B$, sidestepping the electric field singularity as so

$$F = \int_B f \, dV = \int_B \nabla \cdot S^M \, dV = \int_{\partial B} S^M \hat{n} \, dA, \quad (5.9)$$

where $\hat{n}$ denotes the outward unit normal of $\partial B$.

To compute the Maxwell stress tensor, we assume that the electric field
in the deformed elastomer may be locally approximated by the field in an undeformed capacitor. Therefore the electric field can be numerically calculated in the reference domain by solving Laplace’s equation for the potential $\phi$ and numerically differentiating the result. That is, we solve

$$\nabla \cdot (D \nabla \phi) = 0, \quad E = -\nabla \phi,$$

where $D = \epsilon I$ is a matrix that accounts for the presence of a dielectric. Figure 5.17(a) shows the solution and labels the accompanying boundary conditions. We solve the problem in two dimensions in the unit square

![Figure 5.17](image-url)

**Figure 5.17**: Plots of the solution of electric potential $\phi$ for a rigid capacitor in two dimensions. The colour scheme indicates the strength of the potential from $\phi = 0$ (blue) to $\phi = 0.5$ (white) to $\phi = 1$ (red). (a) Full solution in the unit square, representing a quarter of the elastomer, with boundary conditions labelled around the edges. Although the capacitor is too thin to be visible at this scale, the gradual decay of the potential outside the top plate can be seen in the bottom left-hand corner. (b) Highly zoomed plots showing sections of the capacitor corresponding to: (from left to right) the centre of the electrode, the right-hand edge of the electrode and the edge of the dielectric. The $\phi = 1$ boundary condition on the top plate is shown dotted in white. Also indicated are the regions inside and outside the elastomer and their relative dielectric constants. (c) Close-up of the electrode centre indicating the region of integration for numerical computation of $\tau_n$ in dashed pink. (d) Close-up of the electrode edge with contour lines drawn indicating isolines of potential. The integration region for numerical computation of $\tau_t$ is drawn in dashed pink.
and for a quarter of the elastomer only. The boundary conditions account for the rest of the system via symmetry. These are: $\phi = 0$ far from the electrode $([0,1] \times \{1\} \cup \{1\} \times [0,1])$ and along the horizontal symmetry axis $([0,1] \times \{0\}$), $\phi = 1$ along (half) the top electrode $([0,l_0/2] \times \{D_0/2\})$ and $\partial_x \phi = 0$ along the vertical symmetry axis $\{0\} \times [0,1])$. The active region length $l_0$ was chosen such that the electrode surface was far from top and right-hand sides so that the Dirichlet condition there did not affect the solution significantly. Specifically, the geometric parameters used were: $l_0/2 = 0.2$, $L_0/2 = 0.5$ and $D_0/2 = 7.5 \times 10^{-4}$, where $L_0$ denotes the length of the dielectric. The aspect ratios $l_0/D_0$ and $L_0/D_0$ match the geometry used above for the case of a circular disc with circular active region. By solving in two dimensions however, we implicitly assume that the electric field is constant in the third direction, which is equivalent to considering the system in which the elastomer extends infinitely in this ordinate. Therefore the two dimensional system is strictly applicable only to the cases with rectangular active region where the electrode boundary in the third direction is sufficiently far away. Nevertheless, it can also be used as an approximation for elastomers with circular and annular active regions provided the curvatures of the electrode boundaries are not too large.

Figure 5.17(b) shows sections of the solution containing the dielectric in close-up. The region containing the dielectric is labelled with $\varepsilon = 2.6$. This is the relative permittivity measured for the elastomers used in the experimental system. Also visible is the top capacitor plate which is labelled with the $\phi = 1$ boundary condition. Between this and the horizontal axis the gradient of the potential is essentially linear everywhere and is responsible for the constant pressure between the two electrodes. This is given classically by Eq. (5.2) in the rigid case: $\tau_n = -p_{eff} = \frac{1}{2} \varepsilon E^2 = \frac{1}{2} \varepsilon \left( \frac{\partial \phi}{\partial y} \right)^2$. For the geometry used this gives $\tau_n = \frac{1}{2} \varepsilon \left( \frac{1}{D_0/2} \right)^2 = 2.31 \times 10^6$.

To numerically compute the tractions from the electric potential solution we calculate the Maxwell stress tensor via Eq. (5.8) and integrate over a region of our choosing. We do this first for the normal pressure, to check that we can recover the classical result. Figure 5.17(c) shows a close up of the centre of the
electrode. Dashed in pink is a closed rectangular curve bounding a portion of the top capacitor plate. Integrating the Maxwell stress tensor along this curve gives the total force due to the electric field on the enclosed region. Taking the normal component of this integral and dividing by the width of the box (0.02 in this case), we obtain a value for the normal pressure \( \tau_n = 2.307 \times 10^6 \), which is within less than 0.2% relative error with respect to the theoretical value.

Likewise, in Figure 5.17(d) we show a close-up of the electrode edge with a closed rectangular integration region indicated in dashed pink. The width of this region (relative to \( D_0 \)) corresponds to the width over which the tangential force was applied in the model simulations. Integrating the Maxwell stress along the curve, taking the tangential component of the result and dividing by the width gives the tangential traction \( \tau_t = 8.603 \times 10^4 \).

These numerical values for \( \tau_n \) and \( \tau_t \) yield \( \kappa = 0.50 \) to 2 significant figures. This differs from the value used in our model simulations by 17%, indicating that perhaps there is less tangent force than our fitted tractions suggest. A possible reason for this discrepancy is that the experimental system stretches slightly due to curing of the elastomer after it is placed on the frame, adding some tension in the elastomer that is not attributable to the fringing field. Another reason lies in the fact that we computed the electric field in the reference configuration, essentially assuming that the elastic and electrostatic effects can be decoupled from one another. Moreover, solving in two dimensions overlooks any correction arising from the geometry.

At this point, the reader may object that in the deformed case, the effective pressure is twice the classical value (as noted in Sec. 5.3.1) and therefore our predicted value of \( \kappa \) ought to be halved. The reason for this is explained in the Appendix to this chapter—coupling of the compressive pressure to lateral stretching via incompressibility gives rise to additional force between the electrodes. It seems reasonable that the same principle applies to the tangential traction, implying that the effective tangential traction should also be twice the value that is has in the undeformed setting. However, so far a physical derivation eludes us.

Taking these issues into account, the difference between the two results
is not unreasonable. However, deeper understanding is needed in order to establish the accuracy of this prediction. It is worth noting that while $\kappa = 0.6$ provided the best fit across all the results shown in Fig. 5.7, lower values, down to and including $\kappa = 0.5$, produced acceptable results. Furthermore, $\kappa = 0.5$ certainly supports our central thesis that the tangential forces are significant enough to be important for shape selection.

5.6 Discussion

We have presented a straightforward numerical model for capturing the shape of buckled DEs. The electrostatic forces acting on the dielectric are input as boundary conditions to the nonlinear elastostatics equation. We proposed that the aggregate effect of the electric field could be decomposed into an effective pressure due to the constant 'classical' part between the electrodes and a small tangential traction that models the effect of the fringing field. Although these boundary conditions represent a dramatic simplification of the underlying physics, they are nonetheless able to produce close fits to experimental data. In principle, this approach is independent of the geometry of both the elastomer and the active region.

The impact that the edge effect has on solutions is significant, despite its small magnitude. If the effect is left out of the model, we are unable to obtain deformation profiles that are even qualitatively correct. This has clear implications for future modelling work.

The magnitude of the fringing force, relative to the effective pressure is captured by our model in a dimensionless constant $\kappa$. We were able to tune $\kappa$ to a value that produced solutions matching experimental deformation profiles. This value proved robust across different applied voltages and active regions. It was further validated by calculating the relative magnitudes of the corresponding forces in a rigid capacitor from a numerical solution of the Laplace equation.

In the case of an annular active region, we report that the elastomer buckles to produce azimuthal waves, which are localised in the vicinity of the
electrodes. Their wavelength increases in proportion with the width of the annulus. This trend is captured well by our model which produces solutions in qualitative agreement with the experiment.

Another key finding is that care must be taken when solving the elastostatics equation in the case of buckling DEs, since solutions are nonunique. This has significance for any study that uses numerical methods to predict elastomer deformation in a nonlinear elasticity context. In computing a single solution to Eq. (5.1), one cannot guarantee that it corresponds to the equilibrium shape with the lowest possible potential energy. Consequently, it is desirable to find many different solutions and work out which is favoured by the system, either by computing their potential energies via Eq. (5.7), comparing with experimental data, or using some other physical argument. Deflation is one such technique that can be used to find multiple solutions [FBF15]. For the annular active region, this was used to find the lowest energy azimuthal wavelength. Crucially, the most energetically favourable mode is not typically the first solution that our nonlinear solver converges to. In the analogous case of a long rectangular strip, we used deflation to find many interesting deformation patterns including waves of different wavelength, wrinkles and creases.

5.A Appendix: Effective pressure derivation

Our derivation of the effective pressure is similar in spirit to the original by Pelrine et al. [PKJ98], using the classical formula for the energy stored in a capacitor. An equivalent derivation can be performed starting from the Maxwell stress [KSL05]. We diverge from these presentations by permitting the elastomer to have a general Poisson ratio $\nu$. Although the materials used in DE applications may be treated as incompressible materials to a very good approximation, this approach allows us to highlight the difference between our system and that of a rigid capacitor.

Consider the DE depicted in Fig. 5.1(a) and apply fixed charges $+Q$ and $-Q$ on the top and bottom electrodes respectively. For now, we make the
classical assumption that this generates an electric field that is constant with magnitude $E$ directly between the electrodes and zero everywhere else. It compresses this region of the dielectric uniaxially, stretching its diameter from $l_0$ to $l$ and squeezing its thickness from $D_0$ to $D$. Let $A = \pi(l/2)^2$ be the area of the deformed electrodes. Then the electrostatic energy of the two charged surfaces, as commonly derived in textbooks, is

$$U = \frac{Q^2 D}{2 \varepsilon A}. \quad (5.11)$$

The total force $F$ between the electrodes is given by differentiating with respect to the thickness

$$F = -\frac{dU}{dD} = -\frac{Q^2}{2 \varepsilon} \left( \frac{1}{A} - \frac{D}{A^2} \frac{dA}{dD} \right). \quad (5.12)$$

The key difference from the classical case is the appearance of the $dA/dD$ term, which is nonzero for a general elastic material. This derivative can be computed in terms of the Poisson ratio. Let $\varepsilon_{rr}$ and $\varepsilon_{zz}$ denote the principal components of strain in the radial and thickness directions respectively. Then

$$\nu := -\frac{d\varepsilon_{rr}}{d\varepsilon_{zz}} = -\frac{d\varepsilon_{rr}}{dl} \frac{dD}{d\varepsilon_{zz}} \frac{dl}{dD} = -\frac{2}{\pi l} \frac{D_0}{l_0} \frac{dA}{dD}. \quad (5.13)$$

Substituting this into Eq. (5.12) and simplifying using $A = \pi(l/2)^2$ gives

$$F = -\frac{Q^2}{2 \varepsilon A} \left( 1 + 2\nu \frac{l_0}{l} \frac{D}{D_0} \right). \quad (5.14)$$

Since Poisson’s ratio is only makes sense in the small-strain limit, for this analysis we must take $l_0/l \sim 1$, $D/D_0 \sim 1$, though the incompressible case is independent of this assumption. Noting also that the electric field strength is $E = Q/(\varepsilon A)$ and dividing through by the electrode area to obtain the (effective) pressure $p_{\text{eff}}$ acting on the top and bottom surfaces, we obtain

$$p_{\text{eff}} = \frac{F}{A} = -\frac{1}{2} \varepsilon (1 + 2\nu) E^2. \quad (5.15)$$
For $\nu = 0$, the material does no work extending laterally and we recover the classical pressure between two fixed capacitor plates. For $\nu = \frac{1}{2}$ the material is incompressible and we obtain Pelrine’s effective pressure between two electrodes of a dielectric elastomer. Interestingly, in the exotic case of negative Poisson ratios, the pressure changes sign. For materials with $\nu \in [-1, -1/2)$ it is more energetically favourable for the top and bottom surfaces to separate (with a corresponding in-plane dilation) than for them to come together, in spite of the direction of the electric field.
Chapter 6

Conclusions

We conclude this thesis by summarising the work done and the main contributions of each chapter.

In Chapter 2 we presented the results of extensive numerical simulations of the spiral pinball system. This is first and foremost a study on the dynamics of drifting spiral waves, following in the footsteps of Biktashev and Holden [BH93, BH95]. We showed that reflections are nonspecular and that in most cases reflection angle is essentially constant over wide ranges of incident angle. The data obtained were compared with predictions from a phenomenological model due to Biktashev and Holden [BH93], which was shown to capture some, but not all of the aspects of the incidence-reflection angle dependence. This is particularly true of the large-core case, which was not originally considered by Biktashev and Holden. Here, we observed a variety of different behaviours at the boundary: annihilation, phase-dependence and two circumstances where spirals are attracted to the boundary, which we call glancing and binding. The dynamics in these latter two cases are yet to be fully understood. Long-term behaviour in both the small and large core cases was also considered. In the small-core case, we presented a dynamical systems argument for why the pinball trajectories are attracted to a stable square orbit in the small-core regime, while in the large-core case the square is unstable and trajectories collapse towards the diagonal.

The reflections of spiral pinballs bear considerable resemblance to those
of the systems that originally motivated our study. In particular, stable square orbits are ubiquitous, since they emerge naturally from the incidence-reflection dependence of the given system. Moreover, the selection of a ‘preferred’ reflection angle also appears to be emerging as a theme. A recent paper by Pucci et al. demonstrates the reflection behaviour of the bouncing droplet system [PSFB16]. They observe only a small range of reflection angles which depends weakly on system parameters. Certain swimming microorganisms also exhibit similar reflection behaviour [SL12]. This was drawn to our attention after our own study had concluded. In this case, the long-term dynamics have been studied in a manner that generalises to other nonspecular billiard systems [WLST15].

We extended the study of spiral pinballs in Chapter 3, using the asymptotic theory of response functions, which reduces the dynamics of drift to three coupled ODEs whose form can be obtained numerically. We were thereby able to describe the underlying details governing reflections. The fundamental reason for reflection in the small-core regime was already known to Biktašhev and Holden: perturbation of the spiral core alters the rotation frequency, pushing the system off resonance and changing the direction of drift. However, by carefully examining the response function ODEs, one is able to see explicitly the details of the process. We described the dynamics and explained the qualitative form of the dependence of reflection angle on incidence angle in both the small and large core regimes. Reflections of large-core spirals are particularly nontrivial. Response functions explain typical large-core reflections well, but were found not to predict the more exotic glancing and binding trajectories. In the asymptotic limit, we compared response function predictions with direct numerical simulations and found excellent agreement. Moreover, our analysis agrees qualitatively with our observations of spiral pinballs in Chapter 2 that exist far beyond this limit.

In Chapter 4, we directed our attention to spiral waves in deforming media. Though motivated by the case of cardiac tissue, which features reciprocal coupling between nonlinear elastic deformations and excitable wave propagation, our primary interest was a generic system in which spiral wave and mechanics interact with one another. We covered the theory behind such a
system in detail, discussed a suitable model and its implementation via the finite element method. A number of different results using this implementation were demonstrated.

Finally, in Chapter 5 we proposed a new approach to modelling dielectric elastomers (DEs). These are elastic capacitors made with compliant electrodes that deform into interesting shapes when subjected to a potential difference. Our primary contention is thus: in situations where it is important to match out-of-plane deformation shapes accurately, it is necessary to model the effect of the electric fringing field at the boundaries of the electrodes. This edge effect is small and typically considered to be negligible. Through comparison with experiments conducted by our colleagues, we showed that it nevertheless plays an important role in the selection of equilibrium shapes. When the fringing effect was accounted for, agreement between experiment and simulation was excellent. Neglecting the effect led to poor agreement that in some situations was not even qualitatively accurate. The model proposed decomposes the stresses induced by electrostatic forces into two components input via boundary conditions: a normal pressure due to mutual attraction of the electrodes and a tangential traction due to the fringe effect. The advantage of this approach is its simplicity. Simulations are straightforward compared to more detailed models and the fringing effect is easily investigated by varying its magnitude. With this, we captured simple out-of-plane deformations and nontrivial patterns. In the latter case, we demonstrated the nonuniqueness of equilibrium shapes and the importance of computing multiple solutions. An algorithm called ‘deflation’, whose application in the context of nonlinear PDE solving was recently proposed by Farrell et al. [FBF15], was shown to be useful and necessary in this context. At the end of the chapter, by solving the Laplace equation, we independently verified that the applied traction used in the model was of an appropriate magnitude, thereby further justifying our model.
Bibliography


[NJBWHP06] W. Ning-Jie, L. Bing-Wei, and Y. He-Ping, Effects of periodic forcing amplitude on the spiral wave resonance drift, Chinese Phys. Lett. 23 (2006), no. 8, 2030.


154


