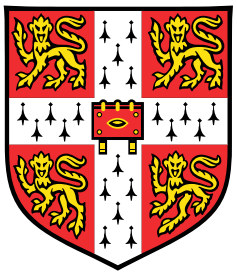
Mathematical models of soft elasticity: elastomers, active solids and topological metamaterials



Ian Roong Chern Tan

University of Cambridge

This thesis is submitted for the degree of
Doctor of Philosophy

Trinity Hall June 2025

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any work that has already been submitted, or, is being concurrently submitted, for any degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.

Ian Roong Chern Tan
June 2025

Abstract

The mechanics of soft solids, from everyday materials like rubber to exotic metamaterials for emerging technologies, is elegantly described by the theory of elasticity, which uses a continuum field approach to capture universal features in the macroscopic physics of solids. Using this approach, we study three types of soft material, ranging from conventional to exotic.

We begin in Part I with elastomers, which are used in many technological applications that rely on hyperelastic constitutive models to predict the nonlinear stress–strain response of these materials. To obtain more accurate models, we use the Irving–Kirkwood–Noll procedure to construct a Cauchy stress tensor from forces along polymer chains. Combining this stress with previously reported experimental results on chain orientations in deformed polymer networks, we present two new chain stretch relations relating the molecular-level stretching of polymer chains to the macroscopic deformation. Our new models have only two fitting parameters, and fit experimental data more accurately than previously proposed two-parameter models.

In Part II, we turn to active solids, which consist of energy-consuming constituents that enable work to be extracted using closed cycles of deformation. Previous work on linear odd elasticity captured their rich phenomenology, which occurs in both living systems and robotic metamaterials. We generalise odd elasticity to the nonlinear regime and identify a new odd modulus for quadratic terms that give rise to deformation-induced odd elasticity. Unlike linear odd elasticity, this odd response does not rely on broken chiral symmetry, so it can exist in three-dimensional isotropic materials, and enables work extraction from closed deformation cycles for which the linear theory predicts zero net work. We show that random networks of active fibres, in which each fibre's stiffness depends on the stretches of surrounding fibres, exhibit this novel form of odd elasticity.

In Part III, we consider topological mechanical metamaterials that exhibit robust floppy modes, characterised by topological invariants. To capture universal features of these modes in the continuum, we augment linear elasticity with additional fields and define a continuum version of Maxwell counting, which expresses the balance between degrees of freedom and mechanical constraints. We define topological invariants for continuum floppy modes and show that one additional field is necessary to capture topological edge modes, whereas two are required for Weyl points, which are topological bulk modes. Although our results are independent of microscopic detail, we show that our theory arises naturally when taking the continuum limit of lattices, enabling efficient prediction of coarse-grained deformations.

Preface

The work presented in Part I was supervised by Thierry Savin and John Biggins, and has been published in the *European Journal of Mechanics - A/Solids* as Tan et al. [1].

The work presented in Part III was supervised by Anton Souslov, and has been published in the *New Journal of Physics* as Tan and Souslov [2].

Acknowledgements

Firstly, I thank Thierry Savin, John Biggins and Anton Souslov for their supervision of the projects that have come together in this PhD thesis. I am thankful to Thierry for his support as a supervisor, providing helpful guidance on research as well as invaluable assistance in securing funding and facilitating my participation at programmes and conferences. I am grateful to John for stimulating conversations that pointed me towards the research questions I ended up working on during my PhD. I thank Anton for his generosity in being willing to take me on as his student, including me in his research group, making time for many fruitful discussions about physics, teaching me how to write well, arranging opportunities like the visit to KITP and giving helpful advice on research careers. I am thankful to Luca Magri for being an excellent supervisor during my Master's project, and for his support and encouragement towards doing a PhD. I gratefully acknowledge PhD funding from the Cambridge Trust.

I am grateful for the many people who have been part of my life in Cambridge: firstly, the past and present members of the Christian Graduate Society, Garang Sheep, CICCU and the church at St A's. Their companionship and hospitality have been great blessings. I am thankful for CH in college, the Division C lunch group, and the Downing College Ultimate team, who have enriched my time here in many ways. I also thank Jia Min for the many discussions about physics and for introducing me to the textbooks that led me towards an interest in continuum mechanics. I thank my Additional Mathematics teacher Mrs Choong for encouraging my interest in mathematics. I thank my family for being supportive and encouraging throughout. Finally, I thank God, who has provided in ways I could not have anticipated or imagined.

Great are the works of the LORD*; they are pondered by all who delight in them.*
Psalm 111:2

Table of contents

List of figures **xiii**

List of tables **xv**

List of key symbols **xvii**

1 Introduction **1**

I New constitutive models for elastomers **5**

2 Introduction to constitutive models for elastomers **7**

2.1 Nonlinear elasticity 9

2.2 Statistical mechanics of polymer chains 12

2.3 Molecular-statistical constitutive models 13

3 New chain stretch relations **19**

3.1 Motivation: experimental observations of chain orientation in vulcanised rubber 19

3.2 Exponential chain stretch relation 20

3.3 Hybrid-exponential chain stretch relation 30

4 Comparison with experimental data **35**

4.1 Constitutive models under consideration 35

4.2 Qualitative prediction of the Poynting effect in simple shear 38

4.3 Quantitative comparison of constitutive models 39

4.4 Summary 52

5 Conclusions: elastomeric constitutive models **53**

Appendix I **55**

I.1 Derivatives with respect to tensors 55
I.2 Relationship between derivatives of the chain stretch 56
I.3 Kinetic stress is a multiple of the identity tensor 56
I.4 Cauchy stress for affine full network model 57
I.5 Modifying the ANM chain stretch 59
I.6 Different choices of the hybrid-exponential interpolation parameter 60
I.7 Comparison of three-parameter constitutive models 61

II Nonlinear active elasticity 65

6 Introduction to active elastic solids 67
6.1 Active matter and odd elasticity 67
6.2 Active elasticity in the nonlinear regime 69

7 Phenomenology of nonlinear active elasticity 73
7.1 Odd elasticity about deformed states 73
7.2 Isotropic elasticity 76
7.3 Weakly nonlinear isotropic elasticity 78

8 Nonlinear active elasticity of random networks 87
8.1 Homogenisation of a random network 87
8.2 Networks of active fibres 91
8.3 Three-dimensional networks 92
8.4 Two-dimensional networks 98

9 Conclusions: nonlinear active elasticity 103

Appendix II 105
II.1 Antisymmetric part of of the elasticity tensor 105
II.2 Hyperelasticity and major symmetry 106
II.3 Representation theorem for isotropic tensor functions in two dimensions . 109
II.4 Antisymmetric part of the elasticity tensor in the weakly nonlinear theory . 112
II.5 Derivatives of shear and normal stress for an isotropic hyperelastic material 113
II.6 Integrals of tensors over the unit sphere 115

III Topological mechanics in the continuum 119

10 Introduction to topological mechanics 121
10.1 Ball-and-spring networks 122
10.2 Topological invariants 128
10.3 Continuum theories for topological mechanics 131

11 Augmented elasticity for topological mechanics 133
11.1 Generalised elasticity and the Maxwell criterion 133
11.2 Continuum limit of a ball-and-spring lattice 140

12 Topological floppy modes in the continuum 153
12.1 Topological edge modes 154
12.2 Weyl zero modes 162

13 Conclusions: topological floppy modes in the continuum 167

Appendix III 171
III.1 Determinant of the compatibility matrix for a geometrically perturbed lattice 171
III.2 Details of the homogenisation procedure 172
III.3 Counting continuum edge modes with a topological invariant 181
III.4 Guest–Hutchinson modes, Poisson ratios and Δ 184
III.5 Analyticity of q_2 as a function of q_1 corresponding to continuum zero modes 190
III.6 Orthogonal soft directions 191
III.7 Characterising edge floppy modes in terms of polarisation directions 192
III.8 Topological polarisation in discrete lattices 193
III.9 Solving partial differential equations for topological floppy modes 195

References 197

List of figures

1.1 Branches of elasticity theory used for the classes of soft material in this thesis. 1

2.1 The definition of a chain stretch relation. 14

3.1 Different assumptions used to model the effect of deformation on the chain orientation distribution. 26

4.1 Model predictions of normal stress T_{22} in simple shear against shear strain γ for different chain stretch relations. 38
4.2 Stress–stretch relations predicted by constitutive models fitted to data from Kawabata et al. 43
4.3 Stress–stretch relations predicted by constitutive models fitted to data from Treloar. 45
4.4 Stress–stretch relations predicted by constitutive models fitted to data from Meunier et al. 47
4.5 Stress–stretch relations predicted by constitutive models fitted to data from Moreira and Nunes. 49
4.6 Stress–stretch relations predicted by constitutive models fitted to data from Janmey et al. 51

I.1 Stress–stretch relations predicted by the hybrid-exponential model for different choices of $\hat{p}(\frac{1}{2}\sqrt{N})$. 61
I.2 Stress–stretch relations predicted by three-parameter constitutive models fitted to Treloar's dataset. 62

7.1 Closed cycles of deformation for work extraction. 82

8.1 Constructing a stress tensor for a random distribution of fibres. 88
8.2 Three-dimensional random networks and their shear responses. 94
8.3 Plots of the normalised odd modulus $d_o/\tilde{\sigma}$ against the parameter $\eta/\bar{k}$. . . . 97

8.4 The dependence of the odd modulus d_o on the fraction p_a of active fibres in a random network consisting of a mixture of active and passive fibres. . . . 98
8.5 Two-dimensional random networks and their shear responses. 100
8.6 Comparison of local average and transverse force networks. 101

10.1 Floppy modes on a strip cut from a two-dimensional lattice. 129

11.1 Physical interpretation of the scaling relationship $\varepsilon = O(\delta)$. 144
11.2 Obtaining continuum fields from a discrete lattice. 147
11.3 Constructing distorted kagome lattices. 151

12.1 Shear-dominant Guest–Hutchinson modes imply topological polarisation in the continuum. 161
12.2 Efficient computation for floppy modes in the continuum. 163
12.3 Constructing a lattice with continuum Weyl points. 165

13.1 Classification of topological phenomena in two dimensions. 169

List of tables

4.1 Model parameters and mean squared error when fitted to experimental data from Kawabata et al. 42
4.2 Model parameters and mean squared error when fitted to experimental data from Treloar. 44
4.3 Model parameters and mean squared error when fitted to experimental data from Meunier et al. 46
4.4 Model parameters and mean squared error when fitted to experimental data from Moreira and Nunes. 48
4.5 Model parameters and mean squared error when fitted to experimental data from Janmey et al. 50

I.1 Model parameters and mean squared error (MSE) for the hybrid-exponential model with different choices of $\hat{p}(\frac{1}{2}\sqrt{N})$, fitted to Treloar's dataset. 61
I.2 Model parameters and mean squared error (MSE) for three-parameter models fitted to Treloar's dataset. 63

List of key symbols

Roman Symbols

$\mathbf{B}$	Left Cauchy–Green deformation tensor
$\mathbf{C}$	Right Cauchy–Green deformation tensor in Parts I, II and compatibility matrix in Part III
C	Elasticity tensor
$\mathbb{C}$	The set of complex numbers
d	Number of spatial dimensions
$\mathbf{F}$	Deformation gradient
$\mathrm{L}(\mathcal{V}_d)$	The set of tensors acting on $\mathcal{V}_d$
$\mathrm{O}(d)$	The set of orthogonal tensors acting on $\mathcal{V}_d$
$\mathbf{P}$	First Piola–Kirchhoff stress tensor in Parts I, II and matrices related to projections in Part III
$\mathbf{Q}$	Orthogonal tensor in Parts I, II and equilibrium matrix in Part III
$\mathbf{q}$	Wavevector
$\mathbf{r}$	Chain or fibre end-to-end vector
$\mathbb{R}$	The set of real numbers
$\mathbf{S}$	Second Piola–Kirchhoff stress tensor
$\mathrm{SO}(d)$	The set of proper orthogonal tensors acting on $\mathcal{V}_d$
$\mathbf{T}$	Cauchy stress tensor

$\mathbf{u}$	Displacement vector
$\mathcal{V}_d$	Real, d-dimensional Euclidean vector space
$\mathbf{W}$	Antisymmetric tensor for anticlockwise rotation by $\frac{1}{2}\pi$ in two-dimensions
W	Strain energy function
$\mathbb{Z}$	The set of integers

Greek Symbols

λ	Stretch or generic eigenvalue
Λ	Generalised strain measures
φ	Additional kinematic field
Ψ	Generalised displacements
ψ	Single chain Helmholtz free energy
$\boldsymbol{\sigma}$	Stress tensor in linear elasticity
Σ	Generalised stress measures

Superscripts

$*$	Complex conjugate
$\dagger$	Conjugate transpose of a tensor or matrix
T	Transpose of a tensor or matrix

Other Symbols

asym	Antisymmetric part of tensor or matrix
det	Determinant of tensor or matrix
$\cdot$	Dot product (single contraction)
:	Double contraction of tensors
sym	Symmetric part of tensor or matrix
$\otimes$	Tensor product between vectors
tr	Trace of tensor or matrix

Chapter 1

Introduction

The theory of elasticity, which uses continuum fields to capture the mechanics of solids, elegantly describes the physics of a broad range of materials. This theory is built on the contributions of many eminent mathematicians and physicists from the late 17th century onwards [3]. Notably, Gottfried Wilhelm Leibniz and Jakob Bernoulli introduced the concept of an internal tension that acts across surfaces in a deformed solid [3]. Another key step was Bernoulli's description of a material fibre under tension by expressing stress as a function of strain [3]. The use of tensors to express three-dimensional states of stress and strain was introduced by Augustin-Louis Cauchy in the early 19th century [3, 4], whose work laid the foundation for the modern theory of elasticity. From the 1940s, this theory was rigorously and systematically developed to account for large material deformations [4–6], forming the basis of models for soft solids such as rubber [7] and biological tissue [8].

In elasticity theory, stress depends only on variables that describe a material body's current state of deformation [6], in contrast to fluid mechanics and plasticity theory, in which stress depends on the rate and history of deformation, respectively [9]. In this thesis, we study the elasticity of three classes of soft material: elastomers, active solids and topological metamaterials, devoting one thesis part to each class. Although these material classes exhibit significantly different properties, our work shows that this broad range of material behaviour

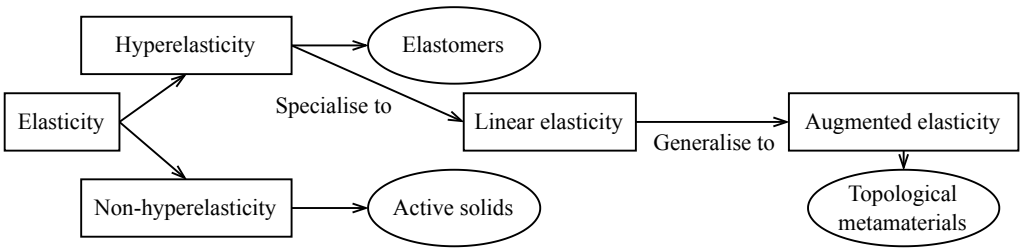


Fig. 1.1 Branches of elasticity theory used for the classes of soft material in this thesis.

can be captured using the approach of elasticity theory, which unifies the different parts of this thesis. Fig. 1.1 provides a taxonomy of the continuum models for the materials we study, showing how they fall under different branches of elasticity theory. Hyperelasticity is the branch suitable for modelling passive materials, like elastomers, whose stress–strain relations are derivable from strain energy functions, which quantify the elastic potential energy due to material deformation [6]. In contrast, no such quantity exists for active solids due to the energy consumption of their constituents, which can lead to non-zero work extraction from closed cycles of deformation [10, 11]. Like elastomers, topological metamaterials are passive, but they exhibit floppy modes [12–14] that cannot be captured by elasticity involving only the usual strain tensor [15–17]. We therefore capture their phenomenology by augmenting standard linear elasticity with additional continuum fields, which characterise the kinematics of soft degrees of freedom in the metamaterials.

Another theme connecting all three parts of this thesis is the use of coarse-graining procedures to obtain continuum models from microscopic physics. Here the term 'microscopic' refers to the length scales of material constituents, whether they are polymer chains in elastomers or metamaterial components that could be several millimetres in size. Bridging macroscopic and microscopic descriptions of these materials not only enables the functional forms of continuum models to be obtained in a principled manner, but also provides design principles for achieving desired material properties. In Parts I and II, we present coarse-graining procedures for random networks of polymer chains and active fibres, respectively, whereas in Part III, we take the continuum limit of ball-and-spring lattices used as models for mechanical metamaterials. Because the active solids in Part II and the topological metamaterials in Part III could be realised in many different ways, we complement these 'bottom-up' coarse-graining approaches with continuum models whose functional forms are obtained using symmetry arguments that are independent of microscopic detail. Such phenomenological approaches that make no reference to microscopic realisation have the advantage of giving insight into the behaviour of a broad range of (meta)materials, even those that are yet to be imagined.

We review relevant information and background theory for each material class in the first chapter of the corresponding thesis part. Because our research questions differ according to the material studied, we present our conclusions for each part separately, in the final chapter of each part. In Part I, we improve the prediction accuracy of nonlinear hyperelastic models for elastomers, using new chain stretch relations that connect continuum deformation to the stretching of polymer chains at the molecular level. In Part II, we continue working with nonlinear elasticity but use non-hyperelastic stress–strain relations to capture the phenomenology of active solids undergoing deformations outside the linear regime. We present

a new elastic modulus for terms quadratic in strain, to characterise the activity of isotropic solids, and show how this modulus arises in random networks of active fibres. In Part III, we use linear elasticity augmented with additional fields to formulate topological mechanics in the continuum, by defining integer-valued invariants for both edge and bulk floppy modes. We classify topological floppy modes in the continuum by showing that at least one additional field is required to capture edge modes and that at least two are necessary for Weyl points, or topological bulk modes, which have not previously been captured in continuum theories.

Part I

New constitutive models for elastomers

Chapter 2

Introduction to constitutive models for elastomers

Elastomers, also known as rubber-like materials, consist of polymer chains cross-linked to form a network, and are characterised by the ability to undergo large, reversible deformations [18]. Because this mechanical property is exploited in myriad technological applications, from tyres and belts [19] to soft robotics [20], models for the mechanical response of elastomers impact many areas of modern life. Crucial to modelling the macroscopic mechanical behaviour of elastomers is the framework of nonlinear elasticity [6, 21], which describes load and deformation in terms of stress and strain instead of the precise molecular interactions and configuration of the material. Within this framework, the mechanical properties of a given material are characterised by a constitutive model, which is a mathematical relation between the stress and deformation of the material [21]. Because constitutive models are used in predictive computational methods to inform technological design, such as finite element analysis [19], improving the accuracy of such models is not only theoretically interesting but also practically relevant.

Although work on the theory of rubber elasticity began in the 1930s [22], and many constitutive models for elastomers have since been proposed, Destrade et al. [23] comment that there is still no model able to describe experimental data from the testing of rubbers 'in a satisfying and universal way'. Anssari-Benam et al. [24] concur: 'Developing models that provide an accurate fit to many deformation modes of rubber-like materials while satisfying the *a priori* mathematical and continuum mechanics requirements remains an exigent and a non-trivial task.' Anssari-Benam and Horgan [25] list three criteria that a constitutive model for elastomers should satisfy:

1. The model should be able to fit data from multiple deformation modes of the same material with a single set of parameter values.

2. The model should have as few parameters as possible.

3. The model should be connected to the molecular structure of the material such that its parameters can be interpreted physically.

The first criterion simply requires that the constitutive model be able to capture the material's mechanical response comprehensively. Motivation for the second criterion arises from the fact that models with more parameters are more likely to have numerical issues when fitting to experimental data, such as the non-uniqueness of optimal sets of model parameters [26]. Furthermore, when many parameters are present, it is not clear whether the model's fitting capability comes from the sheer number of parameters that can be varied or from accurate modelling of the physics involved [27]. Constitutive models that satisfy the third criterion are more likely to satisfy the second, because the functional form of the model is informed by the physics of the polymer network, reducing the number of parameters that need to be obtained by fitting to experimental data. Such models also provide insight into how macroscopic mechanical properties depend on the molecular structure of elastomers.

Many well-known constitutive models for elastomers, such as the Mooney-Rivlin [28], Ogden [29] and Gent [30] models, are phenomenological, i.e. their functional forms are selected to satisfy material symmetries and continuum mechanics principles [6] while capturing the empirically observed behaviour. These phenomenological models make little or no reference to the underlying polymer network structure. Although the parameters in these models can be given molecular interpretations, such as in the case of the Gent model [31], this is usually done *a posteriori*. In some cases, a phenomenological approach leads to excellent fits to experimental data, such as in the case of the three-term Ogden model fitted to vulcanised rubber stress–stretch data [29]. However, this model has a relatively large number of parameters (six) that need to be fitted to the data. Therefore, although some phenomenological models might satisfy the first criterion stated by Anssari-Benam and Horgan [25], they often fail to satisfy the second and/or third criteria. Others, such as the Mooney-Rivlin model, have only two parameters but are unable to fit data over the whole range of stretches [32].

Constitutive models that begin with the statistical mechanics of chains in the polymer network and attempt to derive macroscopic behaviour via some effective representation or averaging procedure are often referred to as *molecular-statistical* models. To connect continuum mechanics and polymer physics, these models adopt *micro-macro chain stretch relations*, which express the stretch of a representative polymer chain as a function of the

macroscopic deformation [33]. Well-known examples of such models include the affine network model [34] and the Arruda–Boyce, or 8-chain, model [32], which are able to capture the large stretch behaviour of rubber under uniaxial tension using only two parameters. Although these models satisfy the second and third criteria of Anssari-Benam and Horgan [25], they fail to capture experimental data from multiple deformation modes simultaneously: if calibrated using data from uniaxial tension, they significantly underpredict data from equibiaxial tension [32]. To improve the prediction accuracy of molecular-statistical models, more sophisticated approaches account for constraints placed on polymer chains by entanglements and excluded-volume interactions between neighbouring chains. Examples of such models are the non-affine microsphere model by Miehe et al. [35], with five parameters, and the extended 8-chain model of Dal et al. [36], with three parameters. Although these models more adequately satisfy the first criterion of Anssari-Benam and Horgan [25], this improvement comes at the expense of introducing more model parameters.

In this part of the thesis, we advance the constitutive modelling of elastomers by presenting new molecular-statistical models that satisfy all three criteria in Anssari-Benam and Horgan [25]. With only two parameters, our models achieve better fits to experimental data than existing models that have the same number of parameters. We accomplish this by proposing new micro-macro chain stretch relations based on experimental observations of polymer chain behaviour [37–39] and the Irving–Kirkwood–Noll procedure [40, 41], which systematically obtains continuum physics from statistical mechanics. In the remainder of this chapter, we review previous approaches to constitutive modelling of elastomers and provide the theoretical background for our contributions. We then develop new chain stretch relations in Chapter 3 and compare the predictions of our new models to experimental data from various elastomeric materials in Chapter 4.

2.1 Nonlinear elasticity

Here we review key concepts in nonlinear elasticity and define the notation used in this part of the thesis. The exposition here is based on Gurtin [6] and Gonzalez and Stuart [21], which contain further details.

2.1.1 Vectors and tensors

Let $\mathbf{v}$ be an element of a real, d-dimensional inner product space $\mathcal{V}_d$, with the inner, or dot, product between vectors $\mathbf{u}, \mathbf{v} \in \mathcal{V}_d$ denoted by $\mathbf{u} \cdot \mathbf{v} \in \mathbb{R}$, where $\mathbb{R}$ is the set of real numbers.

In this thesis, the space of vectors $\mathcal{V}_d$ is isomorphic to $\mathbb{R}^d$, the d-fold Cartesian product of $\mathbb{R}$. We consider only $d = 3$ in this part of the thesis, but both $d = 2$ and $d = 3$ in Parts II and III.

A (second-order) tensor $\mathbf{A}$ on $\mathcal{V}_d$ is a linear map from $\mathbf{v} \in \mathcal{V}_d$ to $\mathbf{A}\mathbf{v} \in \mathcal{V}_d$. We denote the set of such linear maps using $\mathtt{L}(\mathcal{V}_d)$. The transpose and inverse of $\mathbf{A}$ are denoted $\mathbf{A}^{\mathrm{T}}$ and $\mathbf{A}^{-1}$, respectively. These operations on $\mathbf{A}$ commute, so their composition can be written $\mathbf{A}^{-\mathrm{T}}$. The set of tensors $\mathtt{L}(\mathcal{V}_d)$ is a vector space with addition and scalar multiplication defined pointwise: for $\mathbf{A}, \mathbf{B} \in \mathtt{L}(\mathbf{V})$ and $\alpha \in \mathbb{R}$, $\mathbf{A} + \mathbf{B}$ and $\alpha\mathbf{A}$ satisfy $(\mathbf{A} + \mathbf{B})\mathbf{v} = \mathbf{A}\mathbf{v} + \mathbf{B}\mathbf{v}$ and $(\alpha\mathbf{A})\mathbf{v} = \alpha(\mathbf{A}\mathbf{v})$ for all $\mathbf{v} \in \mathcal{V}_d$.

The tensor product between vectors $\mathbf{u}, \mathbf{v} \in \mathcal{V}_d$ is defined to be the tensor $\mathbf{u} \otimes \mathbf{v}$ that maps $\mathbf{w}$ to $(\mathbf{w} \cdot \mathbf{v})\mathbf{u}$. The space of tensors $\mathtt{L}(\mathcal{V}_d)$ has an inner product, also known as the double contraction of tensors, defined by $\mathbf{A} : \tilde{\mathbf{A}} = \mathrm{tr}(\mathbf{A}^{\mathrm{T}}\tilde{\mathbf{A}})$ for $\mathbf{A}, \tilde{\mathbf{A}} \in \mathtt{L}(\mathcal{V}_d)$, where tr denotes the trace of a tensor. If $\{\mathbf{e}_j\}_{j=1}^d$ is an orthonormal basis, then $\{\mathbf{e}_j \otimes \mathbf{e}_k\}_{j,k=1}^d$ is an orthonormal basis for $\mathtt{L}(\mathcal{V}_d)$, and the tensor $\mathbf{A}$ can be expressed as $A_{jk}\mathbf{e}_j \otimes \mathbf{e}_k$, where the Einstein summation convention has been used, and the coefficients are given by $A_{jk} = \mathbf{A} : \mathbf{e}_j \otimes \mathbf{e}_k$. The double contraction between two tensors $\mathbf{A}, \tilde{\mathbf{A}}$ can be expressed in terms of tensor components as $\mathbf{A} : \tilde{\mathbf{A}} = A_{ij}\tilde{A}_{ij}$.

2.1.2 Continuum kinematics

The deformation state of a material body $\mathcal{B}$ is represented by a closed, connected subset of Euclidean space that indicates the positions of body's material particles, with the reference state denoted by $\mathcal{B}_0$. Material points are identified by their positions $\mathbf{x}_0 \in \mathcal{B}_0$ in the reference state. In this part of the thesis, we only consider material bodies for which a stress-free state exists, and take this to be the reference state, which we also refer to as the undeformed state.

A deformation is represented by a mapping $\chi : \mathcal{B}_0 \to \mathcal{B}$ that gives the spatial position $\mathbf{x} = \chi(\mathbf{x}_0)$ of each material point in the deformed state $\mathcal{B}$. The deformation gradient is the tensor $\mathbf{F}(\mathbf{x}_0) = \nabla_{\mathbf{x}_0}\chi(\mathbf{x}_0)$, where $\nabla_{\mathbf{x}_0}$ denotes the gradient with respect to coordinates $\mathbf{x}_0$ in the reference state. The deformation χ is required to be one-to-one, i.e. the body cannot penetrate itself, so $\det\mathbf{F}(\mathbf{x}_0) \neq 0$ for all $\mathbf{x}_0 \in \mathcal{B}_0$. Since the undeformed state corresponds to $\det\mathbf{F}(\mathbf{x}_0) = \det\mathbf{I} = 1 > 0$, where $\mathbf{I}$ is the identity tensor, a deformation with $\det\mathbf{F}(\mathbf{x}_0) < 0$ can only be reached continuously by passing through a deformed state corresponding to $\det\mathbf{F}(\mathbf{x}_0) = 0$, showing that $\det\mathbf{F}(\mathbf{x}_0) < 0$ is impossible. Therefore, all deformations must satisfy $\det\mathbf{F}(\mathbf{x}_0) > 0$ for all $\mathbf{x}_0 \in \mathcal{B}_0$. An *incompressible material* cannot undergo changes in volume, so all its deformations satisfy $\det\mathbf{F}(\mathbf{x}_0) = 1$ for all $\mathbf{x}_0 \in \mathcal{B}_0$.

The polar decomposition [42] expresses the deformation gradient $\mathbf{F}$ as the product of a stretch and a rotation:

$$\mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{V}\mathbf{R}, \tag{2.1}$$

where $\mathbf{R}$ is an orthogonal tensor (i.e. $\mathbf{R}^{\mathrm{T}} = \mathbf{R}^{-1}$), and $\mathbf{U}, \mathbf{V}$ are known as the right and left stretch tensors, respectively. Because only the stretching part of the deformation is relevant to elasticity, the following deformation tensors are often useful:

$$\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F} = \mathbf{U}^2, \qquad \mathbf{B} = \mathbf{F}\mathbf{F}^{\mathrm{T}} = \mathbf{V}^2, \tag{2.2}$$

known as the right and left Cauchy–Green deformation tensors, respectively.

2.1.3 Stress measures

Stress tensors characterise the contact forces exerted across internal or external surfaces of a material. The Cauchy stress tensor $\mathbf{T}$ is a measure of force per unit area in the deformed state, whereas the first Piola–Kirchhoff stress tensor $\mathbf{P}$ is a measure of force per unit reference area. These stress measures are related by the pull-back of the area measure:

$$\mathbf{P} = (\det \mathbf{F})\mathbf{T}\mathbf{F}^{-\mathrm{T}}. \tag{2.3}$$

2.1.4 Elastic materials

If the stress at each point of a material is a function of only the deformation gradient at that point, then the material is *Cauchy elastic*, or simply *elastic*. Elastic constitutive models therefore neglect any dependence of stress on the deformation history or rate.

Passive elastic solids, i.e. those from which no work can be extracted in a closed cycle of deformation, are necessarily *hyperelastic* or *Green elastic* [43]. Hyperelastic materials are characterised by a strain energy function W, from which the first Piola–Kirchhoff stress tensor can be obtained by differentiation with respect to the deformation gradient $\mathbf{F}$, i.e.

$$\mathbf{P} = \hat{\mathbf{P}}(\mathbf{F}) = \frac{\mathrm{d}W}{\mathrm{d}\mathbf{F}}(\mathbf{F}). \tag{2.4}$$

A precise definition of the derivative of a scalar function with respect to a tensor argument is given in Appendix I.1. In component form, (2.4) is written $P_{ij} = \hat{P}_{ij}(\mathbf{F}) = \frac{\partial W}{\partial F_{ij}}(\mathbf{F})$. If the material is incompressible, then an arbitrary hydrostatic pressure p_c must be introduced into the Cauchy stress to account for the incompressibility constraint [44, 45]. The value of p_c is determined by the boundary conditions of the problem. In this case, the relationship between the first Piola–Kirchhoff stress tensor and the strain energy function W is given by

$$\mathbf{P} = \hat{\mathbf{P}}(\mathbf{F}) = \frac{\mathrm{d}W}{\mathrm{d}\mathbf{F}}(\mathbf{F}) - p_c\mathbf{F}^{-\mathrm{T}}. \tag{2.5}$$

In this part of the thesis, we model elastomers as incompressible, hyperelastic materials. We also consider only homogeneous and isotropic elastomers, i.e. we assume that material properties are uniform across all points and that there are no preferred directions in the material. These assumptions have been widely adopted to model the elastic behaviour of a broad range of elastomers, as shown in the review article by Dal et al. [7] and the references therein.

2.2 Statistical mechanics of polymer chains

One of the simplest statistical mechanical models for polymer chains is the *freely jointed chain*. This model is widely used in molecular-statistical constitutive models for elastomers, because it accounts for the strain stiffening behaviour seen at large extensions, known as limiting chain extensibility [32, 35, 46]. The freely jointed chain consists of N rigid links of length ℓ_l that are independently oriented and are equally likely to point in any direction. The root-mean-square of the end-to-end distance r of the freely jointed chain in the absence of external fields is [47]

$$\sqrt{\langle r^2 \rangle} = \sqrt{\langle |\mathbf{r}|^2 \rangle} = \sqrt{N}\ell_\mathrm{l}. \tag{2.6}$$

Using equilibrium statistical mechanics, when the number of links N is sufficiently large [48], the force $\mathbf{f}$ applied to the ends of the freely jointed chain to maintain an end-to-end vector $\mathbf{r}$ is parallel to $\mathbf{r}$ with magnitude f_ch given by

$$f_\mathrm{ch}(r) = \frac{k_\mathrm{B}T}{\ell_\mathrm{l}} \mathcal{L}^{-1}\left(\frac{r}{N\ell_\mathrm{l}}\right), \tag{2.7}$$

where k_B is the Boltzmann constant, T is absolute temperature, and $\mathcal{L}^{-1}$ is the inverse of the Langevin function $\mathcal{L}$, which is defined by $\mathcal{L}(x) = \coth x - \frac{1}{x}$. The chain force (2.7) has a singularity $f_\mathrm{ch}(r) \to \infty$ as $r \to N\ell_\mathrm{l}$ [48, 49]. Thus, the strain stiffening of elastomers at large stretches is due to the singularity in chain force as the chain is straightened so that its end-to-end distance r approaches its contour length $N\ell_\mathrm{l}$. This force–length relation can be obtained by differentiating the Helmholtz free energy [32]

$$\psi(r) \approx Nk_\mathrm{B}T\left(\frac{r}{N\ell_\mathrm{l}}\mathcal{L}^{-1}\left(\frac{r}{N\ell_\mathrm{l}}\right) + \ln\frac{\mathcal{L}^{-1}\left(\frac{r}{N\ell_\mathrm{l}}\right)}{\sinh\mathcal{L}^{-1}\left(\frac{r}{N\ell_\mathrm{l}}\right)}\right). \tag{2.8}$$

2.3 Molecular-statistical constitutive models

Models in this class may be broadly categorised into representative cell models and full network models. In the former, the relationship between macroscopic deformation and polymer chain stretch is obtained by considering a cell containing a few polymer chains oriented along different directions, as a simplified representation of the network. Examples of this approach include the three-chain model [50], the four-chain model [51] and the eight-chain, or Arruda–Boyce model [32]. We focus on the Arruda–Boyce model, because it has the most accurate fits to experimental data among the representative cell models [32]. This model represents the polymer network using a cubic cell whose edges are aligned with the principal stretch directions. Each vertex of the cube is connected to a polymer chain, and the eight chains meet at the centre of the cube. The effect of macroscopic deformation is modelled by stretching each edge of the cube by the corresponding principal stretch. Arruda and Boyce [32] show that each chain in this representative cell is stretched by the same amount

$$\lambda_{\mathrm{AB}}(\mathbf{F}) = \sqrt{\frac{1}{3}\mathrm{tr}(\mathbf{F}^{\mathrm{T}}\mathbf{F})}, \tag{2.9}$$

where $\mathrm{tr}(\mathbf{F}^{\mathrm{T}}\mathbf{F})$ is the first strain invariant I_1.

Instead of considering chains along a few representative directions, full network models account for chains along all possible directions in the network. These approaches typically construct the hyperelastic strain energy function W of an elastomeric material by making the following assumptions.

Assumption 1. *The strain energy function W is equal to the total Helmholtz free energy density (i.e. per unit volume) of the polymer network.*

Assumption 2. *There exists a relationship between the orientation of a chain in the network, represented by the unit vector $\hat{\mathbf{r}}$, and the chain stretch λ induced by the macroscopic deformation $\mathbf{F}$, given by a function $\lambda(\hat{\mathbf{r}},\mathbf{F})$, referred to as* the micro-macro chain stretch relation.

Assumption 3. *All polymer chains in the network have the same end-to-end distance $r_0 = \sqrt{N}\ell_1$, from* (2.6)*, when the sample is undeformed.*

Assumption 4. *The total Helmholtz free energy of the network is the sum of the free energies of individual chains. Therefore, the network Helmholtz free energy density can be obtained by averaging the free energy of a single chain over all orientations and multiplying this average by the number density of chains in the network [35].*

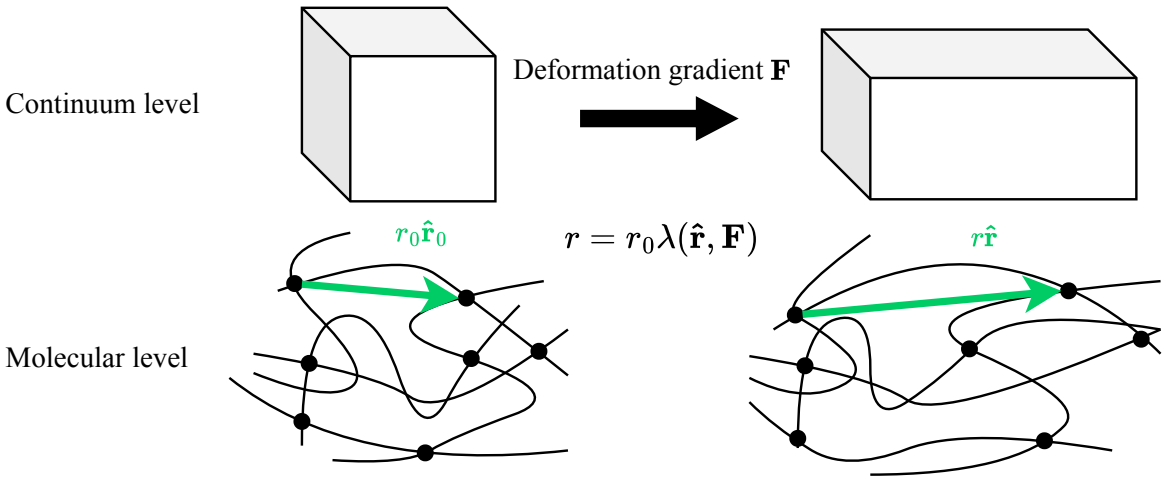


Fig. 2.1 The definition of a chain stretch relation. An elastomeric material is subjected to a deformation specified by the deformation gradient $\mathbf{F}$. At the molecular level, this deformation corresponds to a change in the state of the polymer network, illustrated here using solid curves to represent polymer chains and black circles for cross-links. The end-to-end vector of a representative polymer chain between cross-links is shown using green arrows. The chain stretch relation $\lambda(\hat{\mathbf{r}}, \mathbf{F})$ is the ratio between the deformed and undeformed lengths of this vector, providing a connection between the continuum and molecular levels of description. Although the chain stretch relation is shown here as a function of the chain orientation $\hat{\mathbf{r}}$ in the deformed state, there are chain stretch relations that can be expressed in terms of the orientation $\hat{\mathbf{r}}_0$ in the undeformed state.

Assumption 2 and the definition of the chain stretch relation $\lambda(\hat{\mathbf{r}}, \mathbf{F})$ are illustrated in Fig. 2.1, which shows how a representative chain end-to-end vector is stretched under an imposed continuum-level deformation. This assumption is an idealisation, because the random nature of the polymer network means that individual polymer chains with the same orientation $\hat{\mathbf{r}}$ will not necessarily undergo the same stretch. The chain stretch relation $\lambda(\hat{\mathbf{r}}, \mathbf{F})$ should be interpreted as the mean or representative stretch of chains with the given orientation. This interpretation is valid if the distribution of stretches of these chains is sharply peaked about this representative value.

Averaging over chains in the network is performed by integrating chain orientations $\hat{\mathbf{r}}$ over the unit sphere $\mathcal{S}^2 = \{\mathbf{r} \in \mathcal{V}_3 : |\mathbf{r}| = 1\}$. Therefore, the average value of an orientation-dependent quantity $\xi(\hat{\mathbf{r}})$ is $\int_{\mathcal{S}^2} \xi(\hat{\mathbf{r}})\, \hat{\rho}(\hat{\mathbf{r}}, \mathbf{F})\, \mathrm{d}A(\hat{\mathbf{r}})$, where $\mathrm{d}A(\hat{\mathbf{r}})$ indicates that the integration is performed using the area measure on $\mathcal{S}^2$ with $\hat{\mathbf{r}}$ as the variable of integration and $\hat{\rho}(\hat{\mathbf{r}}, \mathbf{F})$ is the chain orientation distribution function, normalised such that $\int_{\mathcal{S}^2} \hat{\rho}(\hat{\mathbf{r}}, \mathbf{F})\, \mathrm{d}A(\hat{\mathbf{r}}) = 1$. This integration can be performed with respect to chain orientations in either the reference or deformed states, depending on the constitutive model. Some models, like the affine full network model, can be formulated in terms of either state [46], and we will indicate the

meaning of $\hat{\mathbf{r}}$ for each model considered. When integrating over chain orientations in the reference state, $\hat{\rho}(\hat{\mathbf{r}},\mathbf{F})$ is independent of the deformation gradient $\mathbf{F}$. In an isotropic material, $\hat{\rho}(\hat{\mathbf{r}},\mathbf{F}) = (4\pi)^{-1}$ for all $\hat{\mathbf{r}}$ in the reference state.

Using Assumptions 1 – 4 to construct a strain energy function for an elastomer with single chain Helmholtz free energy $\psi(r)$ at chain end-to-end distance r gives [34, 35, 46]

$$W(\mathbf{F}) = n\int_{\mathcal{S}^2} \psi(r_0\lambda(\hat{\mathbf{r}},\mathbf{F}))\hat{\rho}(\hat{\mathbf{r}},\mathbf{F})\,\mathrm{d}A(\hat{\mathbf{r}}), \tag{2.10}$$

where $\lambda(\hat{\mathbf{r}},\mathbf{F})$ is a micro-macro chain stretch relation and n is the number of chains per unit volume. The choice of micro-macro chain stretch relation plays a significant role in determining how well the constitutive model's predictions match experimental data [33, 35], motivating our focus on proposing new chain stretch relations. To set our contributions in context, we first review some previously proposed chain stretch relations for full network models whose strain energy functions take the form (2.10).

The affine chain stretch treats the junctions or nodes where chains are cross-linked together in the polymer network as though they are embedded in the continuum whose deformation is characterised by the deformation gradient $\mathbf{F}$ [33, 52]. Therefore, a chain with end-to-end vector $\mathbf{r}_0$ in the reference state corresponds to an end-to-end vector $\mathbf{F}\mathbf{r}_0$ in the deformed state [46]. The affine stretch of a chain parallel to $\hat{\mathbf{r}}_0$ in the reference state is

$$\lambda^0_{\mathrm{aff}}(\hat{\mathbf{r}}_0,\mathbf{F}) = |\mathbf{F}\hat{\mathbf{r}}_0| = \sqrt{\hat{\mathbf{r}}_0\cdot\mathbf{F}^{\mathrm{T}}\mathbf{F}\hat{\mathbf{r}}_0} = \sqrt{\mathbf{C}:\hat{\mathbf{r}}_0\otimes\hat{\mathbf{r}}_0}, \tag{2.11}$$

where the superscript 0 is used to distinguish this form of the affine stretch from the form expressed in terms of chain orientations $\hat{\mathbf{r}}$ in the deformed state,

$$\lambda_{\mathrm{aff}}(\hat{\mathbf{r}},\mathbf{F}) = 1/\sqrt{\hat{\mathbf{r}}\cdot\mathbf{F}^{-\mathrm{T}}\mathbf{F}^{-1}\hat{\mathbf{r}}} = 1/\sqrt{\mathbf{B}^{-1}:\hat{\mathbf{r}}\otimes\hat{\mathbf{r}}}. \tag{2.12}$$

Although the affine full network model fits the uniaxial tension data of Treloar [53] well, it significantly underpredicts the stress in equibiaxial tension [32].

Although the Arruda–Boyce model is not usually considered a full network model, its strain energy function W can be obtained as a special case of the full network strain energy function (2.10) by setting $\lambda(\hat{\mathbf{r}},\mathbf{F}) = \lambda_{\mathrm{AB}}(\mathbf{F})$ and $\hat{\rho}(\hat{\mathbf{r}},\mathbf{F}) = (4\pi)^{-1}$. Beatty [54] showed that using the orientation-independent chain stretch $\lambda_{\mathrm{AB}}(\mathbf{F})$ is equivalent to assuming that all chains in the network are stretched by the root-mean-square (RMS) of the affine stretch with respect to chain orientations. Although the Arruda–Boyce model fits experimental data better than the affine full network model, it also underpredicts stress in equibiaxial tension [32] if fitted to uniaxial tension data.

Miehe et al. [35] used the constrained minimisation of free energy to arrive at a non-affine chain stretch relation that generalises $\lambda_{\mathrm{AB}}(\mathbf{F})$. Although they began with an orientation-dependent chain stretch $\lambda(\hat{\mathbf{r}},\mathbf{F})$, their constrained minimisation procedure resulted in the orientation-independent stretch

$$\lambda_{\mathrm{M},q}(\mathbf{F}) = \left(\frac{1}{4\pi}\int_{\mathcal{S}^2}\left(\lambda^0_{\mathrm{aff}}(\hat{\mathbf{r}}_0,\mathbf{F})\right)^q \mathrm{d}A(\hat{\mathbf{r}}_0)\right)^{\frac{1}{q}} = \left(\frac{1}{4\pi}\int_{\mathcal{S}^2}\left(\mathbf{C}:\hat{\mathbf{r}}_0\otimes\hat{\mathbf{r}}_0\right)^{\frac{q}{2}} \mathrm{d}A(\hat{\mathbf{r}}_0)\right)^{\frac{1}{q}}, \quad (2.13)$$

where q is a parameter to be fitted to experimental data. This chain stretch is the q-root average of the affine stretch over chain orientations, and reduces to the Arruda–Boyce stretch when $q = 2$. The presence of the additional parameter q results in a three-parameter full network model that achieves better fits to the data than the affine and Arruda–Boyce models.

Amores et al. [33] proposed a non-affine chain stretch relation

$$\lambda_{\mathrm{ANM}}(\hat{\mathbf{r}},\mathbf{F}) = \hat{\mathbf{r}}\cdot\mathbf{U}\hat{\mathbf{r}}, \quad (2.14)$$

where $\mathbf{U}$ is the right stretch tensor defined in Section 2.1. We refer to this chain stretch relation as the ANM chain stretch after the names of the authors of that paper (Amores, Nguyen, Montáns). To justify this chain stretch, Amores et al. [33] argue that the *deformed* polymer network relaxes into a state with an isotropic chain orientation distribution to maximise the configurational entropy of the network. They then construct a first Piola–Kirchhoff stress tensor by summing chain forces and show that the ANM chain stretch leads to consistency between this constructed stress tensor and the stress obtained by differentiating the strain energy function. Zhan et al. [55] also arrive at the ANM chain stretch but use a different motivating argument: the micro-macro chain stretch associated with a given chain orientation $\hat{\mathbf{r}}$ is taken to be the averaged projection of the affine chain stretch onto $\hat{\mathbf{r}}$. Amores et al. [33] show that their chain stretch relation leads to significantly better fits to data from the biaxial deformation of vulcanised rubber in Kawabata et al. [56] than the affine stretch.

Building on the ANM chain stretch, Amores et al. [57] proposed another chain stretch relation to account for the strain stiffening behaviour of elastomers at large stretches, such as seen in the data from Treloar [53]. They argue that the chain orientation distribution remains isotropic only up to moderate stretches and becomes more anisotropic at large stretches. Because the affine stretching of polymer chains leads to anisotropy in the chain orientation distribution, Amores et al. [57] propose a chain stretch that transitions from the ANM chain stretch to the affine chain stretch continuously as strain is increased. We refer to this chain

stretch as the hybrid-ANM stretch. The hybrid-ANM stretch is given by

$$\lambda_{\mathrm{h,ANM}}(\hat{\mathbf{r}},\mathbf{F}) = \left(\hat{\mathbf{r}}\cdot\mathbf{U}^{p(\lambda_{\mathrm{ANM}})}\hat{\mathbf{r}}\right)^{\frac{1}{p(\lambda_{\mathrm{ANM}})}}, \tag{2.15}$$

where

$$p(\lambda_{\mathrm{ANM}}) = 1 + \frac{1}{1+\exp\left(-2\left(\lambda_{\mathrm{ANM}} - \frac{1}{2}\sqrt{N}\right)\right)} \tag{2.16}$$

is a sigmoid function that varies the exponent of $\mathbf{U}$ continuously from $p=1$ to $p=2$. The parameter N corresponds to the number of links in the freely jointed chain (c.f. Section 2.2), and $\sqrt{N}$ is the stretch at which the singularity in chain force occurs in this model, often known as the locking stretch. The hybrid-ANM chain stretch not only gives a good fit to the data from Kawabata et al. [56], but is also able to fit data from the strain stiffening regime in Treloar [53].

Chapter 3

New chain stretch relations

To improve the accuracy of molecular-statistical constitutive models, we propose two new micro-macro chain stretch relations, which we refer to as the exponential and hybrid-exponential chain stretch relations. The justification for these new chain stretch relations begins with a review of experimental observations of polymer chains in deformed samples of vulcanised rubber, in Section 3.1. We then use the Irving–Kirkwood–Noll procedure [40, 41], which systematically derives continuum equations of motion from statistical mechanics, to arrive at the exponential chain stretch. Because this relation is valid only up to moderate stretches, we use an interpolation procedure to obtain the hybrid-exponential stretch, which covers the full range of experimental data from the testing of rubbers.

3.1 Motivation: experimental observations of chain orientation in vulcanised rubber

To determine an appropriate functional form for a micro-macro chain stretch, we review experimental observations of chain orientation in samples of deformed vulcanised rubber, as reported in Toki et al. [37, 38] and Ott et al. [39].

Using *in situ* synchrotron wide-angle X-ray diffraction, Toki et al. [38], following on from Toki et al. [37], quantified the crystalline and amorphous fractions of chains in uniaxially stretched samples of rubber. They used three different types of rubber: (i) sulfur-vulcanised natural rubber, (ii) peroxide-vulcanised natural rubber, and (iii) sulfur-vulcanised synthetic polyisoprene rubber. In loading, for stretches up to 3, more than 95% of the chains were *isotropically oriented* in the amorphous phase, across all three types. For these stretches, the oriented amorphous fraction was below 5% and less than 1% of the chains were in the crystalline phase.

Using proton multiple-quantum nuclear magnetic resonance (NMR), Ott et al. [39] studied chain orientation in uniaxially stretched sulfur-vulcanised natural rubber. To quantify the anisotropy of the chain orientation distribution, they introduced a parameter κ to represent the reorientation of chains towards the stretching direction, where $\kappa = 1$ corresponds to an isotropic orientation distribution. Although the affine chain stretch predicts that κ would decrease from 1 to 0.25 as the uniaxial stretch is increased from 1 to 2.5, the experimental data on Figure 11(A) of Ott et al. [39] showed that κ remained above 0.9 for this stretch range. To explain these results, they suggested that the confining effects of neighbouring polymer chains, often modelled using tube constraints (e.g. Sommer et al. [58] and Mergell and Everaers [59]), lead to the low anisotropy in chain orientations.

In both Toki et al. [38] and Ott et al. [39], experimental observations indicate that polymer chains in deformed samples of rubber remain approximately isotropically oriented even up to moderate stretches. Although a complete theoretical explanation for the these results remains an open problem, we can incorporate these observations into molecular-statistical constitutive models, as demonstrated in the remainder of this chapter. In Chapter 4, we show that the resulting models fit experimental stress–stretch data more accurately than previously proposed models, providing indirect confirmation of the results obtained by Toki et al. [37, 38] and Ott et al. [39].

3.2 Exponential chain stretch relation

To obtain our new chain stretch relation, we begin by constructing a Cauchy stress tensor using the Irving–Kirkwood–Noll procedure [40, 41]. We then require that this constructed stress tensor be consistent with that derived from the full network strain energy function (2.10). This consistency requirement, together with the experimental data on the chain orientation distribution of deformed elastomeric samples reviewed in Section 3.1, leads to a differential equation for a chain stretch relation. We obtain the exponential chain stretch as a solution of this differential equation.

Unlike Amores et al. [33] , who construct a first Piola–Kirchhoff stress tensor by integrating over the contributions that individual chains in the network make to the continuum stress, we construct a Cauchy stress tensor from chain force contributions. Because the first Piola–Kirchhoff stress tensor maps vectors representing oriented areas in the *reference* state to force vectors in the *deformed* state, a relationship between chain orientations in these two states must be assumed in order to construct this stress tensor rigorously. To avoid assuming such a mapping between orientations in the deformed and reference states, we work with the Cauchy stress tensor, which maps oriented areas in the *deformed* state

to forces. Because Amores et al. [33] do not distinguish between reference and deformed states when working with chain orientations, their constitutive model does not satisfy the principle of material frame-indifference, which must be satisfied by physically meaningful constitutive relations [9]. We show how the ANM stretch violates the principle of material frame-indifference in Appendix I.5, where we also modify the ANM stretch to satisfy this principle.

To construct the Cauchy stress tensor, we begin with the procedure proposed by Irving and Kirkwood [40] and reformulated in Noll [41], which we refer to as the Irving–Kirkwood–Noll (IKN) procedure, following Admal and Tadmor [60]. This procedure provides expressions for continuum field quantities such as mass density, momentum flux and stress in terms of the microscopic constituents of a system, and shows that these continuum fields satisfy the mass, momentum and energy balance equations of continuum mechanics.

First, we establish notation adapted from Noll [41]. We consider a system of M particles, whose state at time t is given by the probability density $P(\mathbf{x}_1,\dots,\mathbf{x}_M,\mathbf{v}_1,\dots,\mathbf{v}_M,t)$ over the phase space $\mathbb{R}^{6M}$ of points $(\mathbf{x}_1,\dots,\mathbf{x}_M,\mathbf{v}_1,\dots,\mathbf{v}_M)$ that specify the positions and velocities of the particles. The usual phase space of positions and momenta can be obtained using the transformation $\mathbf{v}_\alpha \mapsto m_\alpha \mathbf{v}_\alpha$, where m_α is the mass of the αth particle. The canonical distribution from the equilibrium statistical mechanics of a system held at fixed temperature is an example of such a probability density. Because the IKN procedure does not require P to have a specific functional form, it also applies to non-equilibrium distributions, provided they satisfy certain regularity conditions [41]. Pointwise continuum fields are expressed as averages over phase space using P. The following notation is adopted for these phase averages: to indicate that a function g of the phase coordinates is averaged over all the coordinates except $\mathbf{x}_\alpha$, which is instead held at a fixed value $\mathbf{x}$, we write

$$\langle g \,|\, \mathbf{x}_\alpha = \mathbf{x}\rangle = \int_{\mathbb{R}^3} \cdots \int_{\mathbb{R}^3} g|_{\mathbf{x}_\alpha=\mathbf{x}} P|_{\mathbf{x}_\alpha=\mathbf{x}} \,\mathrm{d}\mathbf{x}_1 \dots \mathrm{d}\mathbf{x}_{\alpha-1}\mathrm{d}\mathbf{x}_{\alpha+1}\dots\mathrm{d}\mathbf{x}_M \mathrm{d}\mathbf{v}_1 \dots \mathrm{d}\mathbf{v}_M \tag{3.1}$$

for $\alpha = 1,\dots,M$, where $(\cdot)|_{\mathbf{x}_\alpha=\mathbf{x}}$ indicates that the function is evaluated at $\mathbf{x}_\alpha = \mathbf{x}$. This notation extends in an obvious way to fixing more than one variable, e.g. $\langle g \,|\, \mathbf{x}_\alpha = \mathbf{x}, \mathbf{v}_\beta = \mathbf{v}\rangle$. In the notation of Irving and Kirkwood [40], $\langle g \,|\, \mathbf{x}_\alpha = \mathbf{x}\rangle$ is written as $\langle g\delta(\mathbf{x}_\alpha - \mathbf{x}); P\rangle$.

The Cauchy stress tensor from the IKN procedure has two contributions: kinetic stress $\mathbf{T}_K$ and potential stress $\mathbf{T}_V$. Kinetic stress is given by [41, 60]

$$\mathbf{T}_K(\mathbf{x}) = -\sum_{\alpha=1}^{M} m_\alpha \left\langle \left(\mathbf{v}_\alpha^{\mathrm{rel}} \otimes \mathbf{v}_\alpha^{\mathrm{rel}}\right) \middle| \, \mathbf{x}_\alpha = \mathbf{x} \right\rangle,$$

where m_α is the mass of the αth particle and $\mathbf{v}_\alpha^{\text{rel}} = \mathbf{v}_\alpha - \bar{\mathbf{v}}(\mathbf{x})$ is its velocity relative to the pointwise continuum velocity field $\bar{\mathbf{v}}(\mathbf{x})$. The kinetic stress arises from the momentum flux due to the motion of the particles in the system. In a solid, this motion is vibrational. Because we consider only isotropic elastomers, we make the following assumption.

Assumption 5. *The probability that the αth particle is at position* $\mathbf{x}$ *and has velocity* $\mathbf{w}$ *relative to the continuum velocity* $\bar{\mathbf{v}}(\mathbf{x})$, $\langle 1 \,|\, \mathbf{x}_\alpha = \mathbf{x}, \mathbf{v}_\alpha = \mathbf{w} + \bar{\mathbf{v}}(\mathbf{x}) \rangle$, *is independent of the direction of* $\mathbf{w}$.

In Appendix I.3, we show that this assumption implies that $\mathbf{T}_K(\mathbf{x})$ is a (possibly $\mathbf{x}$-dependent) multiple of the identity tensor $\mathbf{I}$.

The potential stress $\mathbf{T}_V$ arises from forces between particles in the system, and is given by [60]

$$\mathbf{T}_V(\mathbf{x}) = \frac{1}{2} \sum_{\substack{\alpha,\beta \\ \alpha \neq \beta}} \int_{\mathbb{R}^3} \int_{s=0}^{1} \left\langle -\mathbf{f}_{\alpha\beta} \,|\, \mathbf{x}_\alpha = \mathbf{x} + s\mathbf{z}, \mathbf{x}_\beta = \mathbf{x} - (1-s)\mathbf{z} \right\rangle \mathrm{d}s \otimes \mathbf{z} \, \mathrm{d}\mathbf{z}, \tag{3.2}$$

where $\mathbf{f}_{\alpha\beta}$ is the force exerted by particle β on particle α. The physical interpretation of this expression is that it is a sum over expectation values of forces between all pairs of particles, where the variable of integration $\mathbf{z}$ represents the vector separating particles α and β. The integration over $\mathbf{z}$ corresponds to taking an expectation value over all possible distances and directions of separation, and integrating over s averages over the position relative to $\mathbf{x}$ of the line segment connecting the particles, with $s = 0$ corresponding to the α end being at $\mathbf{x}$ and $s = 1$ to the β end being at $\mathbf{x}$.

Because we want to express $\mathbf{T}_V$ as a sum over polymer chains rather than over particles, we distinguish between forces along polymer chains in the network, which lead to a stress contribution $\mathbf{T}_\mathrm{c}$, and those that arise from all other inter-atom interactions, such as intra- and inter-chain excluded volume interactions, which contribute to the stress $\mathbf{T}_\mathrm{e}$. To obtain $\mathbf{T}_\mathrm{c}$, we define a chain to be a non-branching sequence of atoms bonded together that connects two cross-links. The chain force is regarded as acting between the two cross-links that are joined by the chain. If the freely jointed chain model is adopted, then this force is given by (2.7). In (3.2), terms corresponding to this contribution have $\mathbf{f}_{\alpha\beta}$ equal to the chain force and $\mathbf{z}$ equal to the chain end-to-end vector $\mathbf{r}$. The other inter-atom interactions that give rise to $\mathbf{T}_\mathrm{e}$ are assumed to be isotropic, so that $\mathbf{T}_\mathrm{e}$ is a multiple of the identity tensor, like the kinetic stress. We summarise this discussion in the following assumption.

Assumption 6. *The potential stress is decomposed such that* $\mathbf{T}_V(\mathbf{x}) = \mathbf{T}_\mathrm{c}(\mathbf{x}) + \mathbf{T}_\mathrm{e}(\mathbf{x})$. *The stress* $\mathbf{T}_\mathrm{c}(\mathbf{x})$ *is a contribution from chain forces given by*

$$\mathbf{T}_\mathrm{c}(\mathbf{x}) = \sum_{j=1}^{N_c} \int_{\mathbb{R}^3} \int_{s=0}^{1} \left\langle -\mathbf{f}_{\alpha(j),\beta(j)} \,|\, \mathbf{x}_{\alpha(j)} = \mathbf{x} + s\mathbf{r}, \mathbf{x}_{\beta(j)} = \mathbf{x} - (1-s)\mathbf{r} \right\rangle \mathrm{d}s \otimes \mathbf{r}\, \mathrm{d}\mathbf{r}, \tag{3.3}$$

where $j = 1, \ldots, N_c$ *indexes the* N_c *polymer chains in the system, and* $\alpha(j), \beta(j)$ *are the indices of particles corresponding to the cross-links connected by the* j*th chain.*
The stress $\mathbf{T}_\mathrm{e}(\mathbf{x})$ *is a contribution from all other interactions (e.g. excluded volume effects), which are assumed to be isotropic, so* $\mathbf{T}_\mathrm{e}(\mathbf{x}) = \xi_\mathrm{e}(\mathbf{x})\mathbf{I}$.

Because the chain force is parallel to its end-to-end vector $\mathbf{r}$ and has magnitude given by a function f_ch of the chain end-to-end length $r = |\mathbf{r}|$, the inner integrand in (3.3) is

$$\begin{aligned}
&\left\langle -\mathbf{f}_{\alpha(j),\beta(j)} \,|\, \mathbf{x}_{\alpha(j)} = \mathbf{x} + s\mathbf{r}, \mathbf{x}_{\beta(j)} = \mathbf{x} - (1-s)\mathbf{r} \right\rangle \\
&= \left\langle -f_\mathrm{ch}(|\mathbf{x}_{\beta(j)} - \mathbf{x}_{\alpha(j)}|) \frac{\mathbf{x}_{\beta(j)} - \mathbf{x}_{\alpha(j)}}{|\mathbf{x}_{\beta(j)} - \mathbf{x}_{\alpha(j)}|} \,\middle|\, \mathbf{x}_{\alpha(j)} = \mathbf{x} + s\mathbf{r}, \mathbf{x}_{\beta(j)} = \mathbf{x} - (1-s)\mathbf{r} \right\rangle \\
&= \left\langle f_\mathrm{ch}(|\mathbf{r}|)\hat{\mathbf{r}} \,|\, \mathbf{x}_{\alpha(j)} = \mathbf{x} + s\mathbf{r}, \mathbf{x}_{\beta(j)} = \mathbf{x} - (1-s)\mathbf{r} \right\rangle \\
&= f_\mathrm{ch}(|\mathbf{r}|)\hat{\mathbf{r}} \left\langle 1 \,|\, \mathbf{x}_{\alpha(j)} = \mathbf{x} + s\mathbf{r}, \mathbf{x}_{\beta(j)} = \mathbf{x} - (1-s)\mathbf{r} \right\rangle,
\end{aligned} \tag{3.4}$$

since $\mathbf{x}_{\beta(j)} - \mathbf{x}_{\alpha(j)} = -\mathbf{r}$. Letting $\sigma_j(\mathbf{x}, \mathbf{r}) = \int_0^1 \left\langle 1 \,|\, \mathbf{x}_{\alpha(j)} = \mathbf{x} + s\mathbf{r}, \mathbf{x}_{\beta(j)} = \mathbf{x} - (1-s)\mathbf{r} \right\rangle \mathrm{d}s$, the expression for $\mathbf{T}_\mathrm{c}$ becomes

$$\mathbf{T}_\mathrm{c}(\mathbf{x}) = \sum_{j=1}^{N_c} \int_{\mathbb{R}^3} f_\mathrm{ch}(|\mathbf{r}|)\hat{\mathbf{r}} \otimes \mathbf{r}\, \sigma_j(\mathbf{x}, \mathbf{r})\, \mathrm{d}\mathbf{r}. \tag{3.5}$$

The quantity $\sigma_j(\mathbf{x}, \mathbf{r})$ is the joint probability density corresponding to the jth polymer chain being positioned with the line segment joining its ends passing through position $\mathbf{x}$ and having end-to-end vector $\mathbf{r}$. We express this joint distribution as $\sigma_j(\mathbf{x}, \mathbf{r}) = \kappa_j(\mathbf{x})\nu_j(\mathbf{r} \,|\, \mathbf{x})$, where $\kappa_j(\mathbf{x})$ is the marginal probability density for the end-to-end line segment of the jth polymer chain passing through position $\mathbf{x}$, and $\nu_j(\mathbf{r} \,|\, \mathbf{x})$ is the conditional probability density for the jth chain having end-to-end vector $\mathbf{r}$ given that its end-to-end line segment passes through position $\mathbf{x}$. Because we consider homogeneous elastomers, $\kappa_j(\mathbf{x})$ is independent of $\mathbf{x}$, giving $\kappa_j(\mathbf{x}) = 1/\Omega$, where Ω is the volume of the system.

We recall that the chain orientations $\hat{\mathbf{r}}$ are distributed with probability density $\hat{\rho}(\hat{\mathbf{r}}, \mathbf{F}(\mathbf{x}))$, and that, by Assumption 2, chain lengths $r = |\mathbf{r}|$ are determined by the chain stretch relation $\tilde{\lambda}(\hat{\mathbf{r}}, \mathbf{F})$. Here we express the chain stretch relation as a function of chain orientations in the deformed state because $\mathbf{r}$ in $\nu_j(\mathbf{r} \,|\, \mathbf{x})$ is an end-to-end vector in the deformed state. Because

$\nu_j(\mathbf{r}\,|\,\mathbf{x})$ assigns very large probability density to regions of $\mathbb{R}^3$ where $r = |\mathbf{r}|$ is close to $r_0\tilde{\lambda}(\hat{\mathbf{r}},\mathbf{F}(\mathbf{x}))$, we make the following assumption.

Assumption 7. *For any function G of the length $|\mathbf{r}|$ and direction $\hat{\mathbf{r}}$ of the chain end-to-end vector $\mathbf{r}$, the integral of G over $\mathbb{R}^3$ can be approximated by an integral over $\mathcal{S}^2$, i.e.*

$$\int_{\mathbb{R}^3} G(|\mathbf{r}|,\hat{\mathbf{r}})\nu_j(\mathbf{r}|\,\mathbf{x})\,\mathrm{d}\mathbf{r} \approx \int_{\mathcal{S}^2} G(r_0\tilde{\lambda}(\hat{\mathbf{r}},\mathbf{F}(\mathbf{x})),\hat{\mathbf{r}})\hat{\rho}(\hat{\mathbf{r}},\mathbf{F}(\mathbf{x}))\,\mathrm{d}A(\hat{\mathbf{r}}),$$

where the chain length $|\mathbf{r}|$ has been replaced by $r_0\tilde{\lambda}(\hat{\mathbf{r}},\mathbf{F}(\mathbf{x}))$ and the integrand is weighted by the chain orientation distribution $\hat{\rho}(\hat{\mathbf{r}},\mathbf{F}(\mathbf{x}))$ instead of the probability density $\nu_j(\mathbf{r}|\,\mathbf{x})$ over chain end-to-end vectors.

With this assumption, $\mathbf{T}_\mathrm{c}$ can be approximated as follows:

$$\begin{aligned}\mathbf{T}_\mathrm{c}(\mathbf{x}) &= \frac{1}{\Omega}\sum_{j=1}^{N_c}\int_{\mathbb{R}^3} f_\mathrm{ch}(|\mathbf{r}|)\hat{\mathbf{r}}\otimes\mathbf{r}\,\nu_j(\mathbf{r}|\,\mathbf{x})\,\mathrm{d}\mathbf{r}\\ &\approx \frac{1}{\Omega}\sum_{j=1}^{N_c}\int_{\mathcal{S}^2}(\hat{\mathbf{r}}\otimes\hat{\mathbf{r}})r_0\tilde{\lambda}(\hat{\mathbf{r}},\mathbf{F}(\mathbf{x}))f_\mathrm{ch}(r_0\tilde{\lambda}(\hat{\mathbf{r}},\mathbf{F}(\mathbf{x})))\hat{\rho}(\hat{\mathbf{r}},\mathbf{F}(\mathbf{x}))\,\mathrm{d}A(\hat{\mathbf{r}}),\end{aligned} \tag{3.6}$$

Because the material is isotropic, $\tilde{\lambda}(\hat{\mathbf{r}},\mathbf{F})$ can be expressed as a function of $\mathbf{B}$ instead of $\mathbf{F}$. Material isotropy implies

$$\tilde{\lambda}(\hat{\mathbf{r}},\mathbf{F}\mathbf{Q}) = \tilde{\lambda}(\hat{\mathbf{r}},\mathbf{F}), \tag{3.7}$$

for all proper orthogonal tensors $\mathbf{Q}$, i.e. tensors that satisfy $\mathbf{Q}^\mathrm{T} = \mathbf{Q}^{-1}$ and $\det\mathbf{Q} = 1$. Physically, this condition corresponds to invariance under all rigid-body rotations. Using the polar decomposition $\mathbf{F} = \mathbf{V}\mathbf{R}$ (cf. Section 2.1), we obtain $\tilde{\lambda}(\hat{\mathbf{r}},\mathbf{F}) = \tilde{\lambda}(\hat{\mathbf{r}},\mathbf{V})$, by setting $\mathbf{Q} = \mathbf{R}^{-1}$. Because $\mathbf{B} = \mathbf{F}\mathbf{F}^\mathrm{T} = \mathbf{V}^2$, we define $\lambda(\hat{\mathbf{r}},\mathbf{B}) = \tilde{\lambda}(\hat{\mathbf{r}},\sqrt{\mathbf{B}})$, leading to

$$\lambda(\hat{\mathbf{r}},\mathbf{B}) = \tilde{\lambda}(\hat{\mathbf{r}},\mathbf{F}). \tag{3.8}$$

We show in Appendix I.2 that the derivatives of these two representations of the chain stretch relation satisfy

$$\frac{\partial\tilde{\lambda}}{\partial\mathbf{F}} = 2\frac{\partial\lambda}{\partial\mathbf{B}}\mathbf{F}. \tag{3.9}$$

Although the use of material isotropy to derive representations for physical quantities in terms of $\mathbf{B}$ is well-established in nonlinear elasticity (e.g. Gurtin [6] and Gonzalez and Stuart [21]), this argument is not usually applied to micro-macro chain stretch relations, as done here.

Using (3.8) in (3.6) gives

$$\mathbf{T}_{\mathrm{c}} = nr_0 \int_{\mathcal{S}^2} \lambda(\hat{\mathbf{r}},\mathbf{B}) f_{\mathrm{ch}}(r_0\lambda(\hat{\mathbf{r}},\mathbf{B}))\hat{\rho}(\hat{\mathbf{r}},\mathbf{F})\hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\,\mathrm{d}A(\hat{\mathbf{r}}), \tag{3.10}$$

where $n = N_c/\Omega$ is the number density of polymer chains in the system and the explicit dependence on $\mathbf{x}$ has been dropped because the expression depends on position $\mathbf{x}$ only through $\mathbf{F}$ and $\mathbf{B}$. We note that the passage from statistical mechanics to continuum mechanics in Admal and Tadmor [60] does not stop with the results of the IKN procedure, but incorporates a spatial averaging step to account for the limited spatial resolution of any measurement made on the macroscale. In the steps leading to (3.10), spatial averaging on the stress fields obtained from the IKN procedure is implicitly included in the assumptions made. In particular, $\mathbf{T}_{\mathrm{c}}$ is smoothed out spatially by assuming that $\kappa_j(\mathbf{x}) = 1/\Omega$ and that the orientations of the vectors $\mathbf{r}$ can be described using the chain orientation distribution $\hat{\rho}(\hat{\mathbf{r}},\mathbf{F}(\mathbf{x}))$. Any $\mathbf{x}$-dependence of the spherical contribution to Cauchy stress (from the kinetic stress $\mathbf{T}_K$ and the non-chain force contribution $\mathbf{T}_{\mathrm{e}}$ to the potential stress) can also be smoothed out by spatial averaging.

Putting together the kinetic and potential contributions to the Cauchy stress tensor gives

$$\mathbf{T} = \mathbf{T}_{\mathrm{c}} + \chi\mathbf{I} = nr_0 \int_{\mathcal{S}^2} \lambda(\hat{\mathbf{r}},\mathbf{B}) f_{\mathrm{ch}}(r_0\lambda(\hat{\mathbf{r}},\mathbf{B}))\hat{\rho}(\hat{\mathbf{r}},\mathbf{F})\hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\,\mathrm{d}A(\hat{\mathbf{r}}) + \chi\mathbf{I}, \tag{3.11}$$

where $\chi\mathbf{I}$ is the isotropic contribution from the kinetic stress $\mathbf{T}_K$ and the non-chain force contribution $\mathbf{T}_{\mathrm{e}}$ to the potential stress.

Besides obtaining a stress tensor using the IKN procedure, we can also obtain a stress tensor by differentiating the strain energy function (2.10), which averages over the Helmholtz free energy contributions of chains in the network. Using (2.5) and (2.3), we obtain a Cauchy stress tensor

$$\mathbf{T} = \frac{\mathrm{d}W}{\mathrm{d}\mathbf{F}}(\mathbf{F})\mathbf{F}^{\mathrm{T}} - p_{\mathrm{c}}\mathbf{I} = \frac{\mathrm{d}}{\mathrm{d}\mathbf{F}}\left(n\int_{\mathcal{S}^2}\psi(r_0\lambda(\hat{\mathbf{r}},\mathbf{F}))\hat{\rho}(\hat{\mathbf{r}},\mathbf{F})\,\mathrm{d}A(\hat{\mathbf{r}})\right)\mathbf{F}^{\mathrm{T}} - p_{\mathrm{c}}\mathbf{I} \tag{3.12}$$

from (2.10). Because both the chain force contribution $\mathbf{T}_{\mathrm{c}}$ in (3.11) and the hyperelastic contribution $\frac{\mathrm{d}W}{\mathrm{d}\mathbf{F}}(\mathbf{F})\mathbf{F}^{\mathrm{T}}$ in (3.12) involve averaging over the polymer chains in the network, we make the the natural assumption that they are equal.

Assumption 8. *The chain force contribution* $\mathbf{T}_{\mathrm{c}}$ *to Cauchy stress is equal to* $\frac{\mathrm{d}W}{\mathrm{d}\mathbf{F}}(\mathbf{F})\mathbf{F}^{\mathrm{T}}$.

This assumption implies that $\chi = -p_{\mathrm{c}}$ in (3.11), where p_{c} is the pressure due to the incompressibility constraint. In Appendix I.4, we show that the Cauchy stress obtained from the affine chain stretch (2.11) using (3.12) is equal to (3.11) with $\mathbf{T}_{\mathrm{c}} = \frac{\mathrm{d}W}{\mathrm{d}\mathbf{F}}(\mathbf{F})\mathbf{F}^{\mathrm{T}}$ and $\chi = -p_{\mathrm{c}}$, demonstrating the validity of Assumption 8 in this case. Because (3.11) and (3.12) are valid

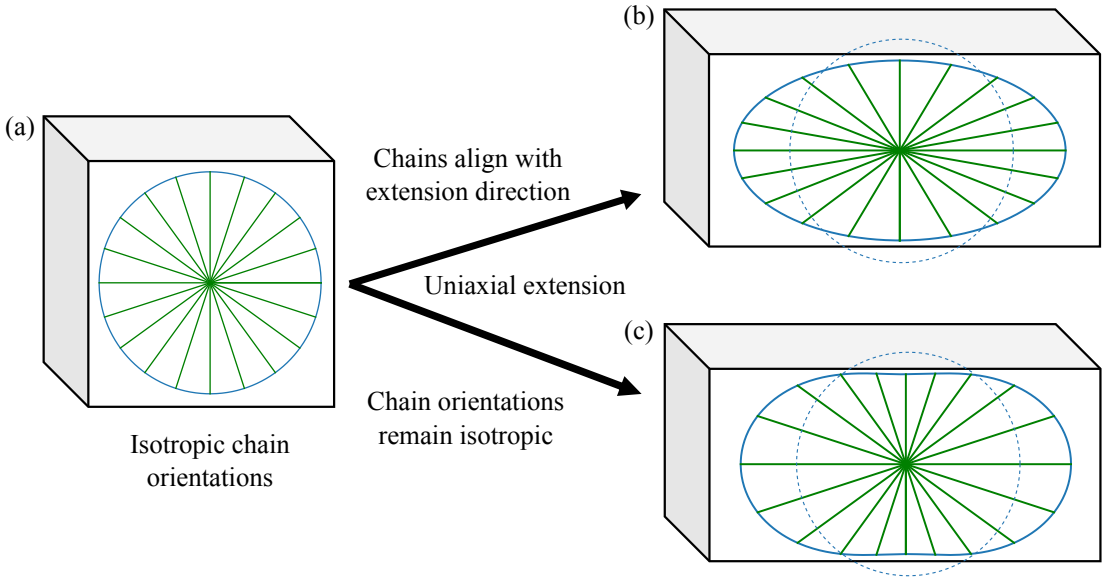


Fig. 3.1 Different assumptions used to model the effect of deformation on the chain orientation distribution. The effect of uniaxial extension on chain lengths and orientations is illustrated using green lines to represent chain end-to-end vectors in a plane containing the extension direction. Importantly, these lines do not show the actual spatial arrangement of chains in the network (cf. Fig. 2.1), but instead represent the statistics of chain lengths and orientations within the network. (a) The undeformed state of an elastomeric material has an isotropic chain orientation distribution and all chain-end-to end vectors are assumed to have equal lengths. (b, c) Under uniaxial extension, the changes in length of the chain end-to-end vectors are represented by the orientation-dependent changes in length of the green lines, and the deformation of the blue circle in (a) into the shapes shown in (b) and (c). In (b), the chain end-to-end vectors transform affinely with the deformation, so they align with the extension direction, as shown by the changes in the orientations of the green lines. In contrast, (c) illustrates Assumption 9, under which chains remain isotropically oriented.

for any chain stretch relation, this assumption is a reasonable generalisation to all chain stretch relations of the result proven in Appendix I.4 for the affine chain stretch. Assumption 8 is the basis for a *consistency requirement* (CR) between the two stress expressions (3.11) and (3.12), which we use to obtain a new chain stretch relation.

To apply CR, we make the following assumption, based on the experimental evidence summarised in Section 3.1.

Assumption 9. *The chain orientation distribution* $\hat{\rho}(\hat{\mathbf{r}},\mathbf{F})$ *for the deformed state is approximately isotropic, i.e.* $\hat{\rho}(\hat{\mathbf{r}},\mathbf{F}) = (4\pi)^{-1}$*, up to moderate stretches.*

In Fig. 3.1, we illustrate this assumption and contrast it with the affine assumption, under which chain orientations in the deformed state are anisotropic. If Ott et al. [39] are correct in

their suggestion that the isotropy of the chain orientation distribution in the deformed state is due to the tube-like confining effects of neighbouring chains, then Assumption 9 can be regarded as capturing these excluded-volume interactions. Unlike the proposals of Miehe et al. [35] and Dal et al. [36], which capture these effects using tube constraint terms, using Assumption 9 does not introduce additional fitting parameters. Although Amores et al. [33] also make this assumption, they justify it using an argument based on entropy maximisation rather than basing it on experimental evidence. Because Assumption 9 is an idealisation, the chain stretch relation proposed here is not expected to provide accurate predictions for large stretches. In the next section, we relax this assumption by introducing the hybrid-exponential chain stretch to account for large stretches. By Assumption 9,

$$\mathbf{T}_{\mathrm{c}} = \frac{nr_0}{4\pi}\int_{\mathcal{S}^2} f_{\mathrm{ch}}(r_0\lambda(\hat{\mathbf{r}},\mathbf{B}))\lambda(\hat{\mathbf{r}},\mathbf{B})\,\hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\,\mathrm{d}A(\hat{\mathbf{r}}) = \frac{nr_0}{4\pi}\int_{\mathcal{S}^2} \psi'(r_0\lambda(\hat{\mathbf{r}},\mathbf{B}))\lambda(\hat{\mathbf{r}},\mathbf{B})\hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\,\mathrm{d}A(\hat{\mathbf{r}}), \tag{3.13}$$

where we have used the relationship between chain force and single chain Helmholtz free energy $f(r) = \psi'(r)$. Next, applying Assumption 9 to (2.10) by taking the variable of integration $\hat{\mathbf{r}}$ to be a chain orientation in the *deformed* state and setting $\hat{\rho}(\hat{\mathbf{r}},\mathbf{F}(\mathbf{x})) = (4\pi)^{-1}$ gives

$$\frac{\mathrm{d}W}{\mathrm{d}\mathbf{F}}(\mathbf{F})\mathbf{F}^{\mathrm{T}} = 2\frac{nr_0}{4\pi}\left(\int_{\mathcal{S}^2}\psi'(r_0\lambda(\hat{\mathbf{r}},\mathbf{B}))\frac{\partial\lambda}{\partial\mathbf{B}}(\hat{\mathbf{r}},\mathbf{B})\,\mathrm{d}A(\hat{\mathbf{r}})\right)\mathbf{B}, \tag{3.14}$$

where (3.9) has been used to replace the derivative with respect to $\mathbf{F}$ by one with respect to $\mathbf{B}$. Assumption 8 leads immediately to

$$\frac{nr_0}{4\pi}\int_{\mathcal{S}^2}\psi'(r_0\lambda(\hat{\mathbf{r}},\mathbf{B}))\left(2\left(\frac{\partial\lambda}{\partial\mathbf{B}}(\hat{\mathbf{r}},\mathbf{B})\right)\mathbf{B} - \lambda(\hat{\mathbf{r}},\mathbf{B})\hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right)\mathrm{d}A(\hat{\mathbf{r}}) = 0, \tag{3.15}$$

which expresses CR as a mathematical relationship satisfied by a chain stretch relation $\lambda(\hat{\mathbf{r}},\mathbf{B})$.

We now seek a chain stretch relation that satisfies the consistency requirement (3.15). To begin, we express (3.15) in terms of the principal directions (eigenvectors) of $\mathbf{B}$, which form an orthonormal basis. In general, $\mathbf{B} = \sum_{i=1}^{3}\mu_i\mathbf{n}_i\otimes\mathbf{n}_i$, where $\{\mu_i = \lambda_i^2\}_{i=1}^3$ and $\{\mathbf{n}_i\}_{i=1}^3$ are the eigenvalues and unit-normalised eigenvectors, respectively, of $\mathbf{B}$, and $\{\lambda_i\}_{i=1}^3$ are the principal stretches of the deformation. Since $\{\mathbf{n}_i\otimes\mathbf{n}_j\}_{i,j=1}^3$ is an orthonormal basis for tensors in the deformed state, the diagonal components of (3.15) in this basis are

$$\frac{nr_0}{4\pi}\int_{\mathcal{S}^2}\psi'(r_0\lambda)\left(2\frac{\partial\lambda}{\partial\mathbf{B}}\mathbf{B} - \lambda\hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right) : \mathbf{n}_i\otimes\mathbf{n}_i\,\mathrm{d}A(\hat{\mathbf{r}}) = 0, \tag{3.16}$$

for $i = 1,2,3$, where $\lambda = \lambda(\hat{\mathbf{r}},\mathbf{B})$ is written for brevity. By the chain rule, $\frac{\partial\lambda}{\partial\mu_i} = \frac{\partial\lambda}{\partial\mathbf{B}} : \mathbf{n}_i \otimes \mathbf{n}_i$, so these equations are equivalent to

$$\int_{\mathcal{S}^2} \psi'(r_0\lambda(\hat{\mathbf{r}},\mathbf{B})) \left(2\frac{\partial\lambda}{\partial\mu_i}\mu_i - \lambda(\hat{\mathbf{r}}\cdot\mathbf{n}_i)^2\right) \mathrm{d}A(\hat{\mathbf{r}}) = 0 \tag{3.17}$$

for $i = 1,2,3$. A sufficient (though not in general necessary) condition for these equations to be satisfied is for the term in parentheses to vanish, i.e.

$$2\frac{1}{\lambda(\hat{\mathbf{r}},\mathbf{B})}\frac{\partial}{\partial\mu_i}\lambda(\hat{\mathbf{r}},\mathbf{B}) = \frac{1}{\mu_i}(\hat{\mathbf{r}}\cdot\mathbf{n}_i)^2 \quad \text{for } i = 1,2,3,$$

which can be integrated to give $\ln\lambda(\hat{\mathbf{r}},\mathbf{B}) = \sum_{i=1}^3 (\ln\lambda_i)(\hat{\mathbf{r}}\cdot\mathbf{n}_i)^2$, where the constant of integration is zero because $\lambda = 1$ when $\lambda_1 = \lambda_2 = \lambda_3 = 1$.

We therefore obtain a new micro-macro chain stretch relation that satisfies the diagonal components of CR, given in (3.16):

$$\lambda(\hat{\mathbf{r}},\mathbf{B}) = \exp\left(\sum_{i=1}^3 (\ln\lambda_i)(\hat{\mathbf{r}}\cdot\mathbf{n}_i)^2\right) = \exp\left(\ln\mathbf{V} : \hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right), \tag{3.18}$$

where $\ln\mathbf{V} = \sum_{i=1}^3 \ln\lambda_i\,\mathbf{n}_i\otimes\mathbf{n}_i$ is known as the logarithmic or natural strain [61]. An alternative representation of this chain stretch relation is the product form

$$\lambda(\hat{\mathbf{r}},\mathbf{B}) = \lambda_1^{(\hat{\mathbf{r}}\cdot\mathbf{n}_1)^2}\lambda_2^{(\hat{\mathbf{r}}\cdot\mathbf{n}_2)^2}\lambda_3^{(\hat{\mathbf{r}}\cdot\mathbf{n}_3)^2}. \tag{3.19}$$

This chain stretch relation, like the affine and ANM relations, predicts that a chain aligned with a principal stretch direction $\mathbf{n}_i$ experiences a stretch equal to the corresponding principal stretch. We refer to the chain stretch in (3.18) as the *exponential chain stretch relation* and write it as λ_{exp}.

To show that λ_{exp} satisfies the full tensor equation (3.15) for CR, we check that the off-diagonal components of the left-hand side of (3.15) are zero by showing that both $\mathbf{T}_{\mathrm{c}}$ and $\frac{\mathrm{d}W}{\mathrm{d}\mathbf{F}}(\mathbf{F})\mathbf{F}^{\mathrm{T}}$ have zero off-diagonal components. By (3.18), λ_{exp} is unchanged when the signs of one or more components of $\hat{\mathbf{r}}$ in the orthonormal basis $\{\mathbf{n}_i\}_{i=1}^3$ are flipped, i.e. when $\hat{\mathbf{r}}\cdot\mathbf{n}_i$ is replaced by $-\hat{\mathbf{r}}\cdot\mathbf{n}_i$. Any physically reasonable chain stretch relation must have this property because each $\mathbf{n}_i$ can always be replaced by $-\mathbf{n}_i$ in the eigendecomposition of $\mathbf{B}$. We use this symmetry to show that the off-diagonal components of the Cauchy stress tensor $\mathbf{T}_{\mathrm{c}}$ are zero. For fixed i, j with $i \neq j$, we decompose the unit sphere $\mathcal{S}^2$ into the two hemispheres $\mathcal{S}_j^+ = \{\hat{\mathbf{r}}\in\mathcal{S}^2 : \hat{\mathbf{r}}\cdot\mathbf{n}_j \geq 0\}$ and $\mathcal{S}_j^- = \{\hat{\mathbf{r}}\in\mathcal{S}^2 : \hat{\mathbf{r}}\cdot\mathbf{n}_j < 0\}$, which satisfy $\mathcal{S}^2 = \mathcal{S}_j^+ \bigcup \mathcal{S}_j^-$. The

off-diagonal components of $\mathbf{T}_\mathrm{c}$ are then given by

$$\mathbf{T}_\mathrm{c} : \mathbf{n}_i \otimes \mathbf{n}_j = \frac{nr_0}{4\pi} \int_{\mathcal{S}_j^+} \psi'(r_0 \lambda_\mathrm{exp}(\hat{\mathbf{r}}, \mathbf{B})) \lambda_\mathrm{exp}(\hat{\mathbf{r}}, \mathbf{B}) (\hat{\mathbf{r}} \cdot \mathbf{n}_i)(\hat{\mathbf{r}} \cdot \mathbf{n}_j) \, \mathrm{d}A(\hat{\mathbf{r}}) + \frac{nr_0}{4\pi} \int_{\mathcal{S}_j^-} \psi'(r_0 \lambda_\mathrm{exp}(\hat{\mathbf{r}}, \mathbf{B})) \lambda_\mathrm{exp}(\hat{\mathbf{r}}, \mathbf{B}) (\hat{\mathbf{r}} \cdot \mathbf{n}_i)(\hat{\mathbf{r}} \cdot \mathbf{n}_j) \, \mathrm{d}A(\hat{\mathbf{r}}).$$

In each integral term, the sign of $\hat{\mathbf{r}} \cdot \mathbf{n}_j$ does not change over the domain of integration, but for every $\hat{\mathbf{r}}$ with $\hat{\mathbf{r}} \cdot \mathbf{n}_i > 0$, there is a corresponding $\hat{\mathbf{r}}$ in the domain of integration with $\hat{\mathbf{r}} \cdot \mathbf{n}_i < 0$ that gives the same value of $\psi' \lambda_\mathrm{exp}$. Therefore, each integral evaluates to zero because of the cancellation due to this symmetry, giving $\mathbf{T}_\mathrm{c} : \mathbf{n}_i \otimes \mathbf{n}_j = 0$, as claimed.

To show that $\frac{\mathrm{d}W}{\mathrm{d}\mathbf{F}}(\mathbf{F})\mathbf{F}^\mathrm{T}$ has zero off-diagonal components, some preliminary properties of λ_exp are necessary. Firstly, this chain stretch satisfies material isotropy because replacing $\mathbf{F}$ by $\mathbf{FQ}$ for any proper orthogonal tensor $\mathbf{Q}$ does not change $\mathbf{B} = \mathbf{F}\mathbf{F}^\mathrm{T}$. Secondly, the principle of material frame-indifference [6], which requires that $W(\mathbf{QF}) = W(\mathbf{F})$ for all proper orthogonal tensors $\mathbf{Q}$, is satisfied because replacing $\mathbf{F}$ by $\mathbf{QF}$ is equivalent to replacing $\mathbf{B}$ by $\mathbf{QBQ}^\mathrm{T}$, so

$$\lambda_\mathrm{exp}(\hat{\mathbf{r}}, \mathbf{QBQ}^\mathrm{T}) = \lambda_\mathrm{exp}(\mathbf{Q}^\mathrm{T}\hat{\mathbf{r}}, \mathbf{B}) \tag{3.20}$$

for all proper orthogonal tensors $\mathbf{Q}$. Therefore, the strain energy function $W(\mathbf{F})$ satisfies

$$W(\mathbf{QF}) = \frac{n}{4\pi} \int_{\mathcal{S}^2} \psi(r_0 \lambda_\mathrm{exp}(\hat{\mathbf{r}}, \mathbf{QBQ}^\mathrm{T})) \, \mathrm{d}A(\hat{\mathbf{r}}) = \frac{n}{4\pi} \int_{\mathcal{S}^2} \psi(r_0 \lambda_\mathrm{exp}(\mathbf{Q}^\mathrm{T}\hat{\mathbf{r}}, \mathbf{B})) \, \mathrm{d}A(\hat{\mathbf{r}}) = W(\mathbf{F}) \tag{3.21}$$

for all proper orthogonal tensors $\mathbf{Q}$, because the Jacobian of the change of variables from $\hat{\mathbf{r}}$ to $\mathbf{Q}^\mathrm{T}\hat{\mathbf{r}}$ is unity. Because the strain energy function $W(\mathbf{F})$ satisfies material frame-indifference and material isotropy, the theorem on page 170 of Gurtin [6] implies that the Cauchy stress derived from this strain energy function has the same eigenbasis as $\mathbf{B}$, so $\frac{\mathrm{d}W}{\mathrm{d}\mathbf{F}}(\mathbf{F})\mathbf{F}^\mathrm{T}$ is diagonal in this eigenbasis.

Because CR as expressed in (3.15) is equivalent to $\mathbf{T}_\mathrm{c} = \frac{\mathrm{d}W}{\mathrm{d}\mathbf{F}}(\mathbf{F})\mathbf{F}^\mathrm{T}$, we have shown that the exponential chain stretch satisfies the full tensor equation (3.15) by showing that both $\mathbf{T}_\mathrm{c}$ and $\frac{\mathrm{d}W}{\mathrm{d}\mathbf{F}}(\mathbf{F})\mathbf{F}^\mathrm{T}$ have zero off-diagonal components. We remark that there could be other chain stretch relations that satisfy the CR (3.15), because we have not shown that chain stretch solutions to this equation are unique.

We conclude this section by summarising the reasoning used to arrive at the exponential chain stretch. Any attempt to propose a chain stretch relation inevitably involves making simplifying assumptions about how chains in a polymer network behave in response to network deformation, including Assumptions 1 – 4 in Section 2.3. Previous approaches also make additional, stronger assumptions about this response, such as assuming that chains

deform affinely [46] or that a representative cell with eight chains captures the effective network response [32]. In contrast, using the IKN procedure enables us to replace strong assumptions, like affine stretching, with weaker ones (Assumptions 5 – 8), to derive a form for the chain force contributions to Cauchy stress (3.10). These assumptions are natural when working with full-network models that have strain energy function (2.10). Assumptions 5 – 8 are also general enough to be consistent with more than one chain stretch relation, as shown in Appendix I.4, where we demonstrate that the affine chain stretch leads to a hyperelastic stress $\frac{\mathrm{d}W}{\mathrm{d}\mathbf{F}}(\mathbf{F})\mathbf{F}^{\mathrm{T}}$ that is a special case of the expression (3.10) from the IKN procedure. Therefore, it is Assumption 9 about the orientation distribution of polymer chains in the deformed state that distinguishes the exponential chain stretch from other chain stretch relations compatible with (3.10). This assumption is not only supported by the experimental evidence reviewed in Section 3.1, but is also confirmed by the comparison to experimental stress–stretch data in Chapter 4.

3.3 Hybrid-exponential chain stretch relation

Because Assumption 9 is not valid in highly deformed elastomeric samples, the exponential chain stretch (3.18) is not expected to give good agreement with experimental data that includes strain stiffening behaviour at large stretches. Indeed, Chapter 4 shows that the exponential chain stretch performs well for low to moderate stretches but does not exhibit the correct form of singularity at large stretches. To obtain a micro-macro chain stretch relation for the whole range of physically attainable stretches, we propose a hybrid chain stretch relation λ_{hyb} to account for the anisotropy of the chain orientation distribution at large stretches.

Under the affine assumption, the chain orientations in the reference and deformed states are related by $\hat{\mathbf{r}} = \mathbf{F}\hat{\mathbf{r}}_0/|\mathbf{F}\hat{\mathbf{r}}_0|$, where the subscript 0 indicates orientations in the reference state. Using this relation, Wu and Giessen [46] showed that the chain orientation distribution in the deformed state is

$$\hat{\rho}(\hat{\mathbf{r}},\mathbf{B}) = \frac{1}{4\pi}\lambda_{\mathrm{aff}}(\hat{\mathbf{r}},\mathbf{B})^3 = \frac{1}{4\pi}\left(\mathbf{B}^{-1} : \hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\right)^{-\frac{3}{2}}, \tag{3.22}$$

if the chains are isotropically oriented in the reference state. Therefore, under the affine assumption, the reorientation of chains is characterised by alignment with directions of larger affine stretch, leading to a highly anisotropic chain stretch distribution in the deformed state. We use this property to construct a hybrid chain stretch that behaves like the exponential chain stretch λ_{exp} at low stretches and like the affine chain stretch at large stretches. By

interpolating between the two chain stretch behaviours, the hybrid chain stretch accounts for the approximately isotropic chain orientation distribution at low to moderate stretches, and the increasing anisotropy at large stretches.

Although Amores et al. [57] also adopt the approach of interpolating between two chain stretches, they interpolate between the ANM and affine stretches, and use a different interpolation function from the one we adopt here. We refer to their hybrid stretch as the hybrid-ANM chain stretch to distinguish it from our hybrid-exponential stretch. Their approach also modifies the freely jointed chain Helmholtz free energy (2.8) to obtain a chain force $f = f_0 + f_{\text{FJC}}$, where f_{FJC} is the force exerted by a freely jointed chain and f_0 is a constant term introduced to account for the internal energy contributions of the material at low stretches. The use of f_0 results in an additional parameter to be fitted to experimental data, so the fits in Amores et al. [57] involve a three-parameter constitutive model. In contrast, we use the standard freely-jointed chain model, so the models we consider involve only two parameters.

A direct approach to capturing chain orientation anisotropy at large stretches would involve interpolating between the isotropic chain orientation distribution $\hat{\rho}(\hat{\mathbf{r}}, \mathbf{B}) = (4\pi)^{-1}$ and the anisotropic distribution (3.22). However, interpolating between distributions is difficult, so we take a different approach that only requires varying one interpolation parameter. To introduce this approach, we observe that the full network strain energy function for the affine chain stretch can be expressed as an integral over chain orientations in either the reference or deformed states [46]:

$$\begin{aligned} W_{\text{aff},0}(\mathbf{C}) &= \frac{n}{4\pi} \int_{\mathcal{S}^2} \psi\left(r_0 \sqrt{\mathbf{C} : \hat{\mathbf{r}}_0 \otimes \hat{\mathbf{r}}_0}\right) \mathrm{d}A(\hat{\mathbf{r}}_0), \\ W_{\text{aff}}(\mathbf{B}) &= n \int_{\mathcal{S}^2} \psi\left(r_0 / \sqrt{\mathbf{B}^{-1} : \hat{\mathbf{r}} \otimes \hat{\mathbf{r}}}\right) \hat{\rho}(\hat{\mathbf{r}}, \mathbf{B}) \, \mathrm{d}A(\hat{\mathbf{r}}) \\ &= \frac{n}{4\pi} \int_{\mathcal{S}^2} \psi\left(r_0 / \sqrt{\mathbf{B}^{-1} : \hat{\mathbf{r}} \otimes \hat{\mathbf{r}}}\right) \left(\mathbf{B}^{-1} : \hat{\mathbf{r}} \otimes \hat{\mathbf{r}}\right)^{-\frac{3}{2}} \mathrm{d}A(\hat{\mathbf{r}}), \end{aligned} \tag{3.23}$$

which are related by the change of variables $\hat{\mathbf{r}} = \mathbf{F}\hat{\mathbf{r}}_0 / |\mathbf{F}\hat{\mathbf{r}}_0|$ and where $\hat{\rho}(\hat{\mathbf{r}}, \mathbf{B})$ is given by (3.22). The anisotropic chain orientation distribution resulting from the affine stretch only needs to be included explicitly when integrating over chain orientations $\hat{\mathbf{r}}$ in the deformed state, as given in the expression for $W_{\text{aff}}(\mathbf{B})$. In contrast, the integral over chain orientations $\hat{\mathbf{r}}_0$ in the reference state for $W_{\text{aff},0}(\mathbf{C})$ captures the same behaviour without using the chain orientation distribution (3.22) because the chains are isotropically oriented in the reference state. This property of the expression for $W_{\text{aff},0}(\mathbf{C})$ suggests an interpolation procedure, which we develop by performing the change of variables $\hat{\mathbf{r}} = \mathbf{R}\hat{\mathbf{r}}_0$ in (3.23), where $\mathbf{R}$ is the rotation tensor from the polar decomposition $\mathbf{F} = \mathbf{V}\mathbf{R}$. Because $\mathbf{B} = \mathbf{R}\mathbf{C}\mathbf{R}^{\mathrm{T}}$, and the Jacobian

of the transformation is unity, we obtain

$$W_{\text{aff},0}(\mathbf{C}) = \frac{n}{4\pi}\int_{\mathcal{S}^2} \psi(r_0\sqrt{\mathbf{B}:\hat{\mathbf{r}}\otimes\hat{\mathbf{r}}})\,\mathrm{d}A(\hat{\mathbf{r}}) = \frac{n}{4\pi}\int_{\mathcal{S}^2} \psi(r_0\lambda_{\text{paf}}(\hat{\mathbf{r}},\mathbf{B}))\,\mathrm{d}A(\hat{\mathbf{r}}), \tag{3.24}$$

where $\lambda_{\text{paf}}(\hat{\mathbf{r}},\mathbf{B}) = \sqrt{\mathbf{B}:\hat{\mathbf{r}}\otimes\hat{\mathbf{r}}}$. Because $\mathbf{R}$ maps vectors from the reference state to the deformed state, the integration in (3.24) is over unit vectors in the deformed state. The chain stretch $\lambda_{\text{paf}}(\hat{\mathbf{r}},\mathbf{B})$ is a pseudo-affine stretch defined on chain orientations $\hat{\mathbf{r}}$ in the deformed state that yields the same strain energy function as the affine chain stretch without requiring (3.22). We emphasise that the unit vectors $\hat{\mathbf{r}}$ in $\lambda_{\text{paf}}(\hat{\mathbf{r}},\mathbf{B})$ should not be interpreted as physically meaningful chain orientations in the deformed state. Rather, the construction of the pseudo-affine stretch is a mathematical manipulation performed to obtain a form for the affine full network strain energy that can be used in an interpolation procedure.

Comparing the strain energy function corresponding to the exponential chain stretch

$$W_{\text{exp}}(\mathbf{B}) = \frac{n}{4\pi}\int_{\mathcal{S}^2} \psi(r_0\lambda_{\text{exp}}(\hat{\mathbf{r}},\mathbf{B}))\,\mathrm{d}A(\hat{\mathbf{r}})$$

to that corresponding to the affine chain stretch in (3.24) points to an interpolation procedure between $\lambda_{\text{exp}}(\hat{\mathbf{r}},\mathbf{B})$ and $\lambda_{\text{paf}}(\hat{\mathbf{r}},\mathbf{B})$. To perform this interpolation, we consider the one-parameter family of chain stretches of the form

$$\lambda_p(\hat{\mathbf{r}},\mathbf{B}) = (\mathbf{B}^{\frac{p}{2}}:\hat{\mathbf{r}}\otimes\hat{\mathbf{r}})^{\frac{1}{p}} = (\mathbf{V}^p:\hat{\mathbf{r}}\otimes\hat{\mathbf{r}})^{\frac{1}{p}}, \tag{3.25}$$

where $p \in \mathbb{R}\setminus\{0\}$. This family of chain stretches is similar to that in Akbari et al. [62] but with $\mathbf{U}$ replaced by $\mathbf{V}$. Setting $p=1$ gives the modified ANM chain stretch in Appendix I.5, and $p=2$ corresponds to the pseudo-affine chain stretch. Using $\mathbf{V}^p = \sum_{i=1}^3 \lambda_i^p \mathbf{n}_i\otimes\mathbf{n}_i$, the limit of λ_p as $p\to 0$ can be obtained as follows:

$$\begin{aligned}\lim_{p\to 0}\lambda_p &= \lim_{p\to 0}\exp(\ln\lambda_p) = \exp\left(\lim_{p\to 0}\ln\lambda_p\right) = \exp\left(\lim_{p\to 0}\frac{\ln\left(\sum_{i=1}^3\lambda_i^p(\hat{\mathbf{r}}\cdot\mathbf{n}_i)^2\right)}{p}\right)\\ &= \exp\left(\lim_{p\to 0}\frac{\left(\sum_{i=1}^3\lambda_i^p\ln\lambda_i(\hat{\mathbf{r}}\cdot\mathbf{n}_i)^2\right)/\left(\sum_{i=1}^3\lambda_i^p(\hat{\mathbf{r}}\cdot\mathbf{n}_i)^2\right)}{1}\right) = \lambda_{\text{exp}},\end{aligned}$$

where in the second-last equality, L'Hôpital's rule [63] is used to evaluate the limit. This result means that λ_p can be defined to be λ_{exp} at $p=0$. Therefore, continuous interpolation between the exponential and pseudo-affine chain stretches can be achieved by continuously varying p from 0 to 2 as the deformation increases.

To choose the dependence of p on the applied deformation, we refer to the experimental results of Toki et al. [38] and Ott et al. [39] discussed in Section 3.1, which show that network chains are mostly isotropically oriented even up to moderate stretches ($\lambda \approx 2.5$). These observations are consistent with our results in Chapter 4, which show that the theoretical predictions based on the exponential stretch match stress–stretch data well up to $\lambda \approx 3$. Therefore, we choose an interpolation function that ensures that p remains close to zero even up to moderate deformations. A suitable candidate for this interpolation function is the 'flat function' [64]

$$f_{\text{flat}}(x) = \begin{cases} \exp\left(-\frac{1}{x^2}\right) & x \neq 0 \\ 0 & x = 0 \end{cases} \tag{3.26}$$

which is continuous at all $x \in \mathbb{R}$. This function has derivatives to all orders, i.e. it is smooth, and is 'flat' because all its derivatives vanish at $x = 0$. As $x \to \infty$, $f_{\text{flat}}(x) \to 1$. In view of these properties, we define a function that links p to chain stretch by

$$\hat{p}(\lambda) = 2 f_{\text{flat}}\left(\phi \frac{\lambda - 1}{\sqrt{N} - 1}\right), \tag{3.27}$$

where N is the number of links in the freely jointed chain representing the polymer molecules (cf. Section 2.2) and ϕ sets the scale over which $\hat{p}$ varies. Using (3.27), we define

$$\lambda_{\text{hyb}}(\hat{\mathbf{r}}, \mathbf{B}) = \lambda_p(\hat{\mathbf{r}}, \mathbf{B})\big|_{p=\hat{p}(\lambda_{\text{exp}}(\hat{\mathbf{r}}, \mathbf{B}))}, \tag{3.28}$$

which ties the value of p in (3.25) to the deformation via the exponential chain stretch. This chain stretch approximates the exponential chain stretch at low stretches and the pseudo-affine chain stretch at large stretches.

The constant ϕ in (3.27) can either be treated as an additional fitting parameter in the constitutive model, or be fixed as part of the model. To complete the definition of a hybrid chain stretch that does not introduce additional fitting parameters, we fix the value of ϕ in the model. Because $\sqrt{N}$ is the locking stretch of the freely jointed chain (the stretch at which the singularity in chain force occurs, cf. Section 2.2), we take $\frac{1}{2}\sqrt{N}$ to be a moderate stretch at which the network chain behaviour is still relatively well-described by the exponential chain stretch relation. We then require that $\hat{p}(\frac{1}{2}\sqrt{N})$ be in the lower half of the interval $0 \leq p \leq 2$, i.e. $0 < \hat{p}(\frac{1}{2}\sqrt{N}) \leq 1$, so that the hybrid stretch remains close to the exponential stretch at $\lambda = \frac{1}{2}\sqrt{N}$. Choosing $\hat{p}(\frac{1}{2}\sqrt{N}) = \frac{1}{2}$ leads to

$$\phi = \frac{\sqrt{2}(\sqrt{N} - 1)}{\sqrt{\ln 2}(\sqrt{N} - 2)}. \tag{3.29}$$

Although this choice of $\hat{p}(\frac{1}{2}\sqrt{N})$ seems arbitrary, we show in Appendix I.6 that other choices such as $\hat{p}(\frac{1}{2}\sqrt{N}) = \frac{1}{4}$ or $\hat{p}(\frac{1}{2}\sqrt{N}) = \frac{3}{4}$ do not significantly affect the fits to experimental data, justifying our choice not to treat ϕ as a fitting parameter. For completeness, we treat ϕ as an additional parameter to be fitted to the data in Appendix I.7, and find that the resulting constitutive model fits experimental data better than other three-parameter models in the literature. We refer to the chain stretch defined in (3.28) with ϕ given in (3.29) as the *hybrid-exponential chain stretch relation*.

Chapter 4

Comparison with experimental data

We evaluate the new chain stretch relations proposed in Chapter 3 by comparing the predictions of constitutive models that use these chain stretch relations to experimental data. These models are compared with previously proposed constitutive models to show that we achieve improvements in fitting performance without introducing additional parameters. After listing the constitutive models that we consider in this chapter, we investigate how the choice of chain stretch relation qualitatively affects model predictions of the Poynting effect in simple shear, a well-known effect in nonlinear elasticity [65]. We then perform quantitative comparisons of fitting performance using experimental data from the testing of vulcanised natural rubber [53, 56], silicone rubber [66], silane-modified polymer adhesive [67], and polyacrylamide gel [68].

4.1 Constitutive models under consideration

The constitutive models that we compare in this chapter are all two-parameter models based on the chain stretch relations in Section 2.3 and Chapter 3. To determine whether our new chain stretch relations lead to improvements in prediction accuracy, we do not need to test all two-parameter molecular statistical models. Instead, we rely on Dal et al. [7], an extensive review of hyperelastic constitutive models for rubber-like materials, which found that the Arruda–Boyce model was the most accurate two-parameter molecular-statistical model with respect to the dataset of Treloar [53]. Because Dal et al. [7] did not include the very recent ANM and hybrid-ANM chain stretches in their review, we include these chain stretches in our comparison here, alongside the Arruda–Boyce and affine models. Each chain stretch

relation is used in a hyperelastic constitutive model with strain energy function

$$W(\mathbf{F}) = \frac{n}{4\pi}\int_{\mathcal{S}^2} \psi(r_0\lambda(\hat{\mathbf{r}},\mathbf{F}))\,\mathrm{d}A(\hat{\mathbf{r}}), \tag{4.1}$$

where $r_0 = \sqrt{N}\ell_1$ (c.f. Section 2.2), the single chain Helmholtz free energy ψ is that of the freely jointed chain given in (2.8), and the chain stretch relation $\lambda(\hat{\mathbf{r}},\mathbf{F})$ is one of the following.

Affine:

$$\lambda^0_{\mathrm{aff}}(\hat{\mathbf{r}},\mathbf{F}) = \sqrt{\hat{\mathbf{r}}\cdot(\mathbf{F}^{\mathrm{T}}\mathbf{F})\hat{\mathbf{r}}} = \sqrt{\hat{\mathbf{r}}\cdot\mathbf{C}\hat{\mathbf{r}}} \tag{4.2}$$

Arruda–Boyce:

$$\lambda_{\mathrm{AB}}(\hat{\mathbf{r}},\mathbf{F}) = \sqrt{\frac{1}{3}\mathrm{tr}(\mathbf{F}^{\mathrm{T}}\mathbf{F})} \tag{4.3}$$

ANM:

$$\lambda_{\mathrm{ANM}}(\hat{\mathbf{r}},\mathbf{F}) = \hat{\mathbf{r}}\cdot\left(\sqrt{\mathbf{F}^{\mathrm{T}}\mathbf{F}}\right)\hat{\mathbf{r}} = \hat{\mathbf{r}}\cdot\mathbf{U}\hat{\mathbf{r}} \tag{4.4}$$

Hybrid-ANM:

$$\begin{aligned}
\lambda_{\mathrm{h,ANM}}(\hat{\mathbf{r}},\mathbf{F}) &= \left(\hat{\mathbf{r}}\cdot\mathbf{U}^{p(\lambda_{\mathrm{ANM}})}\hat{\mathbf{r}}\right)^{\frac{1}{p(\lambda_{\mathrm{ANM}})}} \\
p(\lambda_{\mathrm{ANM}}) &= 1 + \frac{1}{1+\exp\left(-2\left(\lambda_{\mathrm{ANM}} - \frac{1}{2}\sqrt{N}\right)\right)}
\end{aligned} \tag{4.5}$$

Exponential:

$$\lambda_{\mathrm{exp}}(\hat{\mathbf{r}},\mathbf{F}) = \exp\left(\hat{\mathbf{r}}\cdot\ln\left(\sqrt{\mathbf{F}\mathbf{F}^{\mathrm{T}}}\right)\hat{\mathbf{r}}\right) = \exp\left(\hat{\mathbf{r}}\cdot\ln\left(\mathbf{V}\right)\hat{\mathbf{r}}\right) \tag{4.6}$$

Hybrid-Exponential:

$$\begin{aligned}
\lambda_{\mathrm{hyb}}(\hat{\mathbf{r}},\mathbf{F}) &= \begin{cases} \left(\hat{\mathbf{r}}\cdot\mathbf{V}^{\hat{p}(\lambda_{\mathrm{exp}})}\hat{\mathbf{r}}\right)^{\frac{1}{\hat{p}(\lambda_{\mathrm{exp}})}} & \hat{p}(\lambda_{\mathrm{exp}}) \neq 0 \\ \lambda_{\mathrm{exp}}(\hat{\mathbf{r}},\mathbf{F}) & \hat{p}(\lambda_{\mathrm{exp}}) = 0 \end{cases} \\
\hat{p}(\lambda_{\mathrm{exp}}) &= \begin{cases} 2\exp\left(-\left(\frac{\sqrt{2}(\lambda_{\mathrm{exp}}-1)}{\sqrt{\ln 2}(\sqrt{N}-2)}\right)^{-2}\right) & \lambda_{\mathrm{exp}} \neq 1 \\ 0 & \lambda_{\mathrm{exp}} = 1 \end{cases}
\end{aligned} \tag{4.7}$$

In the above equations, we have used $\hat{\mathbf{r}}$ to denote a chain orientation, whether it is in the reference or deformed state. The definitions of the chain stretch relations in Section 2.3 and Chapter 3 indicate whether $\hat{\mathbf{r}}$ is an orientation in the reference or deformed state in each case. The Arruda–Boyce stretch is independent of $\hat{\mathbf{r}}$, so its corresponding strain energy function W does not require integrating over chain orientations.

To obtain predictions of stress, we use (2.5) and (2.7), which gives the first Piola–Kirchhoff stress

$$\begin{aligned}\mathbf{P} &= \frac{\partial W}{\partial \mathbf{F}} - p_c \mathbf{F}^{-\mathrm{T}} = \frac{n r_0}{4\pi} \int_{\mathcal{S}^2} \psi'(r_0 \lambda(\hat{\mathbf{r}}, \mathbf{F})) \frac{\partial \lambda}{\partial \mathbf{F}}(\hat{\mathbf{r}}, \mathbf{F}) \, \mathrm{d}A(\hat{\mathbf{r}}) - p_c \mathbf{F}^{-\mathrm{T}} \\ &= \frac{\bar{\mu}\sqrt{N}}{4\pi} \int_{\mathcal{S}^2} \mathcal{L}^{-1}\left(\frac{\lambda(\hat{\mathbf{r}}, \mathbf{F})}{\sqrt{N}}\right) \frac{\partial \lambda}{\partial \mathbf{F}}(\hat{\mathbf{r}}, \mathbf{F}) \, \mathrm{d}A(\hat{\mathbf{r}}) - p_c \mathbf{F}^{-\mathrm{T}},\end{aligned} \tag{4.8}$$

where $\bar{\mu} = n k_{\mathrm{B}} T$. In (4.8), p_{c} is a hydrostatic pressure determined by the boundary conditions, and $\bar{\mu}, N$ are model parameters to be fitted to experimental data. The two parameters have the following physical interpretations: $\bar{\mu}$ is proportional to the infinitesimal shear modulus μ of the material, and N is the number of links in the freely jointed chain that represents polymer chains in the elastomeric network (c.f. Section 2.2). In this chapter, we only compare chain stretch relations that correspond to two-parameter models, so we do not consider the chain stretch (2.13) introduced by Miehe et al. [35], which uses an additional fitting parameter, leading to a model with at least three parameters.

In the numerical implementation of these models, we approximate the inverse Langevin function using the rational function approximation given in equation (14) of Kröger [69]. Integration over the unit sphere is performed using the algorithm in Bažant and Oh [70], which approximates the integral as a weighted sum. The 21-point version of this algorithm was used by Miehe et al. [35] in their implementation of the non-affine microsphere model. To choose an appropriate number of integration points, we investigated convergence of the numerical results by computing model predictions for several deformation modes using Lebedev quadrature [71–74], which is exact for spherical harmonics up to degree 131. Because this analysis showed that the 61-point integration scheme in Bažant and Oh [70] accurately approximated the relevant integrals, we use this scheme to compute model predictions, as a good balance between accuracy and efficiency.

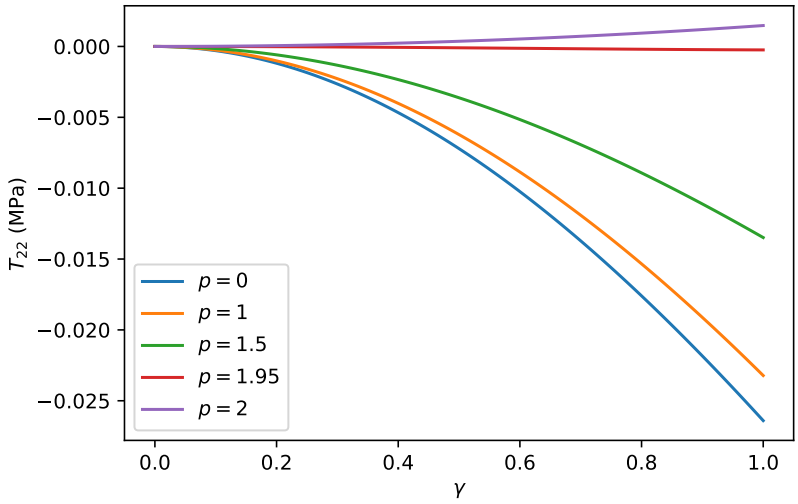


Fig. 4.1 Model predictions of normal stress T_{22} in simple shear against shear strain γ for different values of p in the chain stretch λ_p, defined in (3.25). This chain stretch is used in a strain energy function of the form (4.1). Model parameters are fixed at $\bar{\mu} = 0.35$ MPa, $N = 60$.

4.2 Qualitative prediction of the Poynting effect in simple shear

Simple shear involves homogeneously deforming a material such that the deformation gradient has the matrix representation

$$[\mathbf{F}] = \begin{bmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.9}$$

with respect to an orthonormal basis $\{\mathbf{e}_j\}_{j=1}^3$, where γ is the shear strain. A cube undergoing this deformation would be deformed such that one of the two faces normal to $\mathbf{e}_2$ is displaced relative to the other by γ. Unlike in isotropic linear elasticity, where only the shear stress T_{12} is non-zero in simple shear, this deformation cannot be achieved in isotropic nonlinear elasticity by applying shear stresses alone [6]. In general, the normal stress T_{22} would be non-zero in simple shear, corresponding to the normal forces exerted on the faces being sheared (those orthogonal to $\mathbf{e}_2$) to maintain a constant separation distance between these faces. In most elastomers, this normal stress is compressive, $T_{22} < 0$, indicating a tendency, known as the classical Poynting effect, of the material to expand in the direction perpendicular to the sheared faces [65]. In contrast, biopolymer gels formed by semiflexible polymers have

been shown to exhibit the reverse Poynting effect, where the normal stress is tensile, $T_{22} > 0$, indicating that the material tends to contract in the direction normal to the sheared faces [68].

Here we investigate how the choice of chain stretch relation $\lambda(\hat{\mathbf{r}}, \mathbf{F})$ in (4.1) leads to different signs for the normal stress T_{22} in simple shear. The Arruda–Boyce model is a generalised neo-Hookean model, i.e. its strain energy function is a function of the first strain invariant $I_1 = \mathrm{tr}\,\mathbf{B}$ only. Because it does not depend on the second strain invariant $I_2 = \frac{1}{2}((\mathrm{tr}\,\mathbf{B})^2 - \mathrm{tr}(\mathbf{B}^2))$, it predicts that the normal stress T_{22} in simple shear is identically zero [24]. Therefore, the Arruda–Boyce model is unable to account for any type of Poynting effect in simple shear.

Because the behaviours of the exponential, ANM and affine chain stretches can be reproduced by the chain stretch λ_p in (3.25) by taking $p \to 0$, $p = 1$ and $p = 2$, respectively, we plot the normal stress T_{22} against shear strain γ in Fig. 4.1, for different values of p. The hybrid-exponential stretch is not shown because it predicts the same normal stresses T_{22} as the exponential stretch, for the range of γ considered. Fig. 4.1 shows that the normal stress T_{22} at each $\gamma \in [0,1]$ increases from negative to positive as p is increased from 0 to 2, showing a transition from the classical Poynting effect to the reverse Poynting effect. The transition point occurs at $p \approx 1.95$, where the predicted normal stress $T_{22} \approx 0$ over the range of γ considered. Therefore, the ANM and exponential stretches give qualitative agreement with the experimental observations of the classical Poynting effect in elastomers [65]. In Section 4.3, we compare the predictions of these chain stretches to quantitative data from polyacrylamide gel in simple shear, from Janmey et al. [68].

4.3 Quantitative comparison of constitutive models

We assess the constitutive models listed in Section 4.1 by comparing their predictions to experimental data from Treloar [53] and Kawabata et al. [56] (vulcanised natural rubber), Meunier et al. [66] (silicone rubber), Moreira and Nunes [67] (silane-modified polymer adhesive), and Janmey et al. [68] (polyacrylamide gel). The first three of these datasets, which are widely used to assess the accuracy of constitutive models for rubber-like materials, contain data from samples subjected to uniaxial tension, pure shear and equibiaxial tension. Here we review these deformations, which we express in an orthonormal basis $\{\mathbf{e}_j\}_{j=1}^3$. In these deformations, the principal directions, which are the eigenvectors of the left Cauchy–Green deformation tensor $\mathbf{B} = \mathbf{F}\mathbf{F}^{\mathrm{T}}$, are parallel to the basis $\{\mathbf{e}_j\}_{j=1}^3$.

1. Uniaxial tension

   The sample is stretched by $\lambda_1 = \lambda$ along the $\mathbf{e}_1$-direction. Because the sample is incompressible, the other principal stretches are $\lambda_2 = \lambda_3 = \lambda^{-\frac{1}{2}}$. Stress is applied only

to faces normal to $\mathbf{e}_1$, so the transverse faces are stress-free, implying in particular that $T_{22} = T_{33} = 0$.

2. Pure shear
   The sample is stretched by $\lambda_1 = \lambda$ along the $\mathbf{e}_1$-direction while its length along the $\mathbf{e}_2$-direction is fixed, giving $\lambda_2 = 1$. Incompressibility implies that $\lambda_3 = \lambda^{-1}$, and faces normal to $\mathbf{e}_3$ are stress-free so $T_{33} = 0$.

3. Equibiaxial tension
   The sample is stretched equally along the $\mathbf{e}_1$- and $\mathbf{e}_2$-directions, giving $\lambda_1 = \lambda_2 = \lambda$. Incompressibility implies $\lambda_3 = \lambda^{-2}$, and faces normal to $\mathbf{e}_3$ are stress-free so $T_{33} = 0$.

In each case, the boundary condition $T_{33} = 0$ is used to determine the value of p_c in (4.8).

Unlike the three deformation modes above, simple shear, defined in (4.9), corresponds to deformation gradient $\mathbf{F}$ whose polar decomposition $\mathbf{F} = \mathbf{VR}$ has $\mathbf{R} \neq \mathbf{I}$. Simple shear has principal stretches $\{\lambda, 1, \frac{1}{\lambda}\}$, where $\lambda = \sqrt{\frac{1}{2}(2+\gamma^2+\gamma\sqrt{4+\gamma^2})}$. Therefore, simple shear is kinematically equivalent to pure shear combined with a non-trivial rotation $\mathbf{R}$. To assess the ability of the constitutive models in Section 4.1 to fit simple shear data, we use the datasets of Moreira and Nunes [67] and Janmey et al. [68]. Moreira and Nunes [67] performed separate experimental tests for simple shear and pure shear on a silane-modified polymer adhesive (Flextec® FT 101). Janmey et al. [68] measure shear and normal stresses for both semiflexible biopolymer gels and polyacrylamide gel, which is a flexible polymer gel, using a strain-controlled rheometer. Because our focus is on elastomers formed from flexible polymer chains, we only use the polyacrylamide dataset.

4.3.1 Fitting procedure for model parameters

To fit constitutive models to experimental data, we minimise the residual function equal to the sum of squared differences between model predictions and experimentally measured stress,

$$R(\{\lambda_i, y_i\}_i) = \sum_i \left(y(\lambda_i, \boldsymbol{\varepsilon}) - y_i\right)^2,$$

where $y(\lambda, \boldsymbol{\varepsilon})$ is the stress prediction at stretch λ, with model parameters $\boldsymbol{\varepsilon} = [\bar{\mu}, N]$, and the sum is over all data points (λ_i, y_i), indexed by i. The summation is over data points from all deformation modes in a given dataset, so the model parameters are calibrated by fitting simultaneously to all deformation modes. This simultaneous fitting makes it less likely that the optimal set of model parameters is non-unique [27] and means that the model's capability to account for multiple deformation modes using a single set of parameters can

be assessed [75]. Although using the sum of squared relative errors, $\sum_i \left(\frac{y(\lambda_i,\boldsymbol{\varepsilon})-y_i}{y_i}\right)^2$, as the residual function has the advantage of making the fitting process independent of the stress measure used (e.g. first Piola–Kirchhoff or Cauchy) [23], Anssari-Benam et al. [24] found that minimising squared relative errors gave more weight to data points at smaller stresses, resulting in larger relative errors at moderate to large stretches. Therefore, following Anssari-Benam et al. [24], we choose to minimise the sum over squared absolute errors to assess fitting performance over the whole range of stretches in the experimental data.

4.3.2 Measuring the prediction accuracy of constitutive models

We use the mean squared error (MSE) of each constitutive model, both overall and for individual deformation modes, as a quantitative measure for its prediction accuracy. The MSE is given by $\frac{1}{M}\sum_{i=1}^{M}\left(y(\lambda_i,\boldsymbol{\varepsilon}^*)-y_i\right)^2$, where $\{(\lambda_i,y_i)\}_{i=1}^{M}$ are pairs of stretch and stress values from experimental data and $y(\lambda_i,\boldsymbol{\varepsilon}^*)$ is the theoretical prediction of stress for the stretch λ_i, using the optimum set of parameter values $\boldsymbol{\varepsilon}^*$ obtained through the fitting procedure. The overall MSE is calculated using all the data points from all deformation modes in the dataset, whereas the MSE for each deformation mode is calculated using only the data points corresponding to that deformation mode.

4.3.3 Results and discussion

We fit each constitutive model to experimental data, and compare the fitting performances of the models to each dataset.

Vulcanised natural rubber data from Kawabata et al. [56]

Model predictions are plotted alongside experimental data from Kawabata et al. [56] in Fig. 4.2, and the corresponding model parameter values and MSE are given in Table 4.1. The stresses plotted on Fig. 4.2 correspond to the P_{11} component of the first Piola–Kirchhoff stress $\mathbf{P}$ for all three deformation modes, as well as the transverse stress P_{22} for pure shear. This transverse stress corresponds to the normal force along $\mathbf{e}_2$ required to maintain $\lambda_2 = 1$. For some models, the fitting procedure results in very large values of N, the parameter that controls the location of the stress singularity (c.f. Section 2.2), because this dataset does not include large stretches where significant strain-stiffening occurs.

	$\bar{\mu}$ (MPa)	N	Mean squared error (MPa^2)				
			Overall	Uni.	Sh.	Sh. (T)	Eq.
Affine	0.346	1.48×10^4	0.00789	0.00361	0.00381	0.0186	0.00517
A–B	0.346	2.17×10^4	0.00787	0.00361	0.00381	0.0186	0.00516
ANM	0.501	2.00×10^4	0.00116	0.000356	0.00143	0.00202	**0.000821**
H-ANM	0.501	1.64×10^4	0.00116	**0.000355**	0.00142	0.00202	0.000822
Exp.	0.894	32.6	0.00104†	0.00153	**0.000385**	**0.000814**	0.00147
H-Exp.	0.935	68.2	**0.00104**†	0.00163	0.000408	0.00112	0.000985

Table 4.1 Model parameters and mean squared error (MSE) when fitted to experimental data from Kawabata et al. [56]. The smallest MSE value in each column is displayed in bold font. Abbreviations: "A–B" is "Arruda–Boyce"; "H-" is "Hybrid-"; "Exp." is "Exponential"; "Uni." is "Uniaxial Tension"; "Sh." is "Pure Shear"; "Sh. (T)" is "Pure Shear (Transverse component)"; "Eq." is "Equibiaxial Tension".
†To 3 significant figures these two MSE values are equal, but the hybrid-exponential value is marginally smaller when more significant figures are included (0.001038 vs 0.001039).

The results in Fig. 4.2 and Table 4.1 show that the affine and Arruda–Boyce models have similar predictions, and have the worst fits to the data. Because this dataset does not include large stretches in the strain-stiffening regime, the hybrid-ANM and hybrid-exponential models have similar predictions to the ANM and exponential models, respectively. Overall, the exponential and hybrid-exponential stretches have the most accurate fits to data, followed by the ANM and hybrid-ANM models. Although the difference in overall MSE between the exponential and ANM stretches is small, for stresses less than 1.5, the exponential stretch leads to significantly more accurate predictions than the ANM stretch, as shown on Fig. 4.2. This observation is consistent with the mean squared *relative* errors (ANM: 0.00906 MPa^2, Exponential: 0.00322 MPa^2) associated with this dataset, which weight data points at low stresses much more. The good prediction accuracy of the exponential stretch at low stretches supports the validity of Assumption 9 in this range of stretches.

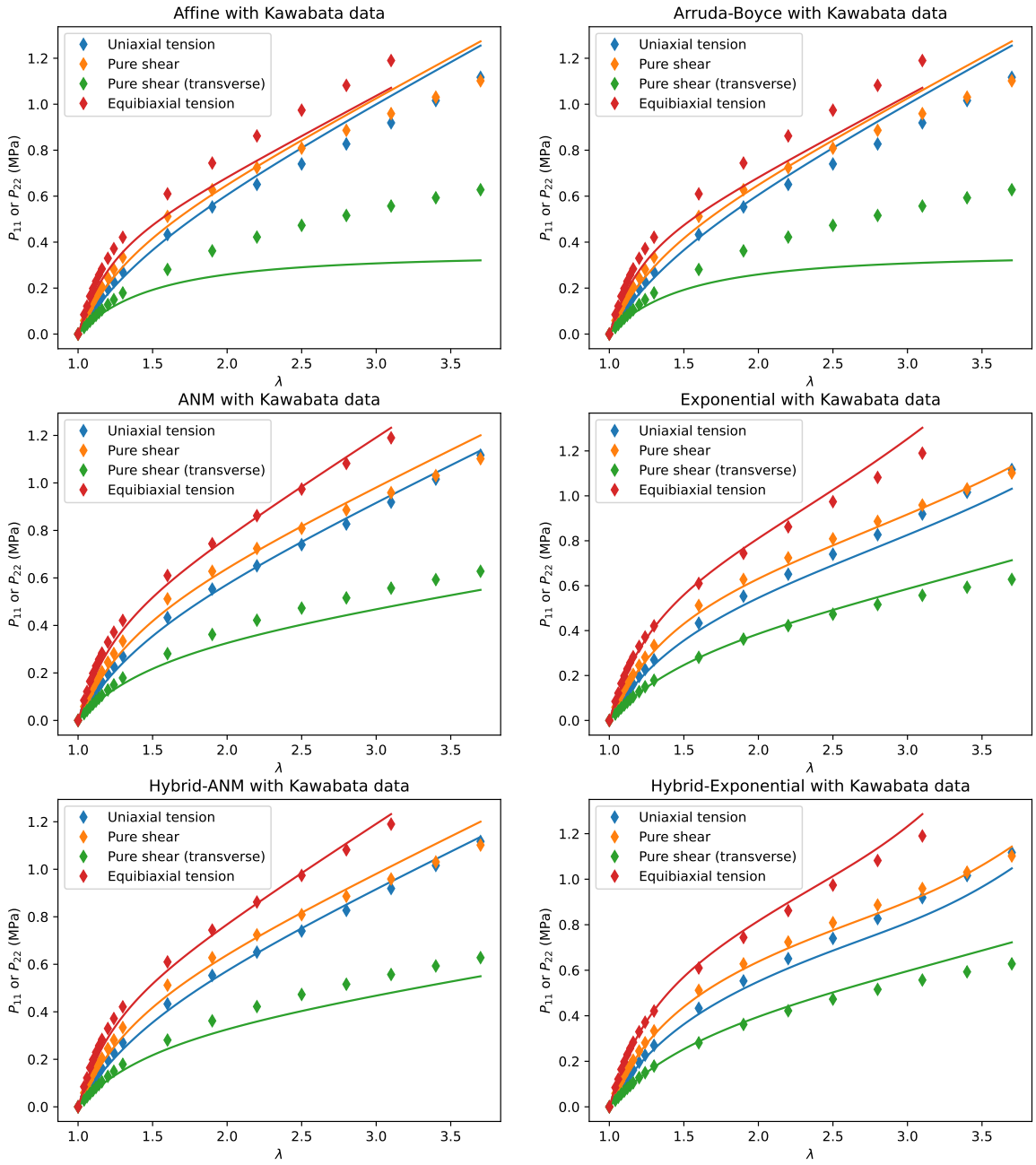


Fig. 4.2 Stress–stretch relations predicted by the constitutive models in Section 4.1 when model parameters are fitted to vulcanised natural rubber data from Kawabata et al. [56]. The model predictions are shown as solid lines, and the experimental data are shown as diamonds.

Vulcanised natural rubber data from Treloar [53]

Model predictions are plotted alongside experimental data from Treloar [53] in Fig. 4.3, and the corresponding model parameter values and MSE are given in Table 4.2. The stresses

plotted on Fig. 4.3 correspond to the P_{11} component of the first Piola–Kirchhoff stress $\mathbf{P}$. As expected, the exponential model does not fit the data well because Assumption 9 is not valid at large stretches. In particular, the exponential model does not capture the form of the singularity in stress at large stretches, unlike the Arruda–Boyce model, which has the second-best fit to the data. The hybrid-exponential model achieves the best fit, with an overall MSE that is an order of magnitude smaller than that of the Arruda–Boyce model. Remarkably, this MSE is lower than that of the extended 8-chain model of Dal et al. [36] given in Appendix I.7, which has the best fit to the dataset of Treloar [53] among the *three-parameter* models included in the extensive review by Dal et al. [7]. Because the extended 8-chain model captures the confining effects of neighbouring chains using a tube-constraint term, the good fit to data achieved by the hybrid-exponential model suggests that these interactions can instead be captured via their effect on the chain orientation distribution: at low stretches, these confining effects are captured by Assumption 9, and at large stretches, the chain orientation distribution approaches that predicted by the affine assumption (c.f. Section 3.3). The advantage of our approach is that we do not need to introduce an additional parameter for the tube-constraint term.

	$\bar{\mu}$ (MPa)	N	Mean squared error (MPa^2)			
			Overall	Uniaxial tension	Pure shear	Equibiaxial tension
Affine	0.332	62.5	0.0288	0.0159	0.0113	0.0622
A–B	0.295	26.2	0.0135	0.00812	0.00355	0.0295
ANM	0.526	61.4	0.0162	0.0235	0.0156	0.00586
H-ANM	0.412	65.8	0.0180	0.0134	0.00473	0.0356
Exp.	1.33	62.1	0.0990	0.135	0.0464	0.0891
H-Exp.	0.908	67.0	**0.00247**	**0.00361**	**0.00196**	**0.00121**

Table 4.2 Model parameters and mean squared error (MSE) when fitted to experimental data from Treloar [53]. The smallest MSE value in each column is displayed in bold font. Abbreviations: “A–B” is “Arruda–Boyce”; “H-” is “Hybrid-”; “Exp.” is “Exponential”.

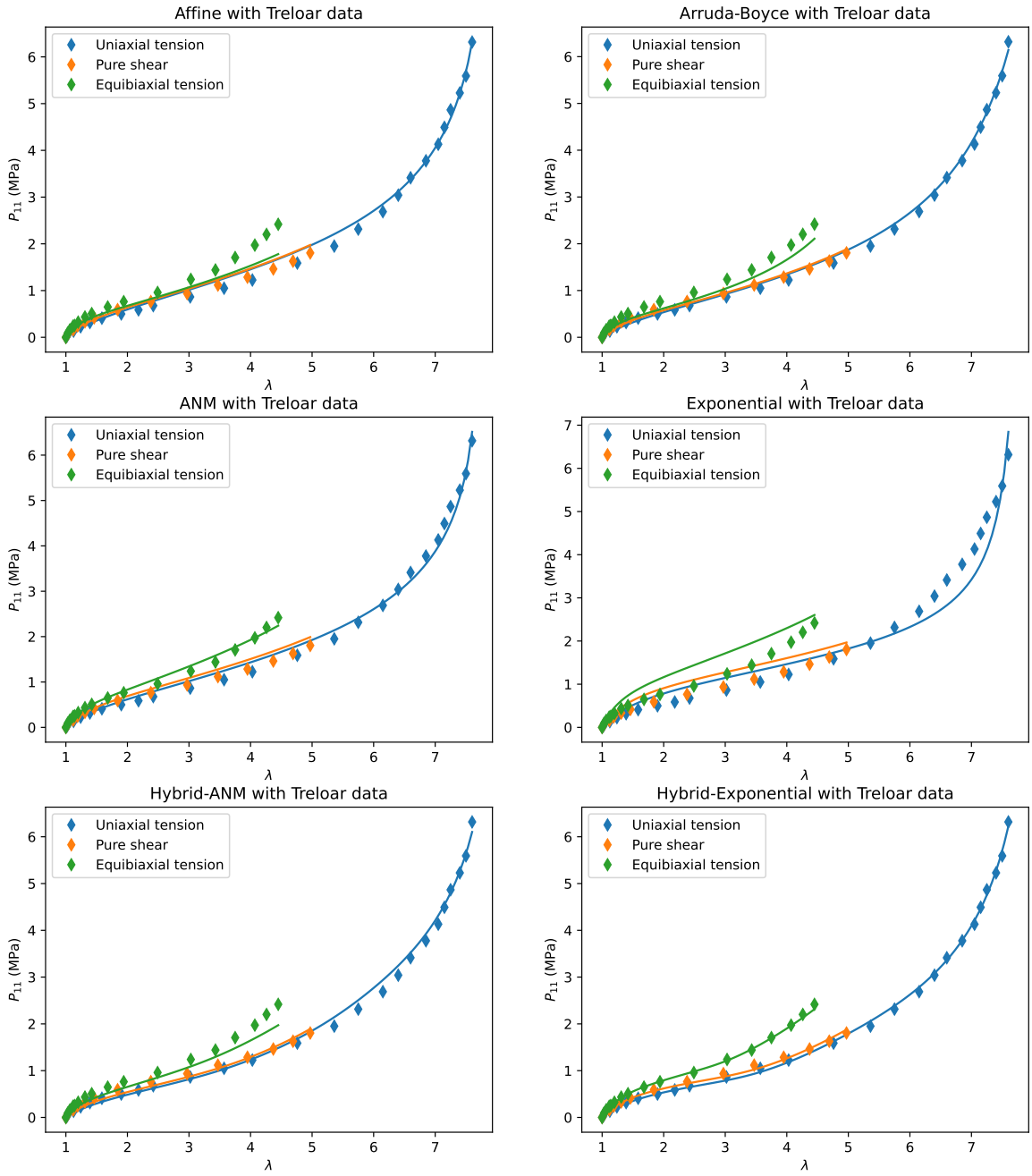


Fig. 4.3 Stress–stretch relations predicted by the constitutive models in Section 4.1 when model parameters are fitted to vulcanised natural rubber data from Treloar [53]. The model predictions are shown as solid lines, and the experimental data are shown as diamonds.

Although the hybrid-ANM model has a larger MSE than the ANM model, Fig. 4.3 shows that the hybrid-ANM model captures the qualitative features of the data better, such as the singularity at large stretches. This qualitative agreement comes from the affine-like behaviour of the hybrid-ANM model at large stretches. In Amores et al. [57], the hybrid-ANM

stretch is fitted to data from Treloar [53] using three parameters, as explained in Section 3.3. Introducing the third parameter leads to an improved fit, as shown in Appendix I.7. However, even with an additional parameter, the hybrid-ANM model does not achieve a better fit than the two-parameter hybrid-exponential model.

Silicone rubber data from Meunier et al. [66]

Model predictions are plotted alongside experimental data from Meunier et al. [66] in Fig. 4.4, and the corresponding model parameter values and MSE are given in Table 4.3. The stresses plotted on Fig. 4.4 correspond to the P_{11} component of the first Piola–Kirchhoff stress $\mathbf{P}$. For this dataset, the Arruda–Boyce and exponential models have the lowest and second-lowest MSE, respectively. However, the Arruda–Boyce model only fits the equibiaxial tension data well, and does not capture the strain-stiffening behaviour displayed by the uniaxial tension and pure shear data points on Fig. 4.4. This strain-stiffening behaviour is better captured by the ANM and exponential models. Interestingly, the hybrid-ANM and hybrid-exponential models do not result in better fits than the ANM and exponential models, respectively, suggesting that the chain orientation distribution in silicone rubber remains approximately isotropic over the range of stretches in this dataset. Based on the features shown in Fig. 4.4 and the MSE values in Table 4.3, we conclude that the exponential model succeeds in capturing the behaviour shown in the dataset of Meunier et al. [66].

	$\bar{\mu}$ (MPa)	N	Mean squared error (MPa^2)			
			Overall	Uniaxial tension	Pure shear	Equibiaxial tension
Affine	0.268	6.34	0.00553	0.00497	0.000935	0.0125
A–B	0.266	4.12	**0.000764**	0.000925	0.000829	**0.000467**
ANM	0.374	6.17	0.00201	0.00198	0.000666	0.00389
H-ANM	0.335	7.25	0.00393	0.00341	0.000846	0.00877
Exp.	0.636	5.9	0.000943	0.000788	0.000486	0.00176
H-Exp.	0.753	12.1	0.00116	**0.000543**	**0.000457**	0.00292

Table 4.3 Model parameters and mean squared error (MSE) when fitted to experimental data from Meunier et al. [66]. The smallest MSE value in each column is displayed in bold font. Abbreviations: “A–B” is “Arruda–Boyce”; “H-” is “Hybrid-”; “Exp.” is “Exponential”.

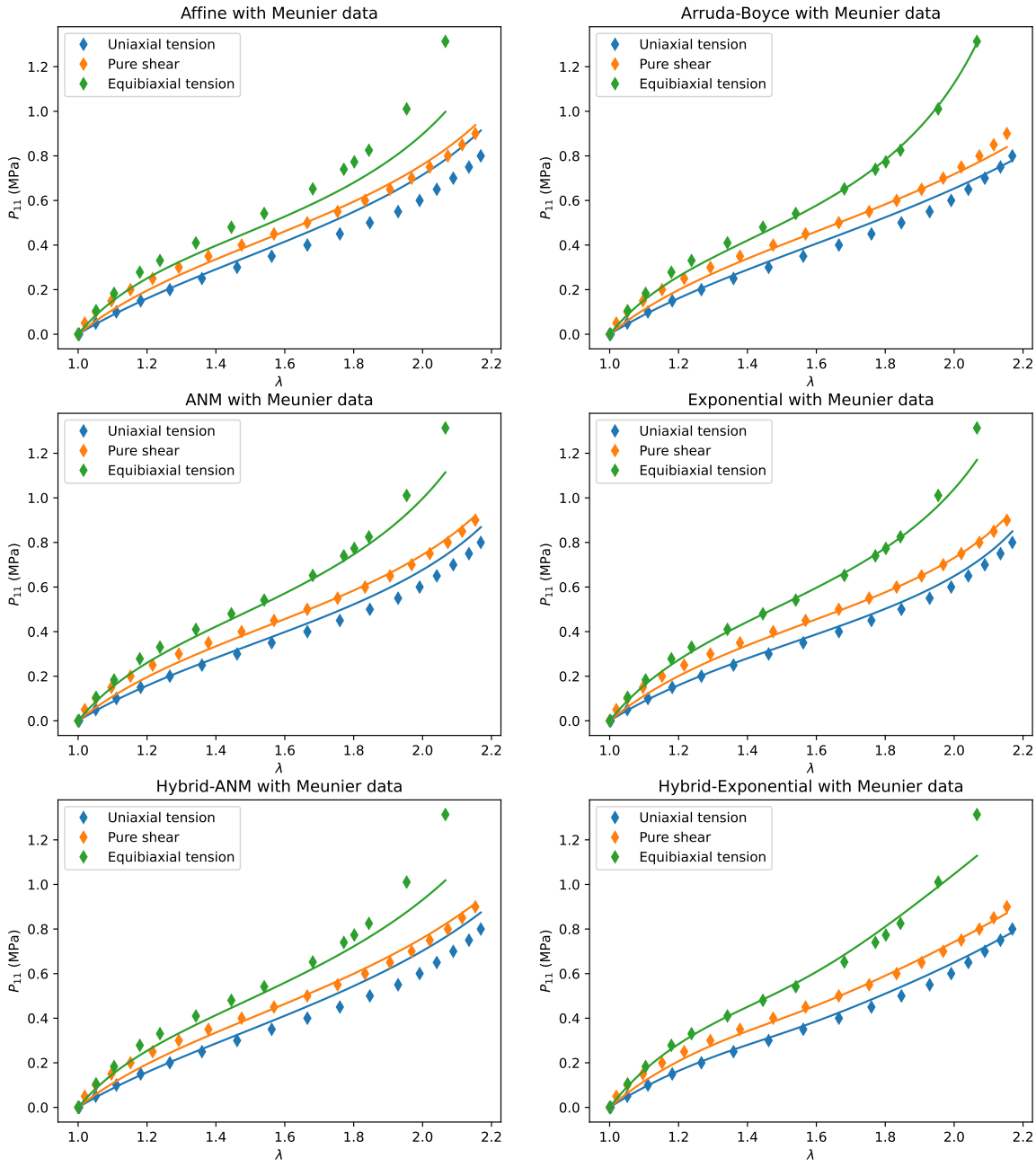


Fig. 4.4 Stress–stretch relations predicted by the constitutive models in Section 4.1 when model parameters are fitted to silicone rubber data from Meunier et al. [66]. The model predictions are shown as solid lines, and the experimental data are shown as diamonds.

Silane-modified polymer adhesive data from Moreira and Nunes [67]

Model predictions are plotted alongside experimental data from Moreira and Nunes [67] in Fig. 4.5, and the corresponding model parameter values and MSE are given in Table 4.4.

The stresses plotted on Fig. 4.5 correspond to the P_{12} component of the first Piola–Kirchhoff stress $\mathbf{P}$ for simple shear and to the P_{11} component for pure shear. In simple shear, stress is plotted against the largest principal stretch, which is given in terms of the shear strain γ by $\lambda = \sqrt{\frac{1}{2}(2+\gamma^2+\gamma\sqrt{4+\gamma^2})}$. For this dataset, the exponential model gives the most accurate predictions. The hybrid-exponential and exponential models give similar predictions because this dataset does not include data from the strain stiffening regime.

	$\bar{\mu}$ (MPa)	N	Mean squared error (MPa^2)		
			Overall	Pure Shear	Simple Shear
Affine	0.426	1.99×10^4	0.0170	0.0174	0.0166
A–B	0.426	2.71×10^4	0.0170	0.0173	0.0166
ANM	0.638	1.16×10^4	0.0136	0.0138	0.0133
H-ANM	0.639	6.97×10^3	0.0136	0.0138	0.0133
Exp.	1.31	2.13×10^4	**0.00450**	**0.00510**	**0.00390**
H-Exp.	1.28	178	0.00502	0.00591	0.00413

Table 4.4 Model parameters and mean squared error when fitted to experimental data from Moreira and Nunes [67]. The smallest MSE value in each column is displayed in bold font. Abbreviations: “A–B” is “Arruda–Boyce”; “H-” is “Hybrid-”; “Exp.” is “Exponential”.

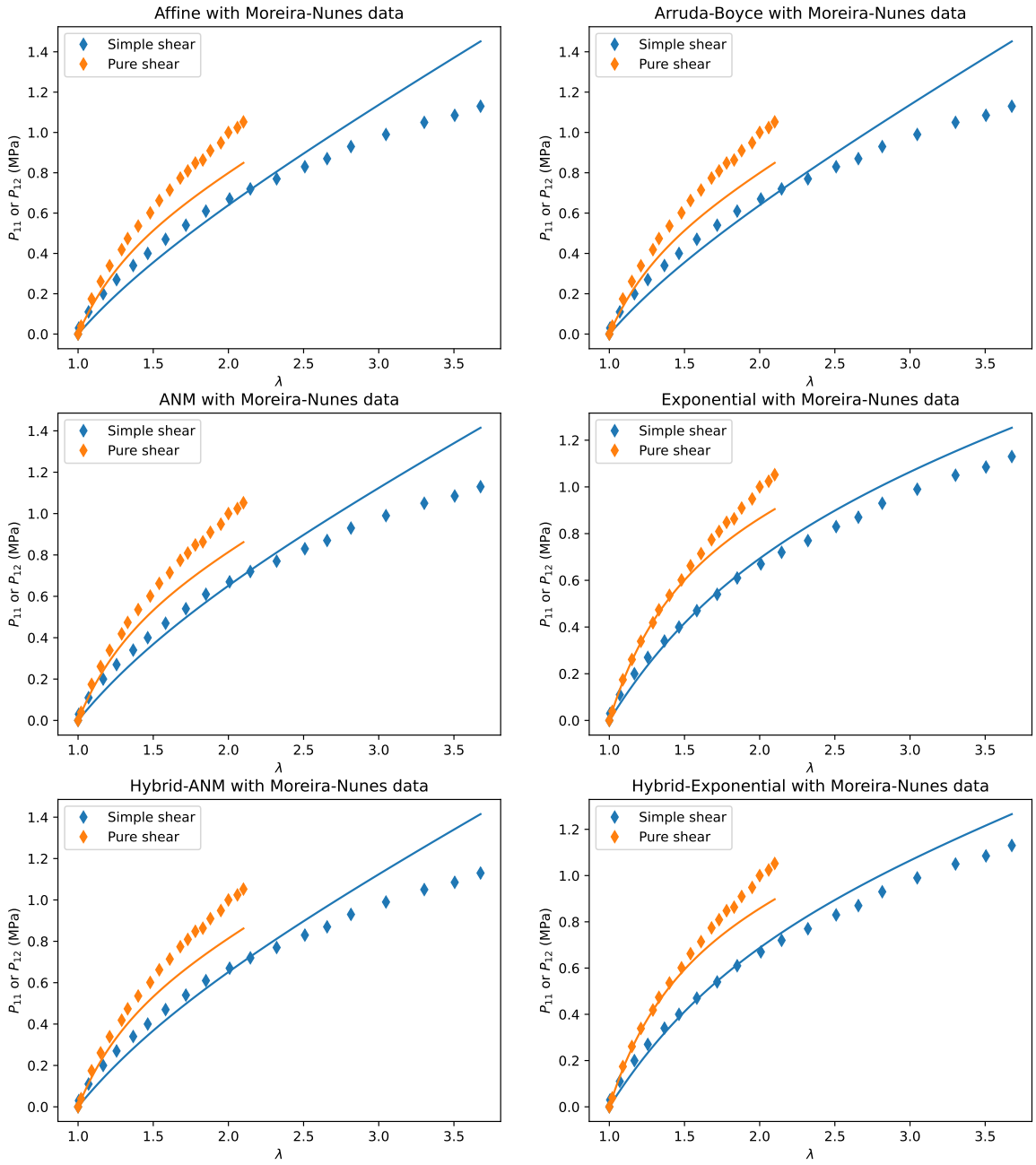


Fig. 4.5 Stress–stretch relations predicted by the constitutive models in Section 4.1 when model parameters are fitted to silane-modified polymer adhesive data from Moreira and Nunes [67]. The model predictions are shown as solid lines, and the experimental data are shown as diamonds. The stresses are plotted against λ, the largest principal stretch associated with each shear strain.

Polyacrylamide gel data from Janmey et al. [68]

Model predictions are plotted alongside simple shear experimental data from Janmey et al. [68] in Fig. 4.6, and the corresponding model parameter values and MSE are given in Table 4.5. The stresses plotted on Fig. 4.5 correspond to the shear P_{12} and normal P_{22} components of the first Piola–Kirchhoff stress $\mathbf{P}$. Stress is plotted against the largest principal stretch, which is given in terms of the shear strain γ by $\lambda = \sqrt{\frac{1}{2}(2+\gamma^2+\gamma\sqrt{4+\gamma^2})}$. For this dataset, although the affine and Arruda–Boyce models have the best fit to the shear stress, both are unable to capture even the qualitative behaviour of the normal stress: as discussed in Section 4.2, the affine model predicts the reverse Poynting effect, whereas the Arruda–Boyce model predicts zero normal stress in simple shear. Both the ANM and exponential models, and their corresponding hybrid models, correctly predict that the normal stress is compressive. Quantitatively, the exponential model provides the best overall fit to both the shear and normal stresses, as shown in Table 4.5. The values of N in Table 4.5 are all equal to the starting guess $N = 60$ used in the fitting procedure because the residual function is insensitive to N in the range considered.

	$\bar{\mu}$ (MPa)	N	Mean squared error (MPa^2)		
			Overall	Shear Stress	Normal Stress
Affine	0.0611	60	0.000325	$\mathbf{7.86\times10^{-7}}$	0.000648
A–B	0.0617	60	0.000319	8.57×10^{-7}	0.000637
ANM	0.0923	60	0.000205	5.14×10^{-6}	0.000404
H-ANM	0.0919	60	0.000206	4.97×10^{-6}	0.000407
Exp.	0.172	60	**0.000111**	1.27×10^{-5}	**0.000209**
H-Exp.	0.172	60	**0.000111**	1.27×10^{-5}	**0.000209**

Table 4.5 Model parameters and mean squared error when fitted to experimental data from Janmey et al. [68]. The smallest MSE value in each column is displayed in bold font. Abbreviations: “A–B” is “Arruda–Boyce”; “H-” is “Hybrid-”; “Exp.” is “Exponential”.

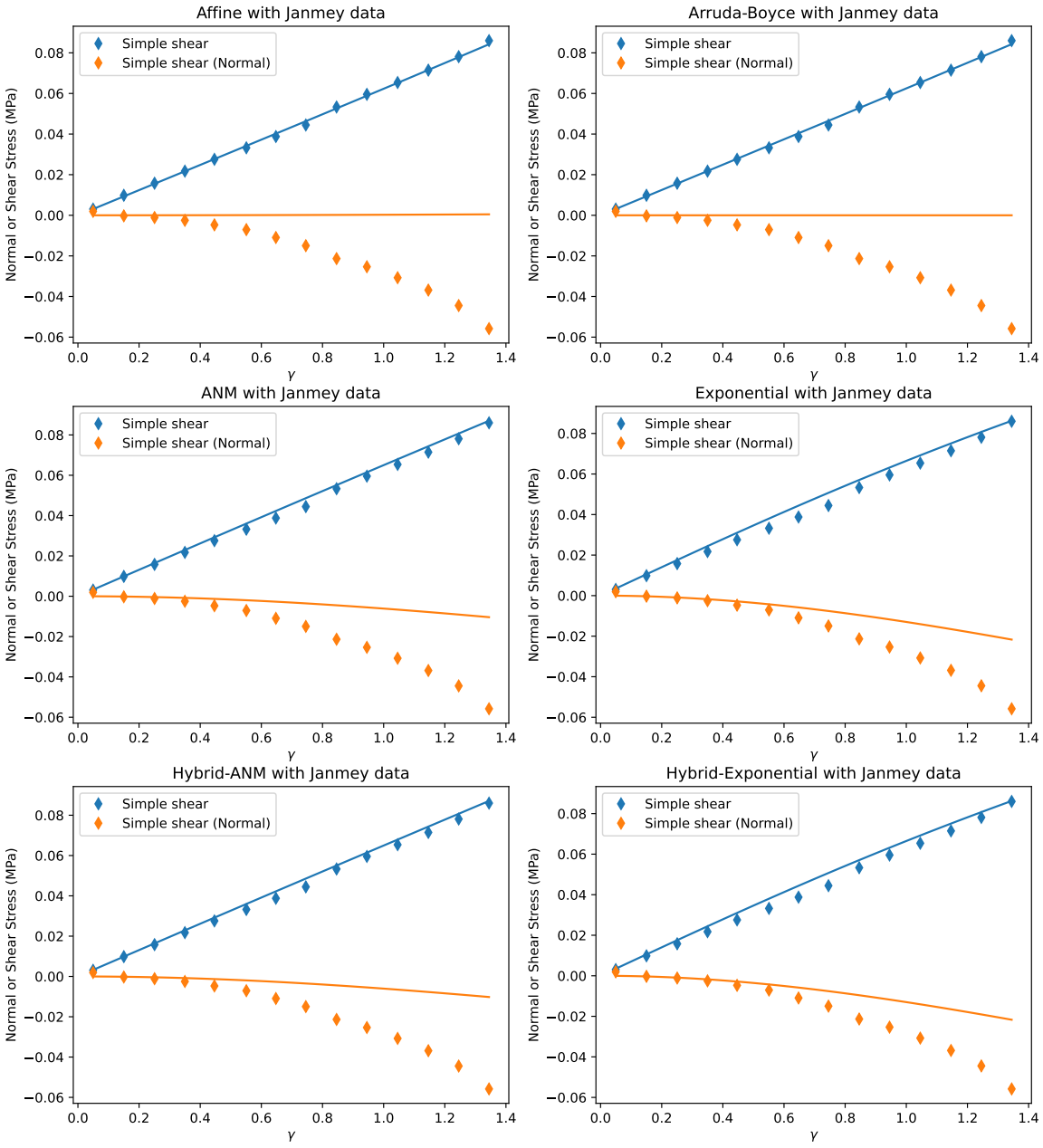


Fig. 4.6 Stress–stretch relations predicted by the constitutive models in Section 4.1 when model parameters are fitted to polyacrylamide gel data from Janmey et al. [68]. The model predictions are shown as solid lines, and the experimental data are shown as diamonds. The stresses are plotted against shear strain γ.

4.4 Summary

In this chapter, we compared the predictions of constitutive models based on our new chain stretch relations and on previously proposed chain stretches to experimental stress–stretch data. We showed that, unlike the affine chain stretch, the exponential and hybrid-exponential chain stretches capture the classical Poynting effect in simple shear, in qualitative agreement with experiments on rubber-like materials [65]. We also performed quantitative comparisons of constitutive model predictions to stress–stretch data from five different experimental datasets. For the vulcanised natural rubber data of Kawabata et al. [56], the silane-modified polymer adhesive data of Moreira and Nunes [67] and the polyacrylamide gel data of Janmey et al. [68], the exponential and hybrid-exponential chain stretches had very similar predictions and achieved the best fits to the data. Although the Arruda–Boyce model's fit to silicone rubber data from Meunier et al. [66] had a smaller mean squared error than that of the exponential model, the strain-stiffening behaviour seen on the stress–stretch plots was better captured by the exponential model, which had the second-lowest mean squared error of the constitutive models tested. This dataset did not display the transition to affine-like behaviour at large stretches, so the hybrid-exponential stretch did not fit the data as well as the exponential stretch. For the vulcanised rubber dataset of Treloar [53], which covered the widest range of stretches among the datasets considered, the hybrid exponential model outperformed all the other models by a large margin, accurately fitting all data from all deformation modes. We have therefore shown that the exponential chain stretch makes accurate predictions for low to moderate stretches, and the hybrid-exponential stretch fits data accurately over the whole range of stretches when significant strain-stiffening is present.

Chapter 5

Conclusions: elastomeric constitutive models

Motivated by the three criteria of Anssari-Benam and Horgan [25] listed in Chapter 2, we have proposed two-parameter constitutive models for elastomers based on new chain stretch relations. To obtain the exponential chain stretch, we first constructed a Cauchy stress tensor from chain force contributions, using the Irving–Kirkwood–Noll procedure to bridge statistical mechanics and nonlinear elasticity. This construction assumed that the chain orientation distribution in the deformed state of the material is isotropic for low to moderate stretches, based on the experimental results of Toki et al. [37, 38] and Ott et al. [39]. Then, we arrived at the exponential stretch by solving a differential equation obtained by requiring that our constructed stress tensor (3.11) be consistent with the hyperelastic stress (3.14) from the strain energy function (2.10). We then obtained the hybrid-exponential chain stretch by interpolating between the exponential chain stretch and a chain stretch that reproduces affine behaviour, to capture the isotropic chain orientation distribution at low stretches and the increase in anisotropy at large stretches.

We found that the constitutive models based on our new chain stretch relations not only capture qualitative aspects of experimental data successfully, such as the Poynting effect in simple shear, but also fit experimental stress–stretch data from five different datasets accurately. The good agreement between our model predictions and these datasets provides indirect experimental confirmation of the theoretical proposals in Chapter 3. Importantly, our new chain stretch relations lead to improvements in prediction accuracy over previously proposed models without increasing the number of fitting parameters. Therefore, our new models satisfy all three criteria of Anssari-Benam and Horgan [25], because they fit multiple deformation modes accurately, have only two parameters, and are molecular-statistical models.

Despite the observation by Dal et al. [7] that, of the 44 constitutive models they surveyed, those that accurately fit the datasets of Treloar [53] and Kawabata et al. [56] have at least three parameters, we showed that two parameters are sufficient for good fits. Dal et al. [7] also note that most of the constitutive models that make accurate predictions consider tube-constraint effects. If the relative isotropy of the chain orientation distribution in the deformed state is due to these effects, as suggested by Ott et al. [39], then our approach accounts for the confining effects of neighbouring chains via the chain orientation distribution, without introducing additional parameters for the tube-constraint terms.

Our findings motivate further work on the deformation dependence of the chain orientation distribution, to explain its approximate isotropy at low to moderate stretches, as observed by Toki et al. [37, 38] and Ott et al. [39], and to characterise its anisotropy at large stretches. Molecular dynamics simulations similar to that by Munoz et al. [76] could provide useful insights, and would enable direct testing of the validity of the assumptions made in Chapter 3. Obtaining a general characterisation of the chain orientation distribution over the whole range of deformations would advance the molecular-statistical constitutive modelling of elastomers by leading to more accurate predictions and providing firmer grounding of model parameters in molecular quantities.

Appendix I

I.1 Derivatives with respect to tensors

Given normed vector spaces $(\mathcal{V}, ||\cdot||_1)$ and $(\widetilde{\mathcal{V}}, ||\cdot||_2)$, the Fréchet derivative $\mathrm{D}f(u)$ of a function $f : \mathcal{V} \to \widetilde{\mathcal{V}}$ at $u \in \mathcal{V}$ is defined to be the bounded linear map $L(u) : \mathcal{V} \to \widetilde{\mathcal{V}}$ satisfying [77]

$$\lim_{||h||_1 \to 0} \frac{||f(u+h) - f(u) - L(u)[h]||_2}{||h||_1} = 0$$

where $h \in \mathcal{V}$. If the Fréchet derivative at $u \in \mathcal{V}$ exists, it is unique. This definition generalises the notion of differentiation to functions defined on any normed vector space, including the space of tensors $\mathtt{L}(\mathcal{V}_d)$, defined in Section 2.1.1. This space has norm $||\mathbf{A}||$ induced by the tensor inner product, given by $||\mathbf{A}|| = \sqrt{\mathbf{A} : \mathbf{A}}$. Therefore, the derivative at $\mathbf{A} \in \mathtt{L}(\mathcal{V}_d)$ of a real-valued function $f : \mathtt{L}(\mathcal{V}_d) \to \mathbb{R}$ is a linear map $\mathrm{D}f(\mathbf{A}) : \mathtt{L}(\mathcal{V}_d) \to \mathbb{R}$, i.e. a linear functional on the space of tensors. By the Riesz Representation Theorem [42], there exists a unique tensor $\mathbf{K}(\mathbf{A})$ such that

$$\mathbf{K}(\mathbf{A}) : \mathbf{H} = \mathrm{D}f(\mathbf{A})[\mathbf{H}]$$

for all $\mathbf{H} \in \mathtt{L}(\mathcal{V}_d)$, i.e. taking the inner product with $\mathbf{K}(\mathbf{A})$ is equivalent to applying the linear map $\mathrm{D}f(\mathbf{A})$. The gradient of f with respect to $\mathbf{X}$ is defined to be [6]

$$\frac{\mathrm{d}f}{\mathrm{d}\mathbf{A}}(\mathbf{A}) = \mathbf{K}(\mathbf{A}).$$

Using a common abuse of terminology, the derivative of a real-valued function of a tensor can be treated as a tensor by employing the gradient instead.

I.2 Relationship between derivatives of the chain stretch

Because the deformation gradient $\mathbf{F}$ necessarily satisfies $\det\mathbf{F} > 0$, the function $\tilde{\lambda}(\hat{\mathbf{r}}, \cdot)$ has domain $\mathtt{L}(\mathcal{V}_d)^+ = \{\mathbf{A} \in \mathtt{L}(\mathcal{V}_d) : \det\mathbf{A} > 0\}$. Using the definitions in Appendix I.1 and subscripts on the differentiation operator D to identify the argument with respect to which differentiation is carried out, we obtain, for all $\mathbf{H} \in \mathtt{L}(\mathcal{V}_d)^+$,

$$\begin{aligned}\frac{\partial\tilde{\lambda}}{\partial\mathbf{F}}(\hat{\mathbf{r}},\mathbf{F}) : \mathbf{H} &= \mathrm{D}_{\mathbf{F}}\tilde{\lambda}(\hat{\mathbf{r}},\mathbf{F})[\mathbf{H}] = \mathrm{D}_{\mathbf{B}}\lambda(\hat{\mathbf{r}},\mathbf{B})[\mathbf{F}\mathbf{H}^{\mathrm{T}} + \mathbf{H}\mathbf{F}^{\mathrm{T}}] \\ &= \frac{\partial\lambda}{\partial\mathbf{B}}(\hat{\mathbf{r}},\mathbf{B}) : (\mathbf{F}\mathbf{H}^{\mathrm{T}} + \mathbf{H}\mathbf{F}^{\mathrm{T}}) = \mathrm{tr}\left[\left(\frac{\partial\lambda}{\partial\mathbf{B}}\right)^{\mathrm{T}}\mathbf{F}\mathbf{H}^{\mathrm{T}}\right] + \mathrm{tr}\left[\left(\frac{\partial\lambda}{\partial\mathbf{B}}\right)^{\mathrm{T}}\mathbf{H}\mathbf{F}^{\mathrm{T}}\right] \\ &= \mathrm{tr}\left[\mathbf{H}^{\mathrm{T}}\frac{\partial\lambda}{\partial\mathbf{B}}\mathbf{F}\right] + \mathrm{tr}\left[\left(\frac{\partial\lambda}{\partial\mathbf{B}}\mathbf{F}\right)^{\mathrm{T}}\mathbf{H}\right] = \left(2\frac{\partial\lambda}{\partial\mathbf{B}}(\hat{\mathbf{r}},\mathbf{B})\mathbf{F}\right) : \mathbf{H}.\end{aligned}$$

where the first equality follows by the definition of the derivative above; the second equality comes from the chain rule; the fourth equality follows from the definition of the tensor inner product in Section 2.1.1, and the fifth equality holds because $\frac{\partial\lambda}{\partial\mathbf{B}}$ is symmetric. Because the above relation holds for any $\mathbf{H} \in \mathtt{L}(\mathcal{V}_d)^+$, the result in (3.9) follows.

I.3 Kinetic stress is a multiple of the identity tensor

Using Assumption 5 together with the easily obtained identity

$$\langle g | \mathbf{x}_\alpha = \mathbf{x}\rangle = \int_{\mathbb{R}^3} \langle g | \mathbf{x}_\alpha = \mathbf{x}, \mathbf{v}_\alpha = \mathbf{v}\rangle \, \mathrm{d}\mathbf{v}$$

and the change of variables $\mathbf{v} = \mathbf{w} + \bar{\mathbf{v}}(\mathbf{x})$, we obtain

$$\begin{aligned}\left\langle \left(\mathbf{v}_\alpha^{\mathrm{rel}} \otimes \mathbf{v}_\alpha^{\mathrm{rel}}\right) \middle| \mathbf{x}_\alpha = \mathbf{x} \right\rangle &= \int_{\mathbb{R}^3} \left\langle \left(\mathbf{v}_\alpha^{\mathrm{rel}} \otimes \mathbf{v}_\alpha^{\mathrm{rel}}\right) \middle| \mathbf{x}_\alpha = \mathbf{x}, \mathbf{v}_\alpha = \mathbf{v} \right\rangle \mathrm{d}\mathbf{v} \\ &= \int_{\mathbb{R}^3} \left\langle (\mathbf{w} \otimes \mathbf{w}) \,|\, \mathbf{x}_\alpha = \mathbf{x}, \mathbf{v}_\alpha = \mathbf{w} + \bar{\mathbf{v}}(\mathbf{x}) \right\rangle \mathrm{d}\mathbf{w},\end{aligned} \tag{I.1}$$

since $\mathbf{v}_\alpha^{\mathrm{rel}} = \mathbf{v}_\alpha - \bar{\mathbf{v}}(\mathbf{x}) = \mathbf{w}$. Letting $\mathbf{w} = w\hat{\mathbf{w}}$, where $w = |\mathbf{w}|$ and $\hat{\mathbf{w}} = \mathbf{w}/w$,

$$\begin{aligned}\left\langle \left(\mathbf{v}_\alpha^{\mathrm{rel}} \otimes \mathbf{v}_\alpha^{\mathrm{rel}}\right) \middle| \mathbf{x}_\alpha = \mathbf{x} \right\rangle &= \int_{\mathcal{S}^2} \int_0^\infty \left\langle w^2 \left(\hat{\mathbf{w}} \otimes \hat{\mathbf{w}}\right) \mid \mathbf{x}_\alpha = \mathbf{x}, \mathbf{v}_\alpha = w\hat{\mathbf{w}} + \bar{\mathbf{v}}(\mathbf{x}) \right\rangle \mathrm{d}w \, \mathrm{d}A(\hat{\mathbf{w}}) \\ &= \int_{\mathcal{S}^2} \left(\hat{\mathbf{w}} \otimes \hat{\mathbf{w}}\right) \int_0^\infty w^2 \left\langle 1 \mid \mathbf{x}_\alpha = \mathbf{x}, \mathbf{v}_\alpha = w\hat{\mathbf{w}} + \bar{\mathbf{v}}(\mathbf{x}) \right\rangle \mathrm{d}w \, \mathrm{d}A(\hat{\mathbf{w}}) \\ &= \int_{\mathcal{S}^2} \left(\hat{\mathbf{w}} \otimes \hat{\mathbf{w}}\right) \eta_\alpha(\mathbf{x}) \, \mathrm{d}A(\hat{\mathbf{w}}),\end{aligned} \tag{I.2}$$

where $\eta_\alpha(\mathbf{x}) = \int_0^\infty w^2 \left\langle 1 \mid \mathbf{x}_\alpha = \mathbf{x}, \mathbf{v}_\alpha = w\hat{\mathbf{w}} + \bar{\mathbf{v}}(\mathbf{x}) \right\rangle \mathrm{d}w$ is independent of $\hat{\mathbf{w}}$ by Assumption 5. If thermal effects are considered, $\eta_\alpha(\mathbf{x})$ would be some function of temperature. Since $\int_{\mathcal{S}^2} \left(\hat{\mathbf{w}} \otimes \hat{\mathbf{w}}\right) \mathrm{d}A(\hat{\mathbf{w}}) = \frac{4\pi}{3}\mathbf{I}$, as proven in Appendix II.6,

$$\mathbf{T}_K(\mathbf{x}) = \left(-\frac{4\pi}{3} \sum_{\alpha=1}^{M} m_\alpha \eta_\alpha(\mathbf{x}) \right) \mathbf{I}, \tag{I.3}$$

showing that $\mathbf{T}_K(\mathbf{x})$ is a multiple of the identity tensor (a spherical tensor).

I.4 Cauchy stress for affine full network model

Using $\det \mathbf{F} = 1$ in the relation between stress tensors $\mathbf{P} = (\det \mathbf{F})\mathbf{T}\mathbf{F}^{-\mathrm{T}}$, the hyperelastic relation (2.5) for incompressible materials, and the strain energy (2.10) with the affine stretch (2.11) leads to

$$\begin{aligned}\mathbf{T} = \mathbf{P}\mathbf{F}^{\mathbf{T}} &= \left(\frac{\mathrm{d}W}{\mathrm{d}\mathbf{F}}(\mathbf{F}) \right) \mathbf{F}^{\mathbf{T}} - p_{\mathrm{c}}\mathbf{I} = \frac{\mathrm{d}}{\mathrm{d}\mathbf{F}} \left(\frac{n}{4\pi} \int_{\mathcal{S}^2} \psi(r_0 \lambda_{\mathrm{aff}}^0(\hat{\mathbf{r}}_0, \mathbf{F})) \, \mathrm{d}A(\hat{\mathbf{r}}_0) \right) \mathbf{F}^{\mathbf{T}} - p_{\mathrm{c}}\mathbf{I} \\ &= \left(\frac{nr_0}{4\pi} \int_{\mathcal{S}^2} \psi'(r_0 \lambda_{\mathrm{aff}}^0(\hat{\mathbf{r}}_0, \mathbf{F})) \frac{\partial \lambda_{\mathrm{aff}}^0}{\partial \mathbf{F}}(\hat{\mathbf{r}}_0, \mathbf{F}) \, \mathrm{d}A(\hat{\mathbf{r}}_0) \right) \mathbf{F}^{\mathbf{T}} - p_{\mathrm{c}}\mathbf{I},\end{aligned} \tag{I.1}$$

where the integration is over chain orientations $\hat{\mathbf{r}}_0$ in the reference state, in which chains are isotropically oriented.

Let $\hat{\lambda}_{\mathrm{aff}}(\hat{\mathbf{r}}_0, \mathbf{C}) = \lambda_{\mathrm{aff}}^0(\hat{\mathbf{r}}_0, \mathbf{F}) = \sqrt{\hat{\mathbf{r}}_0 \cdot \mathbf{C}\hat{\mathbf{r}}_0} = \sqrt{\mathbf{C} : \hat{\mathbf{r}}_0 \otimes \hat{\mathbf{r}}_0}$ be the affine chain stretch expressed as a function of $\mathbf{C}$. Then, differentiation of the affine stretch with respect to $\mathbf{C}$ gives $\frac{\partial \hat{\lambda}_{\mathrm{aff}}}{\partial \mathbf{C}}(\hat{\mathbf{r}}_0, \mathbf{C}) = \frac{1}{2\hat{\lambda}_{\mathrm{aff}}(\hat{\mathbf{r}}_0, \mathbf{C})} \hat{\mathbf{r}}_0 \otimes \hat{\mathbf{r}}_0$. The derivatives $\frac{\partial \hat{\lambda}_{\mathrm{aff}}}{\partial \mathbf{C}}(\hat{\mathbf{r}}_0, \mathbf{C})$ and $\frac{\partial \lambda_{\mathrm{aff}}}{\partial \mathbf{F}}(\hat{\mathbf{r}}_0, \mathbf{F})$ can be related as

follows, using the notation and definitions of Appendix I.1. For any $\mathbf{H} \in \mathtt{L}(\mathcal{V}_d)^+$, we obtain

$$\begin{aligned}\frac{\partial \lambda^0_{\text{aff}}}{\partial \mathbf{F}}(\hat{\mathbf{r}}_0,\mathbf{F}) : \mathbf{H} &= \mathrm{D}_{\mathbf{F}}\lambda^0_{\text{aff}}(\hat{\mathbf{r}}_0,\mathbf{F})[\mathbf{H}] = \mathrm{D}_{\mathbf{C}}\hat{\lambda}_{\text{aff}}(\hat{\mathbf{r}}_0,\mathbf{C})[\mathbf{F}^{\mathrm{T}}\mathbf{H}+\mathbf{H}^{\mathrm{T}}\mathbf{F}] \\ &= \frac{\partial \hat{\lambda}_{\text{aff}}}{\partial \mathbf{C}}(\hat{\mathbf{r}}_0,\mathbf{C}) : (\mathbf{F}^{\mathrm{T}}\mathbf{H}+\mathbf{H}^{\mathrm{T}}\mathbf{F}) = \mathrm{tr}\left[\left(\frac{\partial \hat{\lambda}_{\text{aff}}}{\partial \mathbf{C}}\right)^{\mathrm{T}}\mathbf{F}^{\mathrm{T}}\mathbf{H}\right] + \mathrm{tr}\left[\left(\frac{\partial \hat{\lambda}_{\text{aff}}}{\partial \mathbf{C}}\right)^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{F}\right] \\ &= \mathrm{tr}\left[\left(\mathbf{F}\frac{\partial \hat{\lambda}_{\text{aff}}}{\partial \mathbf{C}}\right)^{\mathrm{T}}\mathbf{H}\right] + \mathrm{tr}\left[\mathbf{H}^{\mathrm{T}}\mathbf{F}\frac{\partial \hat{\lambda}_{\text{aff}}}{\partial \mathbf{C}}\right] = \left(2\mathbf{F}\frac{\partial \hat{\lambda}_{\text{aff}}}{\partial \mathbf{C}}(\hat{\mathbf{r}}_0,\mathbf{C})\right) : \mathbf{H},\end{aligned} \tag{I.2}$$

where the first equality follows from the definition of the derivative; the second equality results from an application of the chain rule; the fourth equality comes from the definition of the tensor inner product and the fifth equality uses the symmetry of the tensor $\frac{\partial \hat{\lambda}_{\text{aff}}}{\partial \mathbf{C}}$. From this, it follows that

$$\frac{\partial \lambda^0_{\text{aff}}}{\partial \mathbf{F}}(\hat{\mathbf{r}}_0,\mathbf{F}) = 2\mathbf{F}\frac{\partial \hat{\lambda}_{\text{aff}}}{\partial \mathbf{C}}(\hat{\mathbf{r}}_0,\mathbf{C}), \tag{I.3}$$

which leads to

$$\frac{\partial \lambda^0_{\text{aff}}}{\partial \mathbf{F}}(\hat{\mathbf{r}}_0,\mathbf{F}) = 2\mathbf{F}\frac{1}{2\hat{\lambda}_{\text{aff}}(\hat{\mathbf{r}}_0,\mathbf{C})}\hat{\mathbf{r}}_0\otimes\hat{\mathbf{r}}_0 = \hat{\mathbf{r}}(\hat{\mathbf{r}}_0)\otimes\hat{\mathbf{r}}_0, \tag{I.4}$$

where $\hat{\mathbf{r}}(\hat{\mathbf{r}}_0) = \frac{\mathbf{F}\hat{\mathbf{r}}_0}{|\mathbf{F}\hat{\mathbf{r}}_0|}$ is the unit vector parallel to the image of $\hat{\mathbf{r}}_0$ under the affine transformation given by $\mathbf{F}$. Substituting the value of $\frac{\partial \lambda_{\text{aff}}}{\partial \mathbf{F}}(\hat{\mathbf{r}}_0,\mathbf{F})$ into the expression (I.1) for the Cauchy stress gives

$$\begin{aligned}\mathbf{T} &= \left(\frac{nr_0}{4\pi}\int_{\mathcal{S}^2}\psi'(r_0\lambda^0_{\text{aff}}(\hat{\mathbf{r}}_0,\mathbf{F}))\,\hat{\mathbf{r}}(\hat{\mathbf{r}}_0)\otimes\hat{\mathbf{r}}_0\,\mathrm{d}A(\hat{\mathbf{r}}_0)\right)\mathbf{F}^{\mathrm{T}} - p_{\mathrm{c}}\mathbf{I} \\ &= \frac{nr_0}{4\pi}\int_{\mathcal{S}^2} f_{\mathrm{ch}}(\hat{\mathbf{r}}(\hat{\mathbf{r}}_0),\mathbf{F})\,\lambda^0_{\text{aff}}(\hat{\mathbf{r}}_0,\mathbf{F})\,\hat{\mathbf{r}}(\hat{\mathbf{r}}_0)\otimes\hat{\mathbf{r}}(\hat{\mathbf{r}}_0)\mathrm{d}A(\hat{\mathbf{r}}_0) - p_{\mathrm{c}}\mathbf{I}\end{aligned}$$

where $f_{\mathrm{ch}}(\hat{\mathbf{r}}(\hat{\mathbf{r}}_0),\mathbf{F}) = \psi'(r_0\lambda^0_{\text{aff}}(\hat{\mathbf{r}}_0,\mathbf{F}))$ is the magnitude of the force exerted on a polymer chain to maintain its end-to-end separation. Next, a change of variables is performed so that the variable of integration is $\hat{\mathbf{r}} = \mathbf{F}\hat{\mathbf{r}}_0/|\mathbf{F}\hat{\mathbf{r}}_0|$. This is accomplished using the polar coordinates (θ_0,ϕ_0) and (θ,ϕ) to parameterise $\hat{\mathbf{r}}_0$ and $\hat{\mathbf{r}}$, respectively, and using the Jacobian of the transformation linking (θ_0,ϕ_0) and (θ,ϕ), as done by Wu and Giessen [46]. The result of this operation is

$$\mathbf{T} = nr_0\int_{\mathcal{S}^2} f(\hat{\mathbf{r}},\mathbf{F})\,\lambda_{\text{aff}}(\hat{\mathbf{r}},\mathbf{F})\,\hat{\rho}(\hat{\mathbf{r}},\mathbf{F})\,\hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\,\mathrm{d}A(\hat{\mathbf{r}}) - p_{\mathrm{c}}\mathbf{I}, \tag{I.5}$$

where $\lambda_{\text{aff}}(\hat{\mathbf{r}},\mathbf{F})$ is the affine stretch (2.12) expressed in terms of chain orientations in the deformed state, and $\hat{\rho}(\hat{\mathbf{r}},\mathbf{F}) = (4\pi)^{-1}\lambda_{\text{aff}}(\hat{\mathbf{r}},\mathbf{F})^3$ is the chain orientation distribution for the deformed state. Thus, the form of Cauchy stress $\mathbf{T}$ for the affine full network model is a special case of (3.11) with $\mathbf{T}_c$ in the form given by (3.10) and $\chi = -p_{\text{c}}$.

I.5 Modifying the ANM chain stretch to satisfy the principle of material frame-indifference

We begin by showing that the ANM chain stretch (4.4) does not satisfy the principle of material-indifference. Let the constitutive relation that relates Cauchy stress $\mathbf{T}$ to the deformation gradient $\mathbf{F}$ be given by $\mathbf{T} = \hat{\mathbf{T}}(\mathbf{F})$. The principle of material frame-indifference is satisfied if and only if $\hat{\mathbf{T}}(\mathbf{QF}) = \mathbf{Q}\hat{\mathbf{T}}(\mathbf{F})\mathbf{Q}^{\text{T}}$ for all proper orthogonal tensors $\mathbf{Q}$ and admissible deformation gradients $\mathbf{F}$ [6]. Amores et al. [33] construct the first Piola–Kirchhoff stress tensor (up to multiplicative constants that do not affect our conclusions here)

$$\mathbf{P} = \frac{nr_0}{4\pi}\int_{\mathcal{S}^2} \psi'(r_0\hat{\mathbf{r}}\cdot\mathbf{U}\hat{\mathbf{r}})\,\hat{\mathbf{r}}\otimes\hat{\mathbf{r}}\,\mathrm{d}A(\hat{\mathbf{r}}) - p_{\text{c}}\mathbf{F}^{-\text{T}}, \tag{I.1}$$

by averaging over the stress contributions of polymer chains. This leads to the Cauchy stress

$$\hat{\mathbf{T}}(\mathbf{F}) = \frac{nr_0}{4\pi}\int_{\mathcal{S}^2} \psi'(r_0\hat{\mathbf{r}}\cdot\mathbf{U}\hat{\mathbf{r}})\,\hat{\mathbf{r}}\otimes(\mathbf{F}\hat{\mathbf{r}})\,\mathrm{d}A(\hat{\mathbf{r}}) - p_{\text{c}}\mathbf{I}. \tag{I.2}$$

Replacing $\mathbf{F}$ by $\mathbf{QF}$ does not change $\mathbf{U}$, giving

$$\begin{aligned}\hat{\mathbf{T}}(\mathbf{QF}) &= \frac{nr_0}{4\pi}\int_{\mathcal{S}^2} \psi'(r_0\hat{\mathbf{r}}\cdot\mathbf{U}\hat{\mathbf{r}})\,\hat{\mathbf{r}}\otimes(\mathbf{QF}\hat{\mathbf{r}})\,\mathrm{d}A(\hat{\mathbf{r}}) - p_{\text{c}}\mathbf{I}\\ &= \frac{nr_0}{4\pi}\int_{\mathcal{S}^2} \psi'(r_0\hat{\mathbf{r}}\cdot\mathbf{U}\hat{\mathbf{r}})\,\hat{\mathbf{r}}\otimes(\mathbf{F}\hat{\mathbf{r}})\,\mathrm{d}A(\hat{\mathbf{r}})\,\mathbf{Q}^{\text{T}} - p_{\text{c}}\mathbf{I}\end{aligned} \tag{I.3}$$

for any orthogonal tensor $\mathbf{Q}$, so this constitutive relation does not satisfy the principle of material frame-indifference.

The chain stretch

$$\lambda(\hat{\mathbf{r}},\mathbf{B}) = \sqrt{\mathbf{B}} : \hat{\mathbf{r}}\otimes\hat{\mathbf{r}} = \mathbf{V} : \hat{\mathbf{r}}\otimes\hat{\mathbf{r}}, \tag{I.4}$$

obtained by replacing $\mathbf{U}$ in the ANM chain stretch with $\mathbf{V}$ agrees with the ANM chain stretch when the deformation gradient $\mathbf{F}$ has the polar decomposition $\mathbf{F} = \mathbf{U} = \mathbf{V}$, i.e. the rotation tensor $\mathbf{R} = \mathbf{I}$. We show that this chain stretch satisfies the principle of material

frame-indifference. The Cauchy stress is

$$\hat{\mathbf{T}}(\mathbf{F}) = \frac{nr_0}{4\pi}\int_{\mathcal{S}^2} \psi'(r_0\hat{\mathbf{r}}\cdot\mathbf{V}\hat{\mathbf{r}})\,\hat{\mathbf{r}}\otimes(\mathbf{V}\hat{\mathbf{r}})\,\mathrm{d}A(\hat{\mathbf{r}}) - p_{\mathrm{c}}\mathbf{I}, \tag{I.5}$$

where the polar decomposition $\mathbf{F} = \mathbf{VR}$ has been used. When $\mathbf{F}$ is replaced by $\mathbf{QF}$, $\mathbf{V}$ becomes $\mathbf{QVQ}^{\mathrm{T}}$, so

$$\begin{aligned}\hat{\mathbf{T}}(\mathbf{QF}) &= \frac{nr_0}{4\pi}\int_{\mathcal{S}^2} \psi'(r_0\hat{\mathbf{r}}\cdot\mathbf{QVQ}^{\mathrm{T}}\hat{\mathbf{r}})\,\hat{\mathbf{r}}\otimes(\mathbf{QVQ}^{\mathrm{T}}\hat{\mathbf{r}})\,\mathrm{d}A(\hat{\mathbf{r}}) - p_{\mathrm{c}}\mathbf{I}\\ &= \frac{nr_0}{4\pi}\int_{\mathcal{S}^2} \psi'(r_0\hat{\mathbf{r}}'\cdot\mathbf{V}\hat{\mathbf{r}}')\,(\mathbf{Q}\hat{\mathbf{r}}')\otimes(\mathbf{QV}\hat{\mathbf{r}}')\,\mathrm{d}A(\hat{\mathbf{r}}') - p_{\mathrm{c}}\mathbf{I}\\ &= \mathbf{Q}\hat{\mathbf{T}}(\mathbf{F})\mathbf{Q}^{\mathrm{T}},\end{aligned} \tag{I.6}$$

where in the second equality the change of variables $\hat{\mathbf{r}}' = \mathbf{Q}^{\mathrm{T}}\hat{\mathbf{r}}$ has been used. Therefore, the chain stretch relation (I.4) results in a constitutive model that satisfies the principle of material frame indifference. We have therefore shown that the correct form of the ANM chain stretch for general deformations is (I.4), which agrees with the original (4.4) in uniaxial tension, pure shear and equibiaxial tension.

I.6 Different choices of the hybrid-exponential interpolation parameter

Fig. I.1 shows the predictions of constitutive models based on the hybrid-exponential stretch with $\hat{p}(\frac{1}{2}\sqrt{N}) = \frac{1}{4}, \frac{3}{4}$, alongside experimental stress–stretch data from Treloar [53]. The predictions for $\hat{p}(\frac{1}{2}\sqrt{N}) = \frac{1}{2}$ are shown in Fig. 4.3. The corresponding mean squared errors and optimal parameter values are shown in Table I.1. The prediction accuracy of the constitutive model based on the hybrid-exponential stretch is not significantly affected by choice of $\hat{p}(\frac{1}{2}\sqrt{N})$ over the range $\frac{1}{4} \le \hat{p}(\frac{1}{2}\sqrt{N}) \le \frac{3}{4}$.

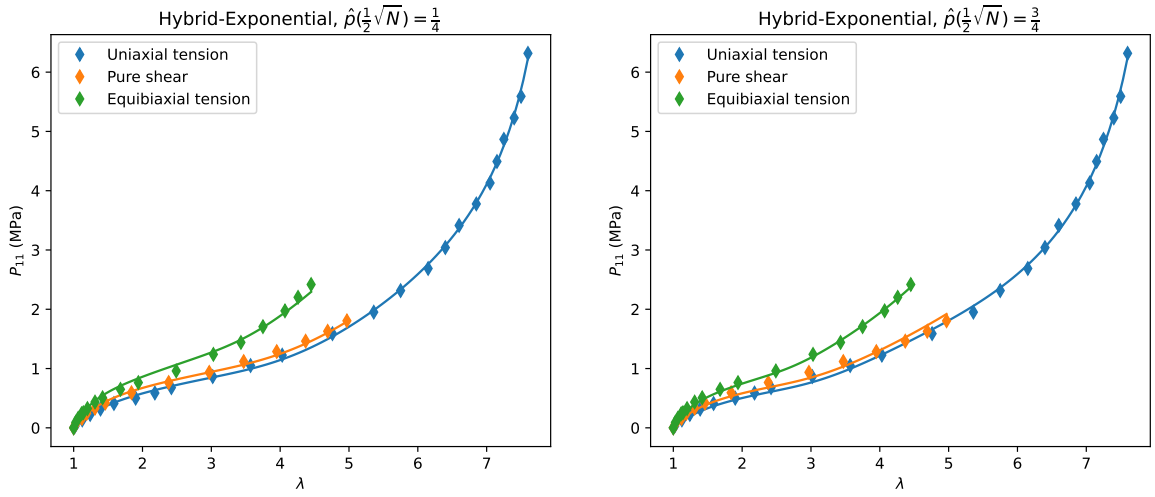


Fig. I.1 Stress–stretch relations predicted by the hybrid-exponential model for different choices of $\hat{p}(\frac{1}{2}\sqrt{N})$, when model parameters are fitted to vulcanised natural rubber data from Treloar [53]. The model predictions are shown as solid lines, and the experimental data are shown as diamonds.

$\hat{p}(\frac{1}{2}\sqrt{N})$	$\bar{\mu}$ (MPa)	N	Mean squared error (MPa^2)			
			Overall	Uniaxial tension	Pure shear	Equibiaxial tension
1/4	0.985	66.6	0.00234	0.00277	0.000767	0.00301
1/2	0.908	67.0	0.00247	0.00361	0.00196	0.00121
3/4	0.852	67.1	0.00313	0.00383	0.00429	0.00113

Table I.1 Model parameters and mean squared error (MSE) for constitutive models based on the hybrid-exponential stretch corresponding to different values of $\hat{p}(\frac{1}{2}\sqrt{N})$, when fitted to experimental data from Treloar [53].

I.7 Comparison of three-parameter constitutive models

If we do not fix a value for $\hat{p}(\frac{1}{2}\sqrt{N})$ for the hybrid-exponential stretch *a priori*, then the constitutive model based on this chain stretch relation has an additional fitting parameter. Here we compare this three-parameter model to previously proposed constitutive models with the same number of parameters: the Gent+Gent model by Pucci and Saccomandi [78], the extended 8-chain model by Dal et al. [36] and the hybrid-ANM model from Amores et al. [57] with the single-chain Helmholtz free energy described in Section 3.3. We chose these models for comparison because the Gent+Gent model is a well-known phenomenological model known for good accuracy, and the extended 8-chain model was found by Dal et al. [7]

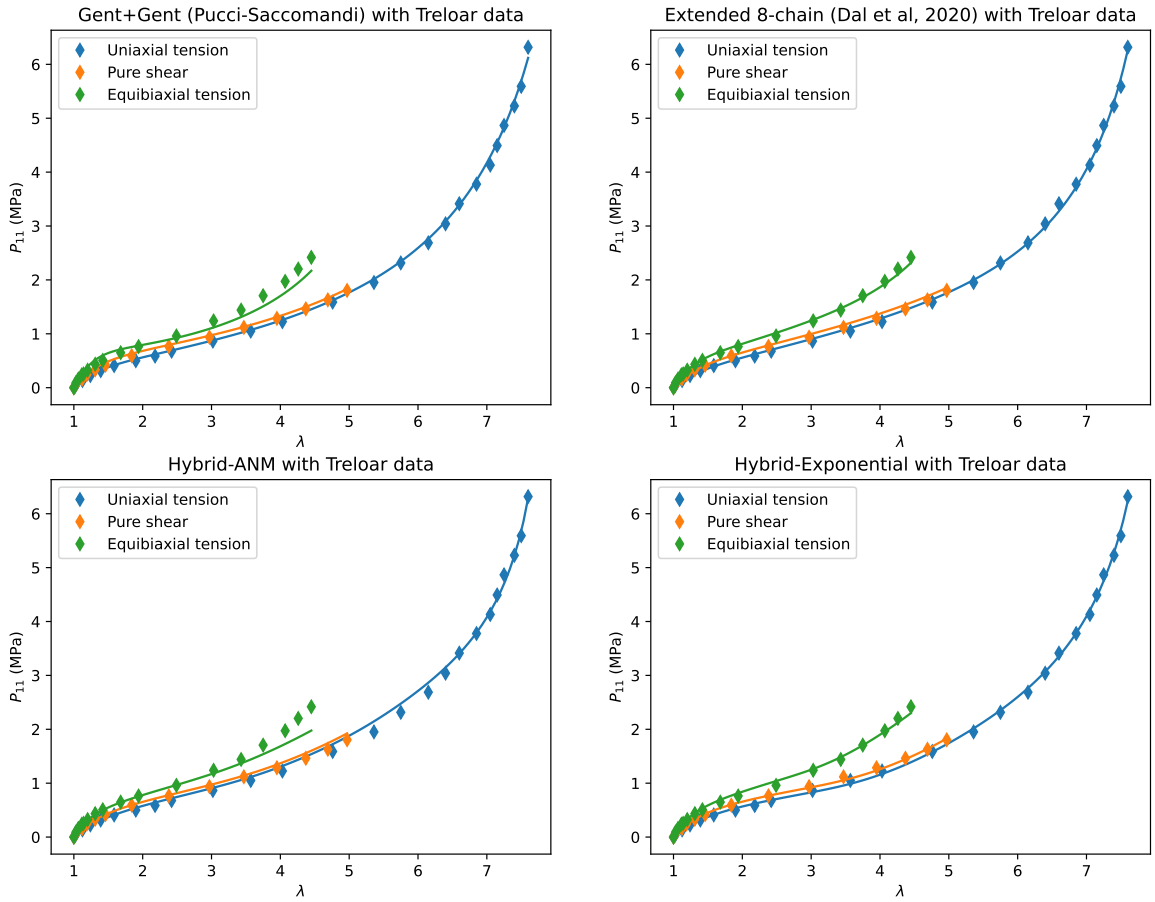


Fig. I.2 Stress–stretch relations predicted by three-parameter constitutive models when model parameters are fitted to vulcanised natural rubber data from Treloar [53]. The model predictions are shown as solid lines, and the experimental data are shown as diamonds.

to have the best fit to the dataset of Treloar [53] among the three-parameter models tested in their extensive review. We also include the hybrid-ANM model in this comparison because its approach to modelling the chain orientation distribution is similar to our approach in the hybrid-exponential stretch, and it was proposed after the publication of the review by Dal et al. [7].

The Gent+Gent model has strain energy function

$$W_{\mathrm{GG}} = \frac{\bar{\mu}}{2} N \log\left(1 - \frac{I_1 - 3}{N}\right) + C \log\left(\frac{I_2}{3}\right), \tag{I.7}$$

where $I_1 = \mathrm{tr}(\mathbf{B})$, $I_2 = \frac{1}{2}\left(I_1^2 - \mathrm{tr}\left(\mathbf{B}^2\right)\right)$, and $\bar{\mu}, N, C$ are the three parameters to be fitted to the data. We have replaced the original parameter name J_m by N for consistency with the other

	μ (MPa)	N	Third parameter	Mean squared error (MPa^2)			
				Overall	Uni.	Shear	Equi.
G+G	0.248	79.6	$C = 0.286$ MPa	0.00694	0.00441	**0.000831**	0.0157
Ext. 8-chain	0.268	25.2	$C = 0.442$ MPa	0.00262	0.00365	0.00175	**0.00185**
H-ANM	0.348	62.9	$P_0 = 0.776$ MPa	0.0117	0.00766	0.00304	0.0247
H-Exp.	0.958	66.8	$\hat{p}(\frac{1}{2}\sqrt{N}) = 0.323$	**0.00223**	**0.00305**	0.000853	0.00214

Table I.2 Model parameters and mean squared error (MSE) for three-parameter models when fitted to experimental data from Treloar [53]. The smallest MSE value in each column is displayed in bold font.
Abbreviations: "G+G" is "Gent+Gent"; "Ext." is "Extended"; "H-" is "Hybrid-"; "Exp." is "Exponential"; "Uni." is "Uniaxial tension"; "Equi." is "Equibiaxial tension".

models considered here, because this parameter controls the Gent+Gent model's singularity in stress.

The extended 8-chain model has strain energy function based on that of the Arruda–Boyce model, with an additional term to model the constraining effects of neighbouring polymer chains. These effects are captured by regarding each polymer chain as being confined to a tube, whose cross-sectional area depends on the material deformation. This tube-constraint term depends on the second strain invariant I_2, leading to a strain energy function

$$W_{\mathrm{E8}} = W_{\mathrm{AB}} + C\left(\left(\frac{I_2}{3}\right)^{\frac{1}{3}} - 1\right), \tag{I.8}$$

where W_{AB} is the Arruda–Boyce strain energy function depending on parameters $\bar{\mu}, N$, and the third parameter C controls the strength of the tube constraint effects.

The third parameter P_0 in the hybrid-ANM model comes from their modified chain force $f = f_0 + f_{\mathrm{FJC}}$, where f_{FJC} is the force (2.7) for a freely jointed chain and f_0 is a constant offset that is fitted to experimental data. The fitting parameter P_0 is related to f_0 by $P_0 = nr_0f_0$, where n is the number density of chains and r_0 is the unstretched chain end-to-end length, as defined in Section 2.3.

The predictions of these models are plotted alongside experimental stress–stretch data from Treloar [53] in Fig. I.2, and the corresponding optimal parameter and MSE values are given in Table I.2. These results show that the hybrid-exponential model has the most accurate fit to the data, followed closely by the extended 8-chain model. Comparing the overall MSE of the extended 8-chain model to that of the *two-parameter* hybrid-exponential model in Table 4.2 shows that, even without treating $\hat{p}(\frac{1}{2}\sqrt{N})$ as an additional fitting parameter, the hybrid-exponential model achieves a slightly more accurate fit than the best three-parameter

model tested by Dal et al. [7]. The conclusion of Appendix I.6 that the fitting performance of the hybrid-exponential stretch is not significantly affected by variations in the value of $\hat{p}(\frac{1}{2}\sqrt{N})$ over the range $\frac{1}{4}$ to $\frac{3}{4}$ is supported by the relatively small decrease in overall MSE when $\hat{p}(\frac{1}{2}\sqrt{N})$ is optimised by fitting to the data.

Part II

Nonlinear active elasticity

Chapter 6

Introduction to active elastic solids

The study of active matter is a broad and rapidly developing subfield of physics that investigates systems consisting of many self-driven, energy-consuming constituents [79, 80]. Such systems occur at vastly different length scales, from the cytoskeleton of a biological cell [81] to agitated rods [82] and the flocking of birds [83]. The interest in active matter is driven by the plethora of fascinating phenomena that would be ruled out by conservative interactions between particles in passive systems. Such phenomena can arise from self-propulsion [84], such as in suspensions of swimming bacteria [85, 86], phoretic active colloids [87, 88], Quincke rollers [89], colloids with active slip [90, 91] and other artificial swimmers [92, 93]. In addition, non-reciprocal interactions, such as between rotating bacteria [94] and starfish embryos [95] or between chemical species [96], can lead to the emergence of novel collective behaviour. In some cases, the energy supply to individual constituents can be made apparent by work extraction from the system, such as the spontaneous rotation of an asymmetric ratchet immersed in a bath of swimming bacteria [97].

Some recent advances in active matter research involve exploration of the interplay between activity and elasticity. Active solids are elastic materials formed by connecting active components, resulting in novel mechanical behaviour, such as self-sustaining oscillations [10, 98–100], and active shape changes [99]. Here we focus on a class of active solid in which stress depends only on the solid's current state of deformation.

6.1 Active matter and odd elasticity

In linear elasticity, the mechanical response of a solid is characterised by the elasticity tensor, a linear transformation that maps strain to stress. The elasticity tensor is a fourth-order tensor, mapping the second-order tensor $\mathbf{A}$ to the second-order tensor $\mathsf{C}[\mathbf{A}]$. Passive elastic solids are subject to the restriction that closed cycles of deformation cannot involve non-zero net work,

implying that they are hyperelastic [6]. Therefore, their stress–strain relations can be derived from a strain energy function W, as expressed by (2.4). This restriction implies that the elasticity tensor C satisfies *major symmetry*, i.e. $\mathbf{A}_2 : \mathsf{C}[\mathbf{A}_1] = \mathbf{A}_1 : \mathsf{C}[\mathbf{A}_2]$ for all second-order tensors $\mathbf{A}_1, \mathbf{A}_2$ [6]. Equivalently, C has major symmetry if its components satisfy $C_{ijkl} = C_{klij}$. In contrast, this restriction does not apply to active solids because energy injection at the micro-scale could result in non-zero work output in closed cycles of deformation [10]. Therefore, the elasticity tensor for an active solid generically has an antisymmetric, or odd, part, corresponding to elastic moduli not present in passive elastic solids. This *odd elasticity* gives rise to novel mechanical behaviour in active solids, such as transverse deflection during uniaxial compression of isotropic solids and elastic wave propagation in the overdamped regime [10]. Odd elasticity has not only been observed in biological systems, such as living crystals formed by starfish embryos [95] and muscles [101], but has also been harnessed to design robotic metamaterials capable of adaptive locomotion [100].

The strain tensor in linear elasticity is the symmetric part of the displacement gradient tensor $\nabla^{\mathrm{s}}\mathbf{u} = \operatorname{sym}\nabla_{\mathbf{x}_0}\mathbf{u}$, where the field $\mathbf{u}(\mathbf{x}_0)$ gives the displacement of a material point from its position $\mathbf{x}_0$ in the undeformed state $\mathcal{B}_0$, and the symmetrisation operator is defined by $\operatorname{sym}\mathbf{A} = \frac{1}{2}(\mathbf{A} + \mathbf{A}^{\mathrm{T}})$. Scheibner et al. [10] showed that the most general elasticity tensor C for two-dimensional isotropic solids maps strain $\nabla^{\mathrm{s}}\mathbf{u}$ to

$$\mathsf{C}[\nabla^{\mathrm{s}}\mathbf{u}] = (B-\mu)\,(\operatorname{tr}\nabla^{\mathrm{s}}\mathbf{u})\mathbf{I} + 2\mu\nabla^{\mathrm{s}}\mathbf{u} + K^o((\mathbf{S}_2 : \nabla^{\mathrm{s}}\mathbf{u})\mathbf{S}_1 - (\mathbf{S}_1 : \nabla^{\mathrm{s}}\mathbf{u})\mathbf{S}_2) + A(\operatorname{tr}\nabla^{\mathrm{s}}\mathbf{u})\mathbf{W}, \tag{6.1}$$

where B, μ, K^o, A are elastic moduli, and the tensors $\mathbf{S}_1$, $\mathbf{S}_2$,$\mathbf{W}$ have matrix representations

$$[\mathbf{S}_1] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad [\mathbf{S}_2] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad [\mathbf{W}] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \tag{6.2}$$

respectively, with respect to some orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2\}$. We have expressed (6.1) using notation different from that used by Scheibner et al. [10] to facilitate generalising this result to the nonlinear regime. The elasticity tensor in (6.1) defines a linear stress–strain relation in which B and μ are the usual bulk and shear moduli, respectively, that characterise the elasticity of isotropic passive solids. The additional moduli K^o and A break major symmetry of the elasticity tensor, and are manifestations of activity. Because an isotropic material that is also symmetric under reflections necessarily has $K^o = 0, A = 0$ [102], both these moduli break chiral symmetry. The odd modulus K^o controls antisymmetric coupling between the two independent shear stresses/strains $\mathbf{S}_1$ and $\mathbf{S}_2$, whereas A couples isotropic dilations to torques. The presence of the antisymmetric tensor $\mathbf{W}$ corresponding to an anticlockwise rotation by $\frac{1}{2}\pi$ is a manifestation of broken chiral symmetry. Although we expressed the

term proportional to K^o in (6.1) using the two independent shears $\mathbf{S}_1$ and $\mathbf{S}_2$, this term can be expressed in terms of any two linearly independent tensors that are traceless and symmetric. We show this using the identities $\mathbf{W}\mathbf{S}_1 = \mathbf{S}_2$ and $\mathbf{W}\mathbf{S}_2 = -\mathbf{S}_1$, which imply that the term proportional to K^o can be expressed as

$$K^o((\mathbf{S}_2 : \nabla^{\mathrm{s}}\mathbf{u})\mathbf{S}_1 - (\mathbf{S}_1 : \nabla^{\mathrm{s}}\mathbf{u})\mathbf{S}_2) = -2K^o \operatorname{sym}(\mathbf{W}\nabla^{\mathrm{s}}\mathbf{u}), \tag{6.3}$$

where the right-hand side does not rely on the specific choice of basis tensors $\mathbf{S}_1, \mathbf{S}_2$. Because the odd moduli K^o and A involve broken chiral symmetry, they are absent in three-dimensional isotropic materials, which are necessarily achiral. Therefore, these materials cannot exhibit odd elasticity in the linear regime [10].

6.2 Active elasticity in the nonlinear regime

In the context of nonlinear elasticity, discussion of closed cycles of deformation involving non-zero net work pre-dates the use of odd elasticity to study active solids. Carroll [43] asked 'Must elastic materials be hyperelastic?' and showed that a non-hyperelastic material undergoing a closed cycle of deformation would perform non-zero net work on its surroundings. Writing before the development of research into active solids, Carroll asserted that this was 'physically unacceptable' and concluded that 'elastic materials must be hyperelastic.' However, recent advances in active matter physics [95, 100, 101] have shown that non-hyperelastic materials exist in both biology and engineering, motivating Yavari and Goriely [11] to revisit non-hyperelastic Cauchy elasticity as a natural generalisation of odd elasticity to the nonlinear regime.

In this part of the thesis, we aim to characterise the soft elasticity of active solids by identifying elastic moduli that control the nonlinear response of these materials, and understand how these moduli control active phenomenology. Because many active solids are soft and undergo large deformations [100, 101], such advances would point towards new design principles for soft materials with novel functionality. In Chapter 7, we use weakly nonlinear isotropic elasticity to identify a novel odd elastic modulus corresponding to stress contributions from terms quadratic in strain. Unlike the odd moduli in linear elasticity, this modulus does not rely on broken chiral symmetry, so three-dimensional isotropic materials can exhibit odd elasticity in the nonlinear regime if this modulus is non-zero. Importantly, we show that the presence of this modulus means that cycles of deformation corresponding to zero net work in the linear regime can result in non-zero net work in the non-linear regime. Furthermore, designing soft active solids requires going beyond purely continuum

descriptions to specific microscopic realisations. We therefore present two- and three-dimensional random networks that exhibit this form of odd elasticity, using both analytical results and numerical simulations to establish that the elastic response of these networks is not hyperelastic, in Chapter 8.

In the remainder of this chapter, we review concepts and define notation for elastic constitutive relations that are not hyperelastic, for use in later chapters. Here we build on the concepts and notation of nonlinear elasticity given in Section 2.1. A generic elastic constitutive relation gives the first Piola–Kirchhoff stress as

$$\mathbf{P} = \hat{\mathbf{P}}(\mathbf{F}), \tag{6.4}$$

where $\hat{\mathbf{P}}$ is a function defined on the set of deformation gradients $\mathbf{F}$ and need not be equal to the derivative of some scalar function W. This constitutive relation can be expressed in terms of the Cauchy stress $\mathbf{T}$ using (2.3), i.e. $\mathbf{T} = \hat{\mathbf{T}}(\mathbf{F})$, where $\hat{\mathbf{T}}(\mathbf{F}) = (\det\mathbf{F})^{-1}\hat{\mathbf{P}}(\mathbf{F})\mathbf{F}^{\mathrm{T}}$ [6]. In terms of the second Piola–Kirchhoff stress $\mathbf{S} = \mathbf{F}^{-1}\mathbf{P}$, the constitutive relation is $\mathbf{S} = \hat{\mathbf{S}}(\mathbf{F})$, where [21]

$$\hat{\mathbf{S}}(\mathbf{F}) = \mathbf{F}^{-1}\hat{\mathbf{P}}(\mathbf{F}) = (\det\mathbf{F})\mathbf{F}^{-1}\hat{\mathbf{T}}(\mathbf{F})\mathbf{F}^{-\mathrm{T}}. \tag{6.5}$$

For a path of deformations represented by the family of deformation gradients $\mathbf{F} = \hat{\mathbf{F}}(\xi)$ parametrised by $\xi \in [\xi_0, \xi_1]$, the work done *on* the elastic material per unit volume is [6]

$$\mathcal{W}\left[\hat{\mathbf{F}}\right] = \int_{\xi_0}^{\xi_1} \hat{\mathbf{P}}\left(\hat{\mathbf{F}}(\xi)\right) : \hat{\mathbf{F}}'(\xi)\,\mathrm{d}\xi, \tag{6.6}$$

where $\hat{\mathbf{F}}'(\xi) = \frac{\mathrm{d}\hat{\mathbf{F}}}{\mathrm{d}\xi}(\xi)$, provided there is no change in the kinetic energy of the material between its initial and final states. If a constitutive relation $\hat{\mathbf{P}}(\mathbf{F})$ satisfies $\mathcal{W}\left[\hat{\mathbf{F}}\right] = 0$ for all closed loops in deformation space, i.e. for all paths satisfying $\hat{\mathbf{F}}(\xi_0) = \hat{\mathbf{F}}(\xi_1)$, then we can define a strain energy function W for this constitutive relation, because the integral in (6.6) depends only on the initial and final deformation gradients, $\hat{\mathbf{F}}(\xi_0)$ and $\hat{\mathbf{F}}(\xi_1)$. This path-independence means that the value of the strain energy function $W(\mathbf{F})$ at each deformation gradient $\mathbf{F}$ can be defined as the integral in (6.6), with path of integration beginning at some fixed reference deformation $\hat{\mathbf{F}}(\xi_0) = \mathbf{F}_0$ and ending at $\hat{\mathbf{F}}(\xi_1) = \mathbf{F}$. This reasoning shows that a material satisfies $\mathcal{W}\left[\hat{\mathbf{F}}\right] = 0$ for all closed loops in deformation space, if and only if the material is hyperelastic. Therefore, a material whose constitutive relation (6.4) cannot be obtained from a strain energy function necessarily has closed loops in deformation space along which non-zero net work is extracted or absorbed by the material.

The functional form of any elastic constitutive relation is constrained by the principle of material frame-indifference [6, 9]. This general principle requires that physical behaviour

be independent of the frame of reference [9]. To articulate how this principle applies to constitutive relations, we first make the following definitions. For a d-dimensional Euclidean vector space $\mathcal{V}_d$, as defined in Section 2.1.1, the set of proper orthogonal tensors acting on $\mathcal{V}_d$ is

$$\mathrm{SO}(d) = \{\mathbf{Q} \in \mathtt{L}(\mathcal{V}_d) : \det\mathbf{Q} = 1, \mathbf{Q}^{\mathrm{T}} = \mathbf{Q}^{-1}\}, \tag{6.7}$$

which is a proper subset of the set of orthogonal tensors

$$\mathrm{O}(d) = \{\mathbf{Q} \in \mathtt{L}(\mathcal{V}_d) : \mathbf{Q}^{\mathrm{T}} = \mathbf{Q}^{-1}\}. \tag{6.8}$$

Physically, the special orthogonal group $\mathrm{SO}(d)$ represents all possible rotations, whereas the orthogonal group $\mathrm{O}(d)$ includes combinations of rotations and reflections. Applying the principle of material-frame indifference to changes of reference frame specified by rotations $\mathbf{Q} \in \mathrm{SO}(d)$, every elastic constitutive relation must satisfy

$$\hat{\mathbf{T}}(\mathbf{QF}) = \mathbf{Q}\hat{\mathbf{T}}(\mathbf{F})\mathbf{Q}^{\mathrm{T}} \tag{6.9}$$

for any $\mathbf{Q} \in \mathrm{SO}(d)$ and any deformation gradient $\mathbf{F}$. In terms of the second Piola–Kirchhoff stress, (6.9) takes the form

$$(\det\mathbf{QF})^{-1}(\mathbf{QF})\hat{\mathbf{S}}(\mathbf{QF})(\mathbf{QF})^{\mathrm{T}} = (\det\mathbf{F})^{-1}\mathbf{QF}\,\hat{\mathbf{S}}(\mathbf{F})\mathbf{F}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}},$$

which implies

$$\hat{\mathbf{S}}(\mathbf{QF}) = \hat{\mathbf{S}}(\mathbf{F}) \tag{6.10}$$

for all $\mathbf{Q} \in \mathrm{SO}(d)$. Using the polar decomposition $\mathbf{F} = \mathbf{RU}$, we set $\mathbf{Q} = \mathbf{R}^{\mathrm{T}}$ in (6.10) to obtain $\hat{\mathbf{S}}(\mathbf{F}) = \hat{\mathbf{S}}(\mathbf{U})$. Therefore, defining the function $\tilde{\mathbf{S}}$ to satisfy $\tilde{\mathbf{S}}(\mathbf{C}) = \hat{\mathbf{S}}(\sqrt{\mathbf{C}})$, we see that the principle of material frame-indifference implies that any elastic constitutive relation takes the form

$$\mathbf{S} = \tilde{\mathbf{S}}(\mathbf{C}), \tag{6.11}$$

where $\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F} = \mathbf{U}^2$ is the right Cauchy–Green deformation tensor. The function $\tilde{\mathbf{S}}$ is defined on the set of positive-definite, symmetric tensors

$$\mathrm{Psym}(d) = \{\mathbf{A} \in \mathtt{L}(\mathcal{V}_d) : \mathbf{A}^{\mathrm{T}} = \mathbf{A},\, \mathbf{v}\cdot\mathbf{Av} > 0 \text{ for all } \mathbf{v} \neq 0\}. \tag{6.12}$$

With this representation, the work done on an elastic material per unit volume (6.6) can be expressed in terms of an integral along a path $\mathbf{C} = \hat{\mathbf{C}}(\xi)$, $\xi \in [\xi_0, \xi_1]$ in $\mathrm{Psym}(d)$, given by

$$\widetilde{\mathcal{W}}\left[\hat{\mathbf{C}}\right] = \frac{1}{2}\int_{\xi_0}^{\xi_1} \tilde{\mathbf{S}}(\hat{\mathbf{C}}(\xi)) : \hat{\mathbf{C}}'(\xi)\,\mathrm{d}\xi, \tag{6.13}$$

provided $\tilde{\mathbf{S}}(\mathbf{C})$ is symmetric for tensors $\mathbf{C} \in \mathrm{Psym}(d)$.

Chapter 7

Phenomenology of nonlinear active elasticity

In this chapter, we study the phenomenology associated with nonlinear elastic constitutive relations that are not hyperelastic, from a purely continuum perspective. Therefore, our results here are independent of microscopic realisation. We focus on two implications of the lack of hyperelasticity in an elastic constitutive relation: the presence of odd moduli when linearising stress–strain relations about deformed reference states, and work extraction from closed cycles of deformation. To characterise this phenomenology in terms of a small number of parameters, we consider weakly nonlinear isotropic elasticity and identify odd moduli within this framework. Because the general form of nonlinear isotropic elasticity for two-dimensional materials has not previously been established, we prove a representation theorem that generalises the linear-elastic constitutive relation (6.1) from Scheibner et al. [10]. We identify an odd modulus d_o in the weakly nonlinear theory corresponding to terms that are quadratic in strain. Unlike the odd moduli K^o, A in (6.1) for two-dimensional isotropic linear elasticity, this odd modulus does not require chiral symmetry to be broken, so it can be present in both two- and three-dimensional isotropic elasticity. We show that this new odd modulus enables work to be extracted using deformation cycles that would lead to zero work extraction in linear odd elasticity.

7.1 Odd elasticity about deformed states

Elasticity tensors characterise the linear stress response of a material to small strains measured from a given reference state. In nonlinear elasticity, an elasticity tensor is obtained by taking the Fréchet derivative of a stress measure with respect to deformation gradient $\mathbf{F}$ [6], where

this derivative is defined in Appendix I.1. In general, elasticity tensors depend on the reference state about which the linearisation is performed. In standard linear elasticity, the derivative is evaluated at the undeformed state $\mathbf{F} = \mathbf{I}$. Although we can define three elasticity tensors, corresponding to the derivatives of the Cauchy stress as well as of the first and second Piola–Kirchhoff stresses, these derivatives are all equal at $\mathbf{F} = \mathbf{I}$ if this state is stress-free [6, 21]. However, these elasticity tensors are not equal in general, if the stress–strain relation is linearised about a stressed reference state.

Braverman et al. [103], in their Supplemental Material, note that an elastic constitutive relation is hyperelastic if and only if the elasticity tensor obtained from linearising the first Piola–Kirchhoff stress has major symmetry. In contrast, the elasticity tensor obtained from the Cauchy stress does not necessarily satisfy major symmetry at reference states $\mathbf{F} \neq \mathbf{I}$ even when the corresponding constitutive relation is hyperelastic. Therefore, the equivalence established in standard linear elasticity between major symmetry of the elasticity tensor and hyperelasticity of the constitutive relation only generalises to some definitions of the elasticity tensor. Because we are interested in odd elasticity as a description of active materials from which non-zero work can be extracted in closed cycles of deformation, we only consider odd elastic moduli that are associated with lack of hyperelasticity.

To identify these odd elastic moduli, we first establish notation for linearising constitutive relations about deformed reference states. In this section, we give a precise definition of an elasticity tensor for which odd moduli are associated with the lack of hyperelasticity, generalising the odd moduli in Scheibner et al. [10] to the nonlinear regime. Although these ideas are present in Braverman et al. [103], we include this section to give more detail and to provide definitions and notation for the rest of this chapter. In this part of the thesis, we consider elastic bodies for which stress-free states $\mathcal{B}_0$ exist, and refer to such states as undeformed states, corresponding to $\mathbf{F} = \mathbf{I}$. A deformed reference state $\mathcal{B}_1$ is obtained by mapping $\mathbf{x}_0 \in \mathcal{B}_0$ to $\mathbf{x}_1 = \chi_1(\mathbf{x}_0) \in \mathcal{B}_1$, and corresponds to deformation gradient $\mathbf{F}_1(\mathbf{x}_0) = \nabla_{\mathbf{x}_0}\chi_1(\mathbf{x}_0)$. Deformations from the reference state $\mathcal{B}_1$ are characterised by the mapping $\bar{\chi}$, which maps $\mathbf{x}_1 \in \mathcal{B}_1$ to $\mathbf{x} = \bar{\chi}(\mathbf{x}_1) \in \mathcal{B}$ in the final deformed state, and are associated with the deformation gradient $\bar{\mathbf{F}}(\mathbf{x}_1) = \nabla_{\mathbf{x}_1}\bar{\chi}(\mathbf{x}_1)$. The strain tensor relative to the deformed reference state $\mathcal{B}_1$ is the gradient $\nabla_{\mathbf{x}_1}\mathbf{u}(\mathbf{x}_1)$ of the displacement field $\mathbf{u}(\mathbf{x}_1) = \bar{\chi}(\mathbf{x}_1) - \mathbf{x}_1$, and satisfies $\bar{\mathbf{F}} = \mathbf{I} + \nabla_{\mathbf{x}_1}\mathbf{u}$. By the chain rule, the deformation gradient with respect to the undeformed state $\mathcal{B}_0$ is

$$\mathbf{F}(\mathbf{x}_0) = \nabla_{\mathbf{x}_0}\bar{\chi}(\chi_1(\mathbf{x}_0)) = \bar{\mathbf{F}}(\chi_1(\mathbf{x}_0))\mathbf{F}_1(\mathbf{x}_0). \tag{7.1}$$

The first Piola–Kirchhoff stress with respect to the deformed reference state $\mathcal{B}_1$ is

$$\mathbf{P}_1 = (\det\bar{\mathbf{F}})\mathbf{T}\bar{\mathbf{F}}^{-\mathrm{T}}, \tag{7.2}$$

which is the Cauchy stress $\mathbf{T}$ pulled back from the final state $\mathcal{B}$ so that it maps vectors representing oriented areas in the deformed reference state $\mathcal{B}_1$ to forces in the final state $\mathcal{B}$. This stress measure is relevant to experimental measurements in which the dimensions of the sample in the deformed reference state $\mathcal{B}_1$ are known. The response of the first Piola–Kirchhoff stress $\mathbf{P}_1$ to deformations $\bar{\mathbf{F}}$ from the reference state $\mathcal{B}_1$ is given by the function

$$\bar{\mathbf{P}}_1(\bar{\mathbf{F}}) = (\det\bar{\mathbf{F}})\hat{\mathbf{T}}(\bar{\mathbf{F}}\mathbf{F}_1)\bar{\mathbf{F}}^{-\mathrm{T}}. \tag{7.3}$$

For small strains $\nabla_{\mathbf{x}_1}\mathbf{u}$ about the reference state $\mathcal{B}_1$, we can expand $\bar{\mathbf{P}}_1(\bar{\mathbf{F}})$ about $\bar{\mathbf{F}} = \mathbf{I}$ to obtain

$$\bar{\mathbf{P}}_1(\mathbf{I}+\nabla_{\mathbf{x}_1}\mathbf{u}) = \hat{\mathbf{T}}(\mathbf{F}_1) + \mathrm{D}_{\bar{\mathbf{F}}}\bar{\mathbf{P}}_1(\mathbf{I})[\nabla_{\mathbf{x}_1}\mathbf{u}] + o(||\nabla_{\mathbf{x}_1}\mathbf{u}||), \tag{7.4}$$

where $o(||\nabla_{\mathbf{x}_1}\mathbf{u}||)$ indicates a function that approaches zero faster than $||\nabla_{\mathbf{x}_1}\mathbf{u}||$ as $||\nabla_{\mathbf{x}_1}\mathbf{u}|| \to 0$. More precisely, a function g has limiting behaviour $o(y)$ if $\lim_{y\to 0}\frac{g(y)}{y} = 0$. The expansion (7.4) shows that the elastic response for small strains $\nabla_{\mathbf{x}_1}\mathbf{u}$ about the reference state $\mathcal{B}_1$ is characterised by the elasticity tensor

$$\mathsf{C}_{\mathbf{F}_1} = \mathrm{D}_{\bar{\mathbf{F}}}\bar{\mathbf{P}}_1(\mathbf{I}) = \mathrm{D}_{\bar{\mathbf{F}}}\left((\det\bar{\mathbf{F}})\hat{\mathbf{T}}(\bar{\mathbf{F}}\mathbf{F}_1)\bar{\mathbf{F}}^{-\mathrm{T}}\right)(\mathbf{I}). \tag{7.5}$$

In Appendix II.1, we show that the antisymmetric part of this elasticity tensor, defined to be $\operatorname{asym}\mathsf{C}_{\mathbf{F}_1} = \frac{1}{2}(\mathsf{C}_{\mathbf{F}_1} - \mathsf{C}_{\mathbf{F}_1}^{\mathrm{T}})$, satisfies

$$\operatorname{asym}\mathsf{C}_{\mathbf{F}_1}[\mathbf{H}] = \frac{2}{\det\mathbf{F}_1}\mathbf{F}_1\left(\operatorname{asym}\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}\left(\mathbf{F}_1^{\mathrm{T}}\mathbf{F}_1\right)\right)[\mathbf{F}_1^{\mathrm{T}}(\operatorname{sym}\mathbf{H})\mathbf{F}_1]\mathbf{F}_1^{\mathrm{T}}, \tag{7.6}$$

for $\mathbf{H} \in \mathrm{L}(\mathcal{V}_d)$, where $\tilde{\mathbf{S}}$ is the stress function (6.11) that maps the right Cauchy–Green deformation tensor $\mathbf{C}$ to the second Piola–Kirchhoff stress in the *undeformed* state. We can now connect the presence of odd elastic moduli in $\mathsf{C}_{\mathbf{F}_1}$ to non-hyperelasticity. In Appendix II.2, we show that an elastic constitutive relation (6.11) is hyperelastic if and only if the derivative of the stress function $\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}(\mathbf{C})$ has major symmetry for all $\mathbf{C} \in \operatorname{Psym}(d)$. Therefore, (7.6) shows that there are no odd elastic moduli associated with $\mathsf{C}_{\mathbf{F}_1}$ for all deformed reference states $\mathbf{F}_1$ (equivalently, $\operatorname{asym}\mathsf{C}_{\mathbf{F}_1} = 0$ for all $\mathbf{F}_1$) if and only if the corresponding constitutive relation is hyperelastic. This result shows that the antisymmetric part of the elasticity tensor $\mathsf{C}_{\mathbf{F}_1}$ appropriately generalises the odd elastic moduli used by Scheibner et al. [10] for linear elasticity.

7.2 Isotropic elasticity

An isotropic elastic material has constitutive relation $\mathbf{T} = \hat{\mathbf{T}}(\mathbf{F})$ that satisfies

$$\hat{\mathbf{T}}(\mathbf{FQ}) = \hat{\mathbf{T}}(\mathbf{F}) \tag{7.7}$$

for all proper orthogonal tensors $\mathbf{Q} \in \mathrm{SO}(d)$ and all deformation gradients $\mathbf{F}$. Setting $\mathbf{Q} = \mathbf{R}^{\mathrm{T}}$, where $\mathbf{R}$ is the rotation tensor from the polar decomposition $\mathbf{F} = \mathbf{VR}$, gives $\hat{\mathbf{T}}(\mathbf{F}) = \hat{\mathbf{T}}(\mathbf{V})$. Therefore, we can define $\tilde{\mathbf{T}}(\mathbf{B}) = \hat{\mathbf{T}}(\sqrt{\mathbf{B}})$ to express the isotropic constitutive relation in the form

$$\mathbf{T} = \tilde{\mathbf{T}}(\mathbf{B}), \tag{7.8}$$

where $\mathbf{B} = \mathbf{FF}^{\mathrm{T}} = \mathbf{V}^2$. By the principle of material frame-indifference,

$$\tilde{\mathbf{T}}(\mathbf{QBQ}^{\mathrm{T}}) = \mathbf{Q}\tilde{\mathbf{T}}(\mathbf{B})\mathbf{Q}^{\mathrm{T}} \tag{7.9}$$

for all $\mathbf{Q} \in \mathrm{SO}(d)$, because replacing $\mathbf{F}$ by $\mathbf{QF}$ is equivalent to replacing $\mathbf{B}$ by $\mathbf{QBQ}^{\mathrm{T}}$. Expressing the constitutive relation in terms of the second Piola–Kirchhoff stress tensor and using the isotropy requirement (7.7) implies that

$$\begin{aligned}\tilde{\mathbf{S}}(\mathbf{Q}^{\mathrm{T}}\mathbf{CQ}) = \hat{\mathbf{S}}(\mathbf{FQ}) &= (\det\mathbf{FQ})(\mathbf{FQ})^{-1}\hat{\mathbf{T}}(\mathbf{FQ})(\mathbf{FQ})^{-\mathrm{T}} = \mathbf{Q}^{\mathrm{T}}(\det\mathbf{F})\mathbf{F}^{-1}\hat{\mathbf{T}}(\mathbf{F})\mathbf{F}^{-\mathrm{T}}\mathbf{Q} \\ &= \mathbf{Q}^{\mathrm{T}}\hat{\mathbf{S}}(\mathbf{F})\mathbf{Q} = \mathbf{Q}^{\mathrm{T}}\tilde{\mathbf{S}}(\mathbf{C})\mathbf{Q}\end{aligned} \tag{7.10}$$

for all $\mathbf{Q} \in \mathrm{SO}(d)$.

Because the stress functions $\tilde{\mathbf{T}}$ and $\tilde{\mathbf{S}}$ satisfy (7.9) and (7.10) respectively, for all $\mathbf{Q} \in \mathrm{SO}(d)$, both functions belong to the class of *isotropic tensor functions* defined on the set $\mathrm{Psym}(d)$. An isotropic tensor function $\Gamma : \mathcal{A} \to \mathtt{L}(\mathcal{V}_d)$ defined on some subset $\mathcal{A} \subseteq \mathtt{L}(\mathcal{V}_d)$ satisfies

$$\Gamma(\mathbf{QAQ}^{\mathrm{T}}) = \mathbf{Q}\Gamma(\mathbf{A})\mathbf{Q}^{\mathrm{T}} \tag{7.11}$$

for all $\mathbf{Q} \in \mathrm{SO}(d)$ and $\mathbf{A} \in \mathcal{A}$. The requirement (7.11) constrains the functional form of Γ, leading to the useful representation theorems for isotropic tensor functions [6]. These theorems enable us to write down the general forms of elastic constitutive relations for isotropic materials, which we use to study how activity in isotropic elastic solids can be represented. We consider tensor functions defined on the domain $\mathcal{A}$ contained in the set of symmetric second-order tensors

$$\mathrm{Sym}(d) = \{\mathbf{A} \in \mathtt{L}(\mathcal{V}_d) : \mathbf{A} = \mathbf{A}^{\mathrm{T}}\}. \tag{7.12}$$

7.2.1 Isotropy for three-dimensional materials

Any $\mathbf{Q} \in \mathrm{O}(3)$ has $\det \mathbf{Q} = \pm 1$, so either $\mathbf{Q}$ or $-\mathbf{Q}$ is in $\mathrm{SO}(3)$, because $\det(-\mathbf{Q}) = -\det \mathbf{Q}$. Because $(-\mathbf{Q})\mathbf{A}(-\mathbf{Q})^{\mathrm{T}} = \mathbf{Q}\mathbf{A}\mathbf{Q}^{\mathrm{T}}$, an isotropic tensor function defined on some subset of $\mathrm{L}(\mathcal{V}_3)$ satisfies (7.11) for all $\mathbf{Q} \in \mathrm{O}(3)$, i.e. $\mathrm{SO}(3)$ can be replaced by $\mathrm{O}(3)$ in the definition of an isotropic tensor function [6]. Physically, invariance under pure rotations in three-dimensional space is equivalent to invariance under combinations of rotations and reflections. Therefore, isotropic three-dimensional materials are necessarily achiral. The representation theorem for isotropic tensor functions defined on subsets of $\mathrm{Sym}(3)$ is proven using this equivalence, e.g. in Gurtin [6]. Because the proof of this result does not require the tensor function to be the derivative of a scalar function, its validity is not restricted to hyperelastic constitutive relations. Therefore, the general form for the stress function $\tilde{\mathbf{S}}$ of an isotropic, three-dimensional elastic material is [6]

$$\tilde{\mathbf{S}}(\mathbf{C}) = \alpha_0(\mathcal{I}(\mathbf{C}))\mathbf{I} + \alpha_1(\mathcal{I}(\mathbf{C}))\mathbf{C} + \alpha_2(\mathcal{I}(\mathbf{C}))\mathbf{C}^2, \tag{7.13}$$

where $\alpha_0, \alpha_1, \alpha_2$ are scalar-valued functions of the strain invariants $\mathcal{I}(\mathbf{C}) = \{I_k(\mathbf{C})\}_{k=1}^3$, conventionally chosen to be

$$I_1(\mathbf{C}) = \operatorname{tr}\mathbf{C}, \qquad I_2(\mathbf{C}) = \frac{1}{2}\left((\operatorname{tr}\mathbf{C})^2 - \operatorname{tr}\left(\mathbf{C}^2\right)\right), \qquad I_3(\mathbf{C}) = \det\mathbf{C}. \tag{7.14}$$

7.2.2 Isotropy for two-dimensional materials

Because $\det(-\mathbf{Q}) = \det\mathbf{Q}$ for $\mathbf{Q} \in \mathrm{O}(2)$, $\mathrm{SO}(2)$ cannot be replaced by $\mathrm{O}(2)$ in the definition of an isotropic tensor function [104]. Physically, invariance under pure rotations in two-dimensional space does not imply invariance under reflections. Therefore, isotropic two-dimensional materials can be chiral, such as the examples of isotropic, odd-elastic materials in Scheibner et al. [10] and Veenstra et al. [100]. Although there are several proofs of representation theorems for isotropic tensor functions defined on subsets of $\mathrm{L}(\mathcal{V}_3)$, e.g. [6, 105–107], the corresponding results for $\mathrm{L}(\mathcal{V}_2)$ have only been proven for *linear* tensor functions [104, 108] and assume that the elasticity tensor is symmetric, thus ruling out odd elasticity. Although the proof by Alexanderian et al. [102] for linear tensor functions does not assume hyperelasticity, it only applies to elasticity tensors that are invariant under reflections as well as rotations, so it does not capture the chiral odd elasticity presented in Scheibner et al. [10] and Veenstra et al. [100].

In Appendix II.3, we establish a representation theorem for isotropic tensor functions defined on subsets of $\mathrm{Sym}(2)$ that applies to nonlinear tensor functions without assuming

invariance under reflections. Because these functions can be chiral, their general form includes terms that depend on an antisymmetric tensor $\mathbf{W}$, given in (6.2). The set of antisymmetric tensors is a one-dimensional subspace of $\mathtt{L}(\mathcal{V}_2)$, so all antisymmetric tensors are scalar multiples of each other. Because $\mathbf{W}$ involves specifying a sense of rotation, any terms involving $\mathbf{W}$ are not invariant under reflections, i.e. they are chiral. Applying the representation theorem proven in Appendix II.3 to elastic constitutive relations shows that the general form for the stress function $\tilde{\mathbf{S}}$ of an isotropic, two-dimensional elastic material is

$$\tilde{\mathbf{S}}(\mathbf{C}) = \beta_0(\mathcal{J}(\mathbf{C}))\mathbf{I} + \beta_1(\mathcal{J}(\mathbf{C}))\mathbf{C} + \beta_\mathrm{s}(\mathcal{J}(\mathbf{C}))\,\mathrm{sym}\,(\mathbf{WC}) + \beta_\mathrm{a}(\mathcal{J}(\mathbf{C}))\mathbf{W}, \tag{7.15}$$

where the scalar-valued functions $\beta_0, \beta_1, \beta_\mathrm{s}, \beta_\mathrm{a}$ depend on the two strain invariants $\mathcal{J}(\mathbf{C}) = \{J_1(\mathbf{C}), J_2(\mathbf{C})\}$ of two-dimensional elasticity:

$$J_1(\mathbf{C}) = \mathrm{tr}\,\mathbf{C}, \quad J_2(\mathbf{C}) = \det\mathbf{C}. \tag{7.16}$$

The general form (7.15) is the nonlinear generalisation of the isotropic linear stress–strain relationship (6.1) characterised by the moduli B, μ, K^o, A, given in Scheibner et al. [10]. Retaining only contributions linear in strain from the first two terms in (7.15) leads to terms involving the bulk modulus B and the shear modulus μ, which characterise passive isotropic elasticity. The term linear in strain from $\beta_\mathrm{s}(\mathcal{J}(\mathbf{C}))\,\mathrm{sym}\,(\mathbf{WC})$ corresponds to antisymmetric coupling between independent shear strains/stresses, characterised by the odd modulus K^o. Linearising $\beta_\mathrm{a}(\mathcal{J}(\mathbf{C}))\mathbf{W}$ leads to a term containing the odd modulus A, which couples dilations to torques.

7.3 Weakly nonlinear isotropic elasticity

To identify elastic moduli characterising soft active solids, we use the framework of *weakly nonlinear elasticity* [109–112], which has been used to model biological tissues and gels [8]. In weakly nonlinear elasticity, stress is expressed as a function of the Green–Lagrange strain tensor

$$\mathbf{E} = \frac{1}{2}(\mathbf{C} - \mathbf{I}) \tag{7.17}$$

and only terms up to second or third order in $||\mathbf{E}||$ are retained, where $||\mathbf{E}|| = \sqrt{\mathbf{E} : \mathbf{E}}$. The Green–Lagrange strain tensor is related to the displacement gradient $\nabla_{\mathbf{x}_0}\mathbf{u} = \mathbf{F} - \mathbf{I}$ by $\mathbf{E} = \nabla^\mathrm{s}\mathbf{u} + \frac{1}{2}(\nabla_{\mathbf{x}_0}\mathbf{u})^\mathrm{T}\nabla_{\mathbf{x}_0}\mathbf{u}$, so it reduces to the usual strain tensor $\nabla^\mathrm{s}\mathbf{u}$ in linear elasticity when terms nonlinear in $\nabla_{\mathbf{x}_0}\mathbf{u}$ are neglected. We consider isotropic materials, so we can expand the stress functions (7.13) and (7.15) for three- and two-dimensional isotropic elasticity,

respectively, in powers of $||\mathbf{E}||$. Our approach therefore generalises the formulation of weakly nonlinear isotropic elasticity in Landau et al. [110], which considers only three-dimensional hyperelastic materials.

7.3.1 Three-dimensional materials

The choice of strain invariants (7.14) is not unique, and there is a one-to-one correspondence between the invariants in (7.14) and the Landau strain invariants $\{\operatorname{tr}\mathbf{E}, \operatorname{tr}(\mathbf{E}^2), \operatorname{tr}(\mathbf{E}^3)\}$ [110, 112]. Therefore, to facilitate expanding $\tilde{\mathbf{S}}(\mathbf{C})$ in powers of $\mathbf{E}$, we express (7.13) in terms of Landau strain invariants. Setting $\mathbf{C} = \mathbf{I} + 2\mathbf{E}$ and retaining terms up to second order in $||\mathbf{E}||$ gives the general form of a weakly nonlinear stress function for isotropic elasticity:

$$\tilde{\mathbf{S}}(\mathbf{I}+2\mathbf{E}) = \big(2a\operatorname{tr}\mathbf{E} + 3c(\operatorname{tr}\mathbf{E})^2 + d_1\operatorname{tr}(\mathbf{E}^2)\big)\mathbf{I} + (2b + 2d_2\operatorname{tr}\mathbf{E})\mathbf{E} + 3e\mathbf{E}^2, \tag{7.18}$$

where a, b, c, d_1, d_2, e are elastic moduli and we assumed that $\tilde{\mathbf{S}}(\mathbf{I}) = \mathbf{0}$, i.e. we consider materials that are unstressed at zero strain. The notation we have chosen for the elastic moduli is motivated by the hyperelastic case, in which the general form of the strain energy function $\widetilde{W}$ for an isotropic material that is unstressed when $\mathbf{E} = \mathbf{0}$ is

$$\widetilde{W}(\mathbf{I}+2\mathbf{E}) = a(\operatorname{tr}\mathbf{E})^2 + b\operatorname{tr}(\mathbf{E}^2) + c(\operatorname{tr}\mathbf{E})^3 + d(\operatorname{tr}\mathbf{E})\operatorname{tr}(\mathbf{E}^2) + e\operatorname{tr}(\mathbf{E}^3), \tag{7.19}$$

where a, b, c, d, e are constants and we have retained terms up to third order in $||\mathbf{E}||$. Using (II.9) and the chain rule, the second Piola–Kirchhoff stress in the hyperelastic case is

$$\begin{aligned}\tilde{\mathbf{S}}_{\text{hyp}}(\mathbf{I}+2\mathbf{E}) &= \frac{\mathrm{d}}{\mathrm{d}\mathbf{E}}\widetilde{W}(\mathbf{I}+2\mathbf{E}) \\ &= \big(2a\operatorname{tr}\mathbf{E} + 3c(\operatorname{tr}\mathbf{E})^2 + d\operatorname{tr}(\mathbf{E}^2)\big)\mathbf{I} + (2b + 2d\operatorname{tr}\mathbf{E})\mathbf{E} + 3e\mathbf{E}^2.\end{aligned} \tag{7.20}$$

We see that in weakly nonlinear isotropic elasticity where terms up to second order in $||\mathbf{E}||$ are retained, the general elastic constitutive relation (7.18) reduces to the hyperelastic relation (7.20) when $d_1 = d_2$, i.e. hyperelasticity imposes one constraint on the elastic moduli characterising the stress–strain relation. Therefore, we define

$$d = \frac{1}{2}(d_1 + d_2), \quad d_o = \frac{1}{2}(d_1 - d_2) \tag{7.21}$$

to express (7.18) as

$$\begin{aligned}\tilde{\mathbf{S}}(\mathbf{I}+2\mathbf{E}) = \frac{\mathrm{d}}{\mathrm{d}\mathbf{E}}\left(a(\mathrm{tr}\,\mathbf{E})^2 + b\,\mathrm{tr}(\mathbf{E}^2) + c(\mathrm{tr}\,\mathbf{E})^3 + d(\mathrm{tr}\,\mathbf{E})\,\mathrm{tr}(\mathbf{E}^2) + e\,\mathrm{tr}(\mathbf{E}^3)\right) \\ + d_o\left(\mathrm{tr}(\mathbf{E}^2)\mathbf{I} - 2(\mathrm{tr}\,\mathbf{E})\mathbf{E}\right),\end{aligned} \tag{7.22}$$

showing that the modulus d_o characterises the departure from hyperelasticity of a general elastic constitutive relation in weakly nonlinear isotropic elasticity.

Odd moduli in the elasticity tensor

This departure from hyperelasticity implies that there exist deformed reference states $\mathbf{F}_1$ for which the elasticity tensor $\mathsf{C}_{\mathbf{F}_1}$ in (7.5) has a non-zero antisymmetric (odd) part, as shown by (7.6). Because the term independent of d_o in (7.22) does not contribute to $\mathrm{asym}\,\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}(\mathbf{C})$, only the term proportional to d_o contributes to the antisymmetric part of $\mathsf{C}_{\mathbf{F}_1}$. In Appendix II.4, we show that this antisymmetric part is

$$\mathrm{asym}\,\mathsf{C}_{\mathbf{F}_1} = \frac{d_o}{\det\mathbf{F}_1}(\mathbf{B}_1\otimes\mathbf{B}_1^2 - \mathbf{B}_1^2\otimes\mathbf{B}_1), \tag{7.23}$$

where $\mathbf{B}_1 = \mathbf{F}_1\mathbf{F}_1^{\mathrm{T}}$ is the left Cauchy–Green deformation tensor for the deformed reference state. Linearising the stress response about the undeformed state, i.e. setting $\mathbf{F}_1 = \mathbf{I}$, results in an elasticity tensor with major symmetry, consistent with previous results [6, 102] that rule out odd elasticity in the linear regime for isotropic three-dimensional materials. In nonlinear elastic constitutive relations, linearising about a deformed reference state can lead to odd moduli because this reference state is not isotropic.

To explore how the odd elastic moduli depend on the deformed reference state, we denote the principal stretches and Eulerian principal directions (eigenvectors of $\mathbf{B}_1$) of the deformation $\mathbf{F}_1$ by $\{(\lambda_j,\mathbf{v}_j)\}_{j=1}^3$. Substituting $\det\mathbf{F}_1 = \lambda_1\lambda_2\lambda_3$ and $\mathbf{B}_1 = \lambda_1^2\mathbf{v}_1\otimes\mathbf{v}_1 + \lambda_2^2\mathbf{v}_2\otimes\mathbf{v}_2 + \lambda_1^3\mathbf{v}_3\otimes\mathbf{v}_3$ into (7.23) shows that the only components contributed by $\mathrm{asym}\,\mathsf{C}_{\mathbf{F}_1}[\nabla_{\mathbf{x}_1}\mathbf{u}]$ to the incremental stress response $\mathsf{C}_{\mathbf{F}_1}[\nabla_{\mathbf{x}_1}\mathbf{u}]$ to small strains from the deformed reference state (cf. (7.4)) are

$$\sigma_j^o = \mathbf{v}_j\otimes\mathbf{v}_j : \mathrm{asym}\,\mathsf{C}_{\mathbf{F}_1}[\nabla_{\mathbf{x}_1}\mathbf{u}], \qquad j=1,2,3, \tag{7.24}$$

which are normal stresses along the principal directions $\mathbf{v}_j$. Similarly, the only components of the strain $\nabla_{\mathbf{x}_1}\mathbf{u}$ relative to the deformed reference state that contribute to $\mathrm{asym}\,\mathsf{C}_{\mathbf{F}_1}[\nabla_{\mathbf{x}_1}\mathbf{u}]$ are

$$\varepsilon_j = \mathbf{v}_j\otimes\mathbf{v}_j : (\nabla_{\mathbf{x}_1}\mathbf{u}), \qquad j=1,2,3, \tag{7.25}$$

which are normal strains along the principal directions $\mathbf{v}_j$. The relationship between these stress σ_j^o and strain ε_j components given by the odd elasticity tensor (7.23) has the matrix form

$$\begin{bmatrix} \sigma_1^o \\ \sigma_2^o \\ \sigma_3^o \end{bmatrix} = \frac{d_o}{\lambda_1\lambda_2\lambda_3}\mathbf{K}_{\text{odd}} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}, \tag{7.26}$$

where $\mathbf{K}_{\text{odd}}$ is an antisymmetric matrix whose (j,k)th element is $(\mathbf{K}_{\text{odd}})_{jk} = \lambda_j^2\lambda_k^2(\lambda_j^2 - \lambda_k^2)$. This equation shows that antisymmetric coupling between normal stresses and strains along two given principal directions $\mathbf{v}_j, \mathbf{v}_k$ is present if and only if the principal stretches λ_j, λ_k of the deformed reference state along those directions are not equal. Significantly, unlike in two-dimensional isotropic linear elasticity, the presence of odd moduli in isotropic three-dimensional elasticity is not due to the breaking of chiral symmetry, but due to the asymmetry between directions induced by the pre-deformation. We therefore refer to these odd moduli as *deformation-induced* odd moduli.

Measuring the odd modulus d_o

Here we present a simple two-parameter deformation that can be used to measure the odd modulus d_o. This deformation consists of simple shear parametrised by γ followed by a uniaxial stretch λ, so the matrix representation of the deformation gradient with respect to some orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is

$$[\mathbf{F}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \gamma & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{7.27}$$

Substituting this deformation gradient into (7.22) and computing derivatives of the Cauchy stress $\mathbf{T}$ with respect to γ and λ gives

$$\frac{\partial T_{22}}{\partial \gamma} - \frac{\partial T_{12}}{\partial \lambda} = 2d_o\gamma\lambda, \tag{7.28}$$

where $T_{12} = \mathbf{e}_1 \otimes \mathbf{e}_2 : \mathbf{T}$ is the shear stress and $T_{22} = \mathbf{e}_2 \otimes \mathbf{e}_2 : \mathbf{T}$ is the normal stress in the shearing plane. For any isotropic hyperelastic material, we show in Appendix II.5 that $\frac{\partial T_{22}}{\partial \gamma} - \frac{\partial T_{12}}{\partial \lambda} = 0$ identically, so the deformation (7.27) can be used to check for non-hyperelasticity in any isotropic elastic constitutive relation.

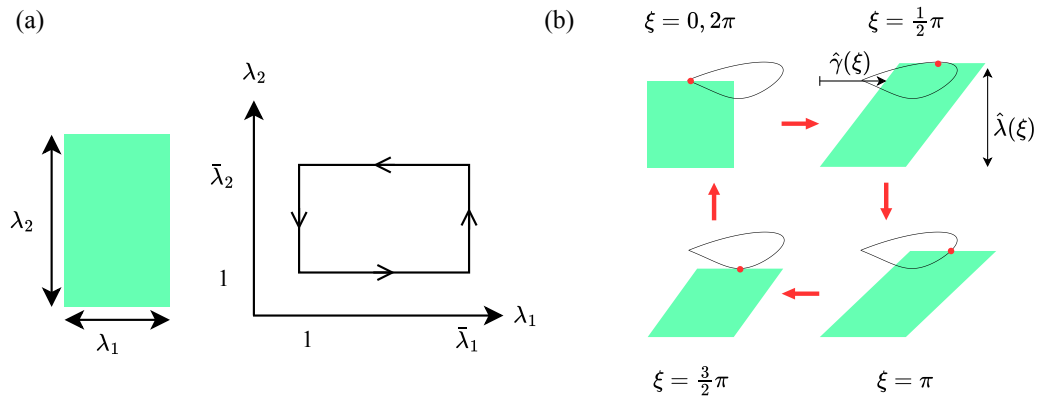


Fig. 7.1 Closed cycles of deformation for work extraction. The solid green shapes represent an isotropic, active elastic material. (a) A cycle of deformation involving stretches along two orthogonal directions, specified by the path in λ_1–λ_2-space. (b) A cycle of deformation involving shearing and stretching, parametrised by $\xi \in [0, 2\pi]$, with the shear $\hat{\gamma}(\xi)$ and stretch $\hat{\lambda}(\xi)$ given by (7.31). The red circular dot indicates the midpoint of the top edge of the material, and traces out the path indicated by the closed black curve.

Work done in closed cycles of deformation

When $d_o \neq 0$, the constitutive relation (7.22) is not hyperelastic, so there exist cycles of deformation such that the work done on the material, given in (6.13), is non-zero. Here we compute the work done over some examples of such cycles, which illustrate how broken symmetries give rise to non-zero work.

We first consider a cycle of deformation proposed by Carroll [43], illustrated in Fig. 7.1(a), in which the path $\mathbf{C} = \hat{\mathbf{C}}(\xi)$, $\xi \in [\xi_0, \xi_1]$ in $\mathrm{Psym}(d)$ is defined piecewise to correspond to a series of in-plane deformations. This series of deformations is described as follows, using the stretches $\lambda_1, \lambda_2, \lambda_3$ that correspond to the orthogonal directions $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, respectively.

1. The material is deformed along $\mathbf{e}_1$ so that λ_1 is varied from 1 to $\bar{\lambda}_1$, while fixing $\lambda_2 = 1$.
2. The material is deformed along $\mathbf{e}_2$ so that λ_2 is varied from 1 to $\bar{\lambda}_2$, while fixing $\lambda_1 = \bar{\lambda}_1$.
3. The material is deformed along $\mathbf{e}_1$ so that λ_1 is varied from $\bar{\lambda}_1$ to 1, while fixing $\lambda_2 = \bar{\lambda}_2$.
4. The material is deformed along $\mathbf{e}_2$ so that λ_2 is varied from $\bar{\lambda}_2$ to 1, while fixing $\lambda_1 = 1$.

The work done per unit volume on the material during this cycle of deformation is computed using (6.13). If the out-of-plane stretch is fixed at $\lambda_3 = 1$, then the work done is

$$\widetilde{\mathcal{W}}\left[\hat{\mathbf{C}}\right] = \frac{d_o}{4}(\bar{\lambda}_1^2 - 1)(\bar{\lambda}_2^2 - 1)(\bar{\lambda}_1^2 - \bar{\lambda}_2^2). \tag{7.29}$$

If we consider instead an incompressible material, for which $\lambda_3 = (\lambda_1\lambda_2)^{-1}$ varies with the in-plane stretches, then the work done per unit volume on the material is

$$\widetilde{\mathcal{W}}\left[\hat{\mathbf{C}}\right] = \frac{d_o}{4}(\bar{\lambda}_1^2 - 1)(\bar{\lambda}_2^2 - 1)(\bar{\lambda}_1^2 - \bar{\lambda}_2^2)\left(1 - \frac{1}{\bar{\lambda}_1^2\bar{\lambda}_2^2}\right)^2. \tag{7.30}$$

Unlike the odd-elastic engine cycles for chiral, isotropic materials in Scheibner et al. [10], in which work done is proportional to the area enclosed by the cycle in strain space, the cycles we consider here can result in zero work done even if the area enclosed by the cycle in stretch space is non-zero. The expressions for work done in (7.29) and (7.30) show that zero work is done when $\bar{\lambda}_1 = \bar{\lambda}_2$. Because the sign of the work done $\widetilde{\mathcal{W}}[\hat{\mathbf{C}}]$ along a cycle of deformation $\mathbf{C} = \hat{\mathbf{C}}(\xi)$, $\xi \in [\xi_0, \xi_1]$ is reversed when the direction of the cycle is reversed, the deformation cycle must be asymmetric with respect to reversal for $\widetilde{\mathcal{W}}[\hat{\mathbf{C}}]$ to be non-zero. Because exchanging $\bar{\lambda}_1$ and $\bar{\lambda}_2$ is equivalent to reversing the cycle, asymmetry with respect to reversal requires $\bar{\lambda}_1 \neq \bar{\lambda}_2$.

We also consider a cycle based on in-plane shearing and stretching, illustrated in Fig. 7.1(b), where the path $\hat{\mathbf{C}}(\xi) = \hat{\mathbf{F}}(\xi)^{\mathrm{T}}\hat{\mathbf{F}}(\xi)$, $\xi \in [0, 2\pi]$ is obtained by defining

$$\hat{\gamma}(\xi) = \bar{\gamma}\sin\left(\frac{\xi}{2}\right), \quad \hat{\lambda}(\xi) = 1 + \bar{\varepsilon}\sin(\xi), \tag{7.31}$$

and $\hat{\mathbf{F}}(\xi)$ is $\mathbf{F}$ in (7.27) with $\gamma = \hat{\gamma}(\xi)$ and $\lambda = \hat{\lambda}(\xi)$. In this case, the work done per unit volume on the material is

$$\widetilde{\mathcal{W}}\left[\hat{\mathbf{C}}\right] = -\frac{d_o}{2}\pi\bar{\gamma}^2\bar{\varepsilon}. \tag{7.32}$$

This expression for work done is not changed when $\bar{\gamma}$ is replaced by $-\bar{\gamma}$, i.e. when the direction of shear is reversed, consistent with the material's symmetry under reflections.

7.3.2 Two-dimensional materials

Second-order tensors in $\mathtt{L}(\mathcal{V}_2)$ are described using two invariants, so we express the isotropic constitutive relation (7.15) in terms of the Landau-style invariants $\{\mathrm{tr}\,\mathbf{E}, \mathrm{tr}(\mathbf{E}^2)\}$. Retaining

terms up to $||\mathbf{E}||^2$ in (7.15), we obtain

$$\begin{aligned}\tilde{\mathbf{S}}(\mathbf{I}+2\mathbf{E}) = (2a\,\mathrm{tr}\,\mathbf{E}+3c(\mathrm{tr}\,\mathbf{E})^2+d_1\,\mathrm{tr}(\mathbf{E}^2))\mathbf{I}+(2b+2d_2\,\mathrm{tr}\,\mathbf{E})\mathbf{E}\\ -2(K^o+\bar{K}^o\,\mathrm{tr}\,\mathbf{E})\,\mathrm{sym}\,(\mathbf{W}\mathbf{E})+(A\,\mathrm{tr}\,\mathbf{E}+\bar{A}(\mathrm{tr}\,\mathbf{E})^2+\tilde{A}\,\mathrm{tr}(\mathbf{E}^2))\mathbf{W},\end{aligned} \tag{7.33}$$

where $a,b,c,d_1,d_2,K^o,\bar{K}^o,A,\bar{A},\tilde{A}$ are elastic moduli and we assumed that $\tilde{\mathbf{S}}(\mathbf{I})=\mathbf{0}$. We use the symbols K^o and A to refer to the same moduli as in (6.1), to facilitate comparison with the linear case. As in the three-dimensional case, the choice of symbols for the constants a,b,c,d_1,d_2 is motivated by comparison to the hyperelastic case, which is derived from the strain energy function

$$\tilde{W}(\mathbf{I}+2\mathbf{E}) = a(\mathrm{tr}\,\mathbf{E})^2+b\,\mathrm{tr}(\mathbf{E}^2)+c(\mathrm{tr}\,\mathbf{E})^3+d(\mathrm{tr}\,\mathbf{E})\,\mathrm{tr}(\mathbf{E}^2). \tag{7.34}$$

The second Piola–Kirchhoff stress derived from this strain energy function is obtained by differentiating $\tilde{W}(\mathbf{I}+2\mathbf{E})$ with respect to $\mathbf{E}$, to obtain

$$\tilde{\mathbf{S}}_{\mathrm{hyp}}(\mathbf{I}+2\mathbf{E}) = (2a\,\mathrm{tr}\,\mathbf{E}+3c(\mathrm{tr}\,\mathbf{E})^2+d\,\mathrm{tr}(\mathbf{E}^2))\mathbf{I}+(2b+2d\,\mathrm{tr}\,\mathbf{E})\mathbf{E}. \tag{7.35}$$

Comparing this expression for stress with that in (7.33) shows that hyperelasticity imposes the constraints $d_1=d_2$ and $K^o,\bar{K}^o,A,\bar{A},\tilde{A}=0$ on the elastic moduli in the weakly nonlinear isotropic stress–strain relation. The moduli $K^o,\bar{K}^o,A,\bar{A},\tilde{A}$ are all associated with broken chiral symmetry, as represented by the antisymmetric tensor $\mathbf{W}$. At small strains, the dominant chiral contributions are controlled by K^o and A. Our analysis here reveals that in the weakly nonlinear isotropic theory, besides the previously explored odd moduli K^o and A associated with terms linear in strain [10], there is an odd modulus $d_o=\frac{1}{2}(d_1-d_2)$, associated with terms quadratic in strain. Because these quadratic terms do not break chiral symmetry, they are also present in the three-dimensional case, as discussed in Section 7.3.1.

Deformation-induced odd moduli in the elasticity tensor

Here we focus on the odd modulus d_o that controls contributions from the non-chiral terms in (7.33), because the moduli K^o and A controlling the dominant chiral terms have previously been studied by Scheibner et al. [10]. Because the terms controlled by d_o are present in the three-dimensional case, their contribution to the antisymmetric part of the elasticity tensor $\mathsf{C}_{\mathbf{F}_1}$ for the linearised stress response about a deformed reference state $\mathbf{F}_1$ is given by (7.23).

By the same reasoning as in the three-dimensional case, the only stress components contributed by terms involving d_o in $\mathrm{asym}\,\mathsf{C}_{\mathbf{F}_1}[\nabla_{\mathbf{x}_1}\mathbf{u}]$ are $\sigma^o_j=\mathbf{v}_j\otimes\mathbf{v}_j:\mathrm{asym}\,\mathsf{C}_{\mathbf{F}_1}[\nabla_{\mathbf{x}_1}\mathbf{u}]$, for $j=1,2$, where $\{(\lambda_j,\mathbf{v}_j)\}_{j=1}^2$ are the principal stretches and principal directions of $\mathbf{B}_1=\mathbf{F}_1\mathbf{F}_1^{\mathrm{T}}$.

Similarly, the only components of the strain $\nabla_{\mathbf{x}_1}\mathbf{u}$ relative to the deformed reference state that contribute to the stress response controlled by d_o in $\operatorname{asym}\mathsf{C}_{\mathbf{F}_1}[\nabla_{\mathbf{x}_1}\mathbf{u}]$ are $\varepsilon_j = \mathbf{v}_j \otimes \mathbf{v}_j : (\nabla_{\mathbf{x}_1}\mathbf{u})$ for $j = 1,2$. The antisymmetric coupling between these stress and strain components has the matrix form

$$\begin{bmatrix} \sigma_1^o \\ \sigma_2^o \end{bmatrix} = d_o \begin{bmatrix} 0 & \lambda_1\lambda_2(\lambda_2^2 - \lambda_1^2) \\ -\lambda_1\lambda_2(\lambda_2^2 - \lambda_1^2) & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}, \tag{7.36}$$

which shows how the odd moduli proportional to d_o depend on the anisotropy of the pre-deformation $\mathbf{F}_1$ used to arrive at the reference state. We have therefore shown that, in two-dimensional materials, in addition to the odd moduli associated with broken chiral symmetry, terms quadratic in strain can be associated with deformation-induced odd moduli that are also present in isotropic three-dimensional elasticity.

Measuring the odd moduli

Because the two-parameter deformation (7.27) only involves two orthogonal directions $\mathbf{e}_1$ and $\mathbf{e}_2$, it has a straighforward analogue for two-dimensional materials:

$$[\mathbf{F}] = \begin{bmatrix} 1 & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} 1 & \gamma \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \gamma \\ 0 & \lambda \end{bmatrix}. \tag{7.37}$$

Substituting this deformation gradient into (7.33) and computing derivatives of the Cauchy stress $\mathbf{T}$ with respect to γ and λ gives

$$\frac{\partial T_{22}}{\partial \gamma} - \frac{\partial T_{12}}{\partial \lambda} = 2d_o\gamma\lambda + (A - 2K^o)\lambda + \left(\bar{A} + \tilde{A} - \frac{3}{2}\bar{K}^o\right)\lambda(\lambda^2 - 1 + \gamma^2), \tag{7.38}$$

where $T_{12} = \mathbf{e}_1 \otimes \mathbf{e}_2 : \mathbf{T}$ is the shear stress and $T_{22} = \mathbf{e}_2 \otimes \mathbf{e}_2 : \mathbf{T}$ is the normal stress in the shearing plane. Importantly, the terms involving the moduli $K^o, \bar{K}^o, A, \bar{A}, \tilde{A}$, which break chiral symmetry, are even in γ, whereas the term involving the modulus d_o is odd in γ. These differences are consistent with the physical origin of the odd moduli: because d_o captures odd-elastic behaviour that arises from deformation-induced anisotropy, as shown in (7.36), the term associated with this modulus changes sign if the direction of shear is reversed. In contrast, the odd moduli K^o, A capture odd-elastic behaviour due to chirality in the material, and do not rely on deformation-induced anisotropy.

Work done in closed cycles of deformation

The cycles of deformation shown in Fig. 7.1(a) and (b) involve stretching or shearing only in the plane spanned by $\mathbf{e}_1$ and $\mathbf{e}_2$, so they can also be performed on two-dimensional materials.

However, because the second Piola–Kirchhoff stress $\mathbf{S}$ is not necessarily symmetric, the work done per unit volume on the material needs to be computed using (6.6) rather than (6.13).

For the cycle of deformation proposed by Carroll [43], shown in Fig. 7.1(a), the work done on the material per unit volume is $\mathcal{W}[\hat{\mathbf{F}}] = \frac{d_o}{4}(\bar{\lambda}_1^2 - 1)(\bar{\lambda}_2^2 - 1)(\bar{\lambda}_1^2 - \bar{\lambda}_2^2)$, the same as that given in (7.29). Importantly, the odd moduli $K^o, \bar{K}^o, A, \bar{A}, \tilde{A}$ do not contribute to the work done over this cycle, because it does not rely on material chirality. Therefore, imposing this cycle of deformation on an isotropic odd-elastic material with stress depending *linearly* on strain would result in zero work extraction. This result illustrates an advantage of considering nonlinear active elasticity: work extraction can be accomplished using larger class of deformation cycles.

The deformation cycle shown in Fig. 7.1(b) involves work done per unit volume of material given by

$$\mathcal{W}[\hat{\mathbf{F}}] = -\frac{d_o}{2}\pi\bar{\gamma}^2\bar{\varepsilon} + \frac{1}{210}\left(560K^o - 280A - 8\left(2(\bar{A}+\tilde{A}) - 3\bar{K}^o\right)(12\bar{\varepsilon}^2 + 7\bar{\gamma}^2)\right)\bar{\gamma}\bar{\varepsilon}. \quad (7.39)$$

The first term, involving d_o, is the same as that obtained in (7.32), for which chirality is absent. In contrast, replacing $\bar{\gamma}$ with $-\bar{\gamma}$ reverses the signs of the terms involving $K^o, \bar{K}^o, A, \bar{A}, \tilde{A}$, consistent with the chiral nature of these terms. We can consider another deformation cycle in which the parameter ξ is varied over $[0, 4\pi]$ instead of $[0, 2\pi]$ but using the same expressions for $\hat{\gamma}(\xi)$ and $\hat{\lambda}(\xi)$ as in (7.31). The first half of this modified cycle has the same form as that shown in Fig. 7.1(b), and its second half is a mirror image of the first half, i.e. the second half is the first half with the direction of shear reversed. For this cycle, the work done per unit volume is

$$\begin{aligned}\mathcal{W}[\hat{\mathbf{F}}] &= -\frac{d_o}{2}\pi\bar{\gamma}^2\bar{\varepsilon} + \frac{1}{210}\left(560K^o - 280A - 8\left(2(\bar{A}+\tilde{A}) - 3\bar{K}^o\right)(12\bar{\varepsilon}^2 + 7\bar{\gamma}^2)\right)\bar{\gamma}\bar{\varepsilon} \\ &\quad - \frac{d_o}{2}\pi\bar{\gamma}^2\bar{\varepsilon} - \frac{1}{210}\left(560K^o - 280A - 8\left(2(\bar{A}+\tilde{A}) - 3\bar{K}^o\right)(12\bar{\varepsilon}^2 + 7\bar{\gamma}^2)\right)\bar{\gamma}\bar{\varepsilon} \\ &= -d_o\pi\bar{\gamma}^2\bar{\varepsilon},\end{aligned} \quad (7.40)$$

showing that the contributions from the chiral terms cancel out, leaving just the terms that depend on d_o. This cycle of deformation is therefore another member of the class of cycles for which work extraction from isotropic, two-dimensional elastic materials requires the presence of terms nonlinear in strain.

Chapter 8

Nonlinear active elasticity of random networks

In this chapter, we show how the phenomenology of nonlinear active elasticity discussed in Chapter 7 can be realised in isotropic random networks. First, we present a homogenisation procedure to construct a stress tensor from the fibre forces in a random network, which we use to obtain analytical constitutive relations for networks of active fibres. We then describe a class of active fibres that gives rise to non-hyperelastic behaviour in both two- and three-dimensional random networks. Using numerical simulations, we confirm that networks of these fibres exhibit non-hyperelastic behaviour controlled by the odd modulus d_o. Our work therefore points towards new realisations of active elastic materials that exhibit this novel form of odd elasticity.

8.1 Homogenisation of a random network

We present a homogenisation procedure that constructs a stress tensor from microscopic forces, to obtain constitutive relations from microscopic realisations of active solids. Because we consider microscopic forces that do not necessarily arise from potentials, the Irving–Kirkwood–Noll procedure [40, 41, 60] used in Part I is not applicable. However, the virial stress tensor, used to compute stress in molecular dynamics simulations [60, 113], can be derived without assuming that microscopic forces are conservative. An example of such a derivation can be found in Appendix A of Admal and Tadmor [60], which uses time averages. Because we focus on athermal random networks instead of molecular systems, we present a different argument for constructing a stress tensor from microscopic forces, which does not rely on the virial theorem. Our argument is based on spatial averaging of forces and leads

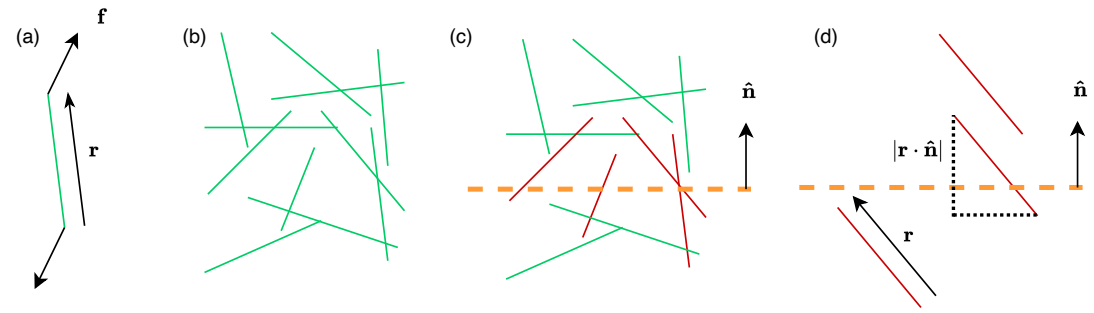


Fig. 8.1 Constructing a stress tensor for a random distribution of fibres. (a) Fibre with end-to-end vector $\mathbf{r}$ maintained by a force $\mathbf{f}$. (b) Random distribution of fibres in an elastic material. (c) A plane surface, normal to $\hat{\mathbf{n}}$, separating two parts of the material is shown as an orange dashed line. Fibres that intersect the surface are shown in red. These fibres contribute to the force $\mathbf{T}\hat{\mathbf{n}}$, where $\mathbf{T}$ is the Cauchy stress tensor. (d) For fibre positions homogeneously distributed throughout the material, the probability that a fibre with end-to-end vector $\mathbf{r}$ crosses a given plane surface with normal $\hat{\mathbf{n}}$ is proportional to its projected length $|\mathbf{r}\cdot\hat{\mathbf{n}}|$.

to a stress tensor similar to the part of the virial stress tensor that arises from inter-particle forces.

We consider materials consisting of active force-exerting elements, which we refer to as fibres. Fig. 8.1(a) shows a fibre with end-to-end vector $\mathbf{r}$ and the force $\mathbf{f}(\mathbf{r})$ required to maintain the fibre in its current state. Because each fibre is subjected to a pair of forces with equal magnitude and opposite directions, we adopt the convention that $\mathbf{f}(\mathbf{r})$ is the force on the end of the fibre to which $\mathbf{r}$ points, so the force on the other end is $-\mathbf{f}(\mathbf{r})$.

The Cauchy stress tensor $\mathbf{T}$ characterises the contact forces between parts of a material body. At each point on the contact surface between two parts, the contact force per unit area exerted by one part on the other is $\mathbf{T}\hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ is the unit normal vector that specifies the local orientation of the contact surface at that point. Therefore, in this section, we construct the Cauchy stress tensor by summing over the force contributions of the fibres intersecting specified contact surfaces.

We consider fibres randomly distributed throughout the solid, as illustrated in Fig. 8.1(b), where the fibre end-to-end vectors $\mathbf{r}$ in the deformed state are characterised by the probability density $\tilde{\rho}(\mathbf{r},\mathbf{F})$, which depends on the deformation gradient $\mathbf{F}$. Fig. 8.1(c) shows a plane surface normal to $\hat{\mathbf{n}}$, shown as a dashed orange line, separating two parts of a material containing N_f fibres distributed over a volume Ω. The fibres that contribute to the force per unit area $\mathbf{T}\hat{\mathbf{n}}$ exerted across this surface are the fibres, coloured in red on Fig. 8.1(c), that intersect the surface. If the fibres are homogeneously distributed throughout the volume Ω of material, then the probability that a fibre with end-to-end vector $\mathbf{r}$ crosses the surface is proportional to the projection of its length $|\mathbf{r}\cdot\hat{\mathbf{n}}|$ normal to the surface, as shown on Fig. 8.1(d).

Because $\mathbf{T}\hat{\mathbf{n}}$ is the force per unit area exerted across the surface *by* the material part to which $\hat{\mathbf{n}}$ points, each fibre intersecting the surface contributes a force $\operatorname{sgn}(\mathbf{r}\cdot\hat{\mathbf{n}})\mathbf{f}(\mathbf{r})$, where the sign of $\mathbf{r}\cdot\hat{\mathbf{n}}$ accounts for the sign convention adopted for $\mathbf{f}(\mathbf{r})$, shown in Fig. 8.1(a). The average force exerted across a plane surface normal to $\hat{\mathbf{n}}$ with area S is obtained by summing the force contributions from all the fibres, weighted by the probability of a given fibre intersecting the surface. The probability that a fibre intersects the surface, given that it has end-to-end vector $\mathbf{r}$, is $\Omega^{-1}S|\mathbf{r}\cdot\hat{\mathbf{n}}|$, so this average force is

$$S\mathbf{T}\hat{\mathbf{n}} = \sum_{m=1}^{N_\mathrm{f}} \operatorname{sgn}(\mathbf{r}_m\cdot\hat{\mathbf{n}})\mathbf{f}(\mathbf{r}_m)\frac{S|\mathbf{r}_m\cdot\hat{\mathbf{n}}|}{\Omega} = S\left(\frac{1}{\Omega}\sum_{m=1}^{N_\mathrm{f}}\mathbf{f}(\mathbf{r}_m)\otimes\mathbf{r}_m\right)\hat{\mathbf{n}}, \tag{8.1}$$

where $\mathbf{r}_m$ is the end-to-end vector of the mth fibre. Because this expression for average force is valid for any surface orientation $\hat{\mathbf{n}}$, the Cauchy stress is given by

$$\mathbf{T} = \frac{1}{\Omega}\sum_{m=1}^{N_\mathrm{f}}\mathbf{f}(\mathbf{r}_m)\otimes\mathbf{r}_m, \tag{8.2}$$

which is similar in form to the part of the virial stress tensor arising from inter-particle forces [60]. We can replace the sum over fibres in (8.2) by an integral over end-to-end vectors $\mathbf{r}\in\mathcal{V}_d$ weighted by the end-to-end vector probability density $\tilde{\rho}(\mathbf{r},\mathbf{F})$, to obtain

$$\mathbf{T} = n\int_{\mathcal{V}_d}\mathbf{f}(\mathbf{r})\otimes\mathbf{r}\,\tilde{\rho}(\mathbf{r},\mathbf{F})\,\mathrm{d}\mathbf{r}, \tag{8.3}$$

where $n = N_\mathrm{f}/\Omega$ is the number density of fibres in the deformed state. Comparing this expression to (3.10) shows that the form of the Cauchy stress tensor constructed from chain forces using the Irving–Kirkwood–Noll procedure is a special case of (8.3). Although the reasoning we use in this section is more general than that used to arrive at (3.10) in not requiring microscopic forces to be conservative, it does not replace the Irving–Kirkwood–Noll procedure because it does not account for statistical-mechanical averages over phase space.

Here we consider materials for which we know the end-to-end vector distribution in the undeformed state rather than in the deformed state, so we perform a change of variables to carry out the integration (8.3) over end-to-end vectors $\mathbf{r}_0$ in the undeformed state. To capture the change in fibre end-to-end vectors due to material deformation, we assume that there is a one-to-one function $\nu_\mathbf{F}:\mathcal{V}_d\to\mathcal{V}_d$ that maps each undeformed fibre end-to-end vector $\mathbf{r}_0$ to a deformed end-to-end vector $\mathbf{r} = \nu_\mathbf{F}(\mathbf{r}_0)$, capturing the average behaviour of the fibres under deformation. Letting $\tilde{\rho}_0(\mathbf{r}_0)$ be the probability density for $\mathbf{r}_0$, the conservation of the number

of fibres implies that, for any subset $\mathcal{D} \subseteq \mathcal{V}_d$, $\tilde{\rho}_0$ is related to $\tilde{\rho}$ by

$$\int_{\nu_{\mathbf{F}}^{-1}(\mathcal{D})} \tilde{\rho}_0(\mathbf{r}_0)\,\mathrm{d}\mathbf{r}_0 = \int_{\mathcal{D}} \tilde{\rho}(\mathbf{r},\mathbf{F})\,\mathrm{d}\mathbf{r},$$

where $\nu_{\mathbf{F}}^{-1}(\mathcal{D}) = \{\mathbf{r}_0 \in \mathcal{V}_d : \nu_{\mathbf{F}}(\mathbf{r}_0) \in \mathcal{D}\}$. By the change of variables formula,

$$\int_{\mathcal{D}} \tilde{\rho}(\mathbf{r},\mathbf{F})\,\mathrm{d}\mathbf{r} = \int_{\nu_{\mathbf{F}}^{-1}(\mathcal{D})} \tilde{\rho}(\nu_{\mathbf{F}}(\mathbf{r}_0),\mathbf{F})\,|\det \mathrm{D}_{\mathbf{r}_0}\nu_{\mathbf{F}}(\mathbf{r}_0)|\,\mathrm{d}\mathbf{r}_0,$$

so that the usual localisation arguments give

$$\tilde{\rho}_0(\mathbf{r}_0) = \tilde{\rho}(\nu_{\mathbf{F}}(\mathbf{r}_0),\mathbf{F})\,|\det \mathrm{D}_{\mathbf{r}_0}\nu_{\mathbf{F}}(\mathbf{r}_0)|. \tag{8.4}$$

Performing the change of variables $\mathbf{r} = \nu_{\mathbf{F}}(\mathbf{r}_0)$ in (8.3) and using (8.4) gives

$$\begin{aligned}\mathbf{T} &= n\int_{\mathcal{V}_d} \mathbf{f}(\nu_{\mathbf{F}}(\mathbf{r}_0)) \otimes \nu_{\mathbf{F}}(\mathbf{r}_0)\,\tilde{\rho}(\nu_{\mathbf{F}}(\mathbf{r}_0),\mathbf{F})\,|\det \mathrm{D}_{\mathbf{r}_0}\nu_{\mathbf{F}}(\mathbf{r}_0)|\,\mathrm{d}\mathbf{r}_0 \\ &= \frac{n_0}{\det\mathbf{F}}\int_{\mathcal{V}_d} \mathbf{f}(\nu_{\mathbf{F}}(\mathbf{r}_0)) \otimes \nu_{\mathbf{F}}(\mathbf{r}_0)\,\tilde{\rho}_0(\mathbf{r}_0)\,\mathrm{d}\mathbf{r}_0,\end{aligned} \tag{8.5}$$

where $n_0 = n\det\mathbf{F}$ is the number density of fibres in the undeformed state. If all the fibres have approximately the same undeformed length $r_0 = \bar{r}_0$, then we can approximate the integral over end-to-end vectors $\mathbf{r}_0$, weighted by the probability density $\tilde{\rho}(\mathbf{r}_0)$, with an integral over fibre orientations $\hat{\mathbf{r}}_0$, weighted by the probability density for fibre orientations $\hat{\rho}(\hat{\mathbf{r}}_0)$. With this approximation, the Cauchy stress simplifies to

$$\mathbf{T} = \frac{n_0}{\det\mathbf{F}}\int_{\mathcal{S}^{d-1}} \mathbf{f}(\nu_{\mathbf{F}}(\bar{r}_0\hat{\mathbf{r}}_0)) \otimes \nu_{\mathbf{F}}(\bar{r}_0\hat{\mathbf{r}}_0)\,\hat{\rho}_0(\hat{\mathbf{r}}_0)\,\mathrm{d}A(\hat{\mathbf{r}}_0), \tag{8.6}$$

where the domain of integration is the unit sphere $\mathcal{S}^{d-1}$ rather than the whole vector space $\mathcal{V}_d$. Because we focus on the nonlinear elasticity of isotropic materials, we also assume that the fibres are isotropically oriented in the undeformed state, i.e. $\hat{\rho}_0(\hat{\mathbf{r}}_0) = \frac{1}{2^{d-1}\pi}$, so

$$\mathbf{T} = \frac{n_0}{2^{d-1}\pi\det\mathbf{F}}\int_{\mathcal{S}^{d-1}} \mathbf{f}(\nu_{\mathbf{F}}(\bar{r}_0\hat{\mathbf{r}}_0)) \otimes \nu_{\mathbf{F}}(\bar{r}_0\hat{\mathbf{r}}_0)\,\mathrm{d}A(\hat{\mathbf{r}}_0). \tag{8.7}$$

To obtain predictions of stress for a material consisting of an isotropic distribution of fibres, we need to specify the functional form of the mapping $\nu_{\mathbf{F}}$ from fibre end-to-end vectors in the undeformed state to vectors in the deformed state. The simplest candidate for such a mapping is the affine transformation $\nu_{\mathbf{F}}(\bar{r}_0\hat{\mathbf{r}}_0) = \bar{r}_0\mathbf{F}\hat{\mathbf{r}}_0$, which is expected to be valid in highly over-constrained networks of stretching-dominated fibres [13, 114–116], i.e.

networks with high coordination numbers. When fibres deform affinely with the imposed deformation $\mathbf{F}$, the Cauchy stress is

$$\mathbf{T} = \frac{n_0 \bar{r}_0}{2^{d-1}\pi \det \mathbf{F}} \int_{\mathcal{S}^{d-1}} \mathbf{f}(\bar{r}_0 \mathbf{F}\hat{\mathbf{r}}_0) \otimes (\mathbf{F}\hat{\mathbf{r}}_0)\, \mathrm{d}A(\hat{\mathbf{r}}_0). \tag{8.8}$$

We use (8.8) to compute theoretical predictions of stress for the random networks considered in the rest of this chapter.

8.2 Networks of active fibres

We consider random networks consisting of fibres joined at nodes, which behave as freely-rotating joints, such as shown in Fig. 8.2(a). To obtain non-hyperelastic behaviour from such systems, we take inspiration from biology, specifically collagen networks with embedded fibroblasts [117–119]. Fibroblasts are cells that not only synthesise collagen fibres [118] but also develop protrusions which exert forces on the surrounding collagen matrix in response to network strain [117–119]. In this interaction, the cells integrate force and strain information from different directions in the network and respond by aligning and stiffening nearby fibres [118]. Similar behaviour, involving a positive mechanical feedback loop between cells and surrounding fibres, has also been observed with tumour cells embedded in collagen matrices [120]. The active mechanical response of fibroblasts is complex and not fully understood [117, 121], so proposing a model that accurately captures their response is beyond the scope of this thesis. Instead, we consider a simple random network inspired by these biological systems, in which the integration of strain information from different directions is accomplished by considering fibres behaving as springs whose stiffnesses depend on the stretches of neighbouring fibres.

We use $\mathcal{N}(m)$ to denote the set of indices corresponding to fibres that share a node with the mth fibre. Then, the force–length relation for the mth fibre has stiffness k_m that depends on the deformed and undeformed end-to-end vectors $\{(\mathbf{r}_l, \mathbf{r}_{l,0})\}_{l\in\mathcal{N}(m)}$ of the neighbouring fibres, taking the form

$$\mathbf{f}_m(\mathbf{r}_m) = k_m\left(\{(\mathbf{r}_l, \mathbf{r}_{l,0})\}_{l\in\mathcal{N}(m)}\right)\,(|\mathbf{r}_m| - |\mathbf{r}_{m,0}|)\hat{\mathbf{r}}_m, \tag{8.9}$$

where $\hat{\mathbf{r}}_m$ is the unit vector parallel to the fibre with deformed and undeformed end-to-end vectors $\mathbf{r}_m$ and $\mathbf{r}_{m,0}$, respectively. We consider stiffnesses k_m that depend on the average

squared stretch of the neighbouring fibres $\mathcal{N}(m)$, i.e.

$$k_m\left(\{(\mathbf{r}_l,\mathbf{r}_{l,0})\}_{l\in\mathcal{N}(m)}\right)=\bar{k}+\eta d\left(\left(\frac{1}{|\mathcal{N}(m)|}\sum_{l\in\mathcal{N}(m)}\left(\frac{|\mathbf{r}_l|}{|\mathbf{r}_{l,0}|}\right)^2\right)-1\right), \tag{8.10}$$

where $\bar{k}$ is a baseline stiffness, η controls how strongly the stretches of neighbouring fibres affect k_m, and $|\mathcal{N}(m)|$ is the number of elements in the set $\mathcal{N}(m)$. We refer to networks whose fibres have the force–length relation (8.9) with stiffness (8.10) as *local average networks*.

We use a mean-field approach to obtain an analytical constitutive model for local average networks. Assuming that fibres deform affinely with the imposed deformation $\mathbf{F}$, the average squared stretch of a fibre in an initially isotropic random network is

$$\frac{1}{2^{d-1}\pi}\int_{\mathcal{S}^{d-1}}|\mathbf{F}\hat{\mathbf{r}}_0|^2\,\mathrm{d}A(\hat{\mathbf{r}}_0)=\mathbf{C}:\frac{1}{2^{d-1}\pi}\int_{\mathcal{S}^{d-1}}\hat{\mathbf{r}}_0\otimes\hat{\mathbf{r}}_0\,\mathrm{d}A(\hat{\mathbf{r}}_0)=\frac{1}{d}\operatorname{tr}\mathbf{C}=1+\frac{2}{d}\operatorname{tr}\mathbf{E}, \tag{8.11}$$

where $d=2,3$ and we have used (II.38) to evaluate the integral. Assuming that all fibres have the same natural length $\bar{r}_0$, replacing the average squared stretch in (8.10) with (8.11) shows that the mean-field version of (8.9) is

$$\mathbf{f}(\mathbf{r})=\left(\bar{k}+2\eta\operatorname{tr}\mathbf{E}\right)\bar{r}_0(|\mathbf{F}\hat{\mathbf{r}}_0|-1)\frac{1}{|\mathbf{F}\hat{\mathbf{r}}_0|}\mathbf{F}\hat{\mathbf{r}}_0, \tag{8.12}$$

where $\mathbf{r}=\bar{r}_0\mathbf{F}\hat{\mathbf{r}}_0$.

8.3 Three-dimensional networks

Here we show that three-dimensional local average networks exhibit non-hyperelastic behaviour. We use the weakly nonlinear theory to obtain a theoretical prediction for the odd modulus d_o in terms of the active fibre parameter η and show that this prediction is a good approximation to numerical results. Substituting (8.12) into (8.8) leads to a constitutive relation, which we can express in terms of the second Piola–Kirchhoff stress using (6.5):

$$\tilde{\mathbf{S}}(\mathbf{C})=\frac{n_0\bar{r}_0^2}{4\pi}\left(\bar{k}+\eta(\operatorname{tr}\mathbf{C}-3)\right)\int_{\mathcal{S}^2}\frac{\sqrt{\mathbf{C}:\hat{\mathbf{r}}_0\otimes\hat{\mathbf{r}}_0}-1}{\sqrt{\mathbf{C}:\hat{\mathbf{r}}_0\otimes\hat{\mathbf{r}}_0}}\hat{\mathbf{r}}_0\otimes\hat{\mathbf{r}}_0\,\mathrm{d}A(\hat{\mathbf{r}}_0). \tag{8.13}$$

Using $\mathbf{C} = \mathbf{I} + 2\mathbf{E}$ and retaining terms up to $||\mathbf{E}||^2$ leads to

$$\tilde{\mathbf{S}}(\mathbf{I}+2\mathbf{E}) = \frac{n_0\bar{r}_0^2}{4\pi}\bar{k}\int_{\mathcal{S}^2}\left(\mathbf{E}:\hat{\mathbf{r}}_0\otimes\hat{\mathbf{r}}_0 - \frac{1}{2}(\mathbf{E}:\hat{\mathbf{r}}_0\otimes\hat{\mathbf{r}}_0)^2\right)(1-\mathbf{E}:\hat{\mathbf{r}}_0\otimes\hat{\mathbf{r}}_0)\hat{\mathbf{r}}_0\otimes\hat{\mathbf{r}}_0\,\mathrm{d}A(\hat{\mathbf{r}}_0)$$
$$+\frac{n_0\bar{r}_0^2}{4\pi}2\eta\,\mathrm{tr}\mathbf{E}\int_{\mathcal{S}^2}(\mathbf{E}:\hat{\mathbf{r}}_0\otimes\hat{\mathbf{r}}_0)\,(1-\mathbf{E}:\hat{\mathbf{r}}_0\otimes\hat{\mathbf{r}}_0)\hat{\mathbf{r}}_0\otimes\hat{\mathbf{r}}_0\,\mathrm{d}A(\hat{\mathbf{r}}_0)+O(||\mathbf{E}||^3).$$

We use the identities (II.39) and (II.44) to perform the integration over the unit sphere to obtain

$$\tilde{\mathbf{S}}(\mathbf{I}+2\mathbf{E}) = \frac{n_0\bar{r}_0^2}{105}\left(\left(7\bar{k}+\left(14\eta-\frac{3}{2}\bar{k}\right)\mathrm{tr}\mathbf{E}\right)(\mathrm{tr}\mathbf{E})\mathbf{I}+14\bar{k}\mathbf{E}-12\bar{k}\mathbf{E}^2\right)$$
$$+\frac{n_0\bar{r}_0^2}{105}(7\eta-3\bar{k})(\mathrm{tr}(\mathbf{E}^2)\mathbf{I}+2(\mathrm{tr}\mathbf{E})\mathbf{E})-\frac{n_0\bar{r}_0^2}{15}\eta(\mathrm{tr}(\mathbf{E}^2)\mathbf{I}-2(\mathrm{tr}\mathbf{E})\mathbf{E})+O(||\mathbf{E}||^3). \tag{8.14}$$

Comparing (8.14) and (7.22) shows that the odd elastic modulus is given by

$$\frac{d_o}{n_0\bar{r}_0^2\bar{k}} = -\frac{1}{15}\frac{\eta}{\bar{k}}, \tag{8.15}$$

so this mean-field constitutive relation is not hyperelastic when $\eta \neq 0$. We validate this mean-field prediction by comparing it to numerical simulations of random networks.

8.3.1 Network generation

The random networks used in our simulations, such as the example in Fig. 8.2(a), consist of two types of node. Interior nodes are randomly placed in the interior of the unit cube, whereas boundary nodes are placed with regular spacing on the boundaries of the unit cube. We use the boundary nodes to impose specified deformations on the network. The positions of the N_{int} interior nodes are generated using the following algorithm, adapted from Pósfai et al. [122], which ensures that nodes are separated by a distance of at least ℓ_{min}.

1. Initialisation: Each of the coordinates specifying the position of the first node is obtained as an independent random sample from the uniform distribution over the unit interval $(0,1)$.

2. Placement of the $(k+1)$th node: A candidate node position is obtained in the same way as the position of the first node. The candidate position is accepted if it is at least ℓ_{min} from the nodes $1,\ldots,k$ already placed. Otherwise, this step is repeated until a candidate position is accepted, or until a pre-set number of attempts is exceeded.

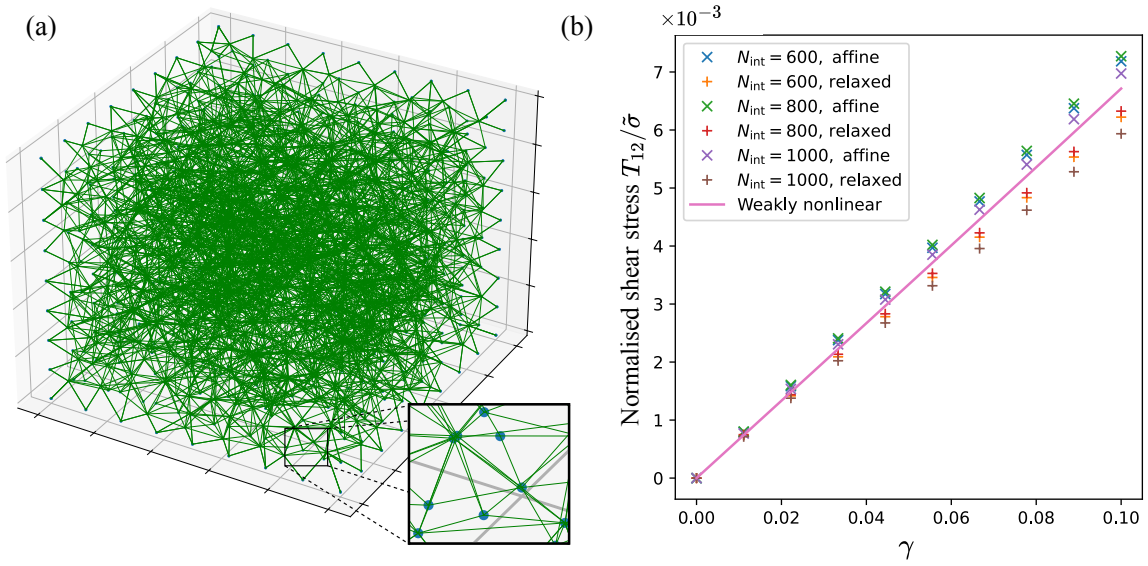


Fig. 8.2 Three-dimensional random networks and their shear responses. (a) An instance of a three-dimensional random network with $N_{\text{int}} = 1000$ interior nodes, connected by fibres, shown in green. (b) Plot of normalised shear stress $T_{12}/\tilde{\sigma}$ against shear strain γ, where $\tilde{\sigma}$ is defined in (8.17), when random networks with numbers of interior nodes $N_{\text{int}} = 600, 800, 1000$ are subjected to simple shear. Stresses computed for affine displacements of the interior nodes are shown using $\times$ markers and stresses for the network relaxed to force equilibrium are shown using $+$ markers. The analytical predictions from the weakly nonlinear theory (8.14) are shown as a solid line.

3. Termination: The algorithm terminates if all N_{int} interior nodes have been placed, or if a node cannot be placed within the number of allowed attempts.

We choose $\ell_{\text{min}} = \tilde{\ell} N_{\text{int}}^{-\frac{1}{3}}$ when placing interior points, where larger $\tilde{\ell}$ leads to a more homogeneous distribution of interior points but a lower probability of successful placement of all N_{int} points. We found that $\tilde{\ell} = 0.78$ worked well for the successful generation of homogeneous networks.

Each of the N_{f} fibres is indexed by an unordered pair $\{k, k'\}$ of node indices, where these indices can refer to either boundary or interior nodes. To generate networks in which two nearby nodes are more likely to be connected by a fibre than two nodes that are far apart, we compute the distances between all pairs of nodes, excluding boundary–boundary pairs, and connect the pairs with the N_{f} shortest separation distances with fibres. We do not place fibres corresponding to boundary–boundary pairs because these fibres would not be isotropically oriented, on average. The number of fibres N_{f} is chosen so that the coordination number of the network is approximately 12, ensuring that it is highly over-constrained.

8.3.2 Measuring stresses

We impose a specified deformation $\mathbf{F}$ on each network by displacing boundary nodes affinely, i.e. a boundary node with position vector $\mathbf{x}$ in the undeformed state has position vector $\mathbf{Fx}$ in the deformed state. We then let the network relax to force equilibrium by allowing the interior nodes to move into positions where they experience no net forces from the fibres to which they are connected. The Cauchy stress of the network for given node positions is computed using (8.2), for force–length relation (8.9) with stiffness (8.10). This approach to calculating stress is consistent with that used in molecular dynamics simulations [60, 113], because the stress (8.2) corresponds to the part of the virial stress tensor that arises from inter-particle forces.

Network relaxation is modelled using overdamped Newtonian dynamics, in which the position $\mathbf{x}_k(t)$ of the kth interior node at time t satisfies

$$\zeta_d \dot{\mathbf{x}}_k(t) = \mathbf{f}_{\mathrm{tot},k}(\{\mathbf{x}_{k'}(t)\}), \tag{8.16}$$

where ζ_d is a damping factor and $\mathbf{f}_{\mathrm{tot},k}(\{\mathbf{x}_{k'}\})$ is the total force on the kth node due to the fibres attached to it, computed from the positions $\mathbf{x}_{k'}$ of other nodes, both boundary and interior, that are connected to the kth node by fibres. We obtain the equilibrium positions by solving (8.16) numerically, with the positions of the boundary nodes fixed by the imposed deformation. Because non-affine displacements of the network nodes are not expected to be large, we choose the initial positions $\mathbf{x}_k(0)$ to be the positions of interior nodes displaced affinely according to the imposed deformation $\mathbf{F}$. The relaxed positions of the interior nodes are the values of $\{\mathbf{x}_k(t)\}$, where k indexes the interior nodes, at time t sufficiently large for all the node velocity components to be closer to zero than some pre-set threshold. We choose this velocity threshold to be $10^{-3}\frac{\bar{k}}{\zeta_d}L_{\mathrm{net}}$, where L_{net} is the network size, here equal to 1 because the network occupies the unit cube. This threshold is established by imposing a simple shear deformation on the network with $\gamma = 0.1$ and measuring the shear stress after relaxation for decreasing velocity thresholds $10^{-r}\frac{\bar{k}}{\zeta_d}L_{\mathrm{net}}$, $r = 2,3,4,5,6$. The change in measured stress between the thresholds $10^{-5}\frac{\bar{k}}{\zeta_d}L_{\mathrm{net}}$ and $10^{-6}\frac{\bar{k}}{\zeta_d}L_{\mathrm{net}}$ is 0.004%, so we can take the stress at velocity threshold $10^{-6}\frac{\bar{k}}{\zeta_d}L_{\mathrm{net}}$ to be the actual stress in the relaxed network. The difference between this actual stress and the stress at the velocity threshold $10^{-3}\frac{\bar{k}}{\zeta_d}L_{\mathrm{net}}$ is 0.7%, so we use this threshold in our simulations to obtain good accuracy within a reasonable computation time.

8.3.3 Results and discussion

We non-dimensionalise network stress by expressing it in units of

$$\tilde{\sigma} = \frac{\bar{k}}{\Omega} \sum_{m=1}^{N_\mathrm{f}} r_{0,m}^2, \tag{8.17}$$

where $\bar{k}$ is the baseline fibre stiffness in (8.10), Ω is the network volume, and $r_{0,m}$ is the unstretched length of the mth fibre. Analytical stress predictions for each random network are obtained from the weakly nonlinear theory (8.14) by replacing $n_0\bar{r}_0^2$ with $\Omega^{-1}\sum_{m=1}^{N_\mathrm{f}} r_{0,m}^2$. With this normalisation, the analytical predictions are independent of network size and number of fibres. Therefore, the theory predicts the same normalised stresses for different networks that are subjected to the same deformation, provided the networks have the same $\eta/\bar{k}$ and are sufficiently isotropic for the theory to be applicable.

We impose the simple shear deformation (4.9) on random networks of the type shown in Fig. 8.2(a) and plot the normalised shear stress $\tilde{\sigma}^{-1}T_{12}$ against shear strain γ in Fig. 8.2(b) for networks containing different numbers $N_\mathrm{int} = 600, 800, 1000$ of interior nodes. Besides the stresses computed from the positions of the interior nodes obtained by relaxation to force equilibrium in the deformed network, we also plot the stresses computed by assuming that all nodes are displaced affinely with the imposed deformation gradient $\mathbf{F}$. Although the affine stresses overestimate the stresses obtained from relaxation to force equilibrium, we see that the affine assumption is a reasonable approximation for small shear strains in these highly over-constrained networks. The similarity between the values of normalised stress for different networks indicates that the networks we consider have a sufficiently large number of nodes to be treated in the continuum limit. The predictions of the weakly nonlinear theory (8.14) are a good approximation to the actual stresses in this regime.

To show that the stress–strain relations of these random networks are non-hyperelastic, we measure the odd modulus d_o by imposing the deformation (7.27) and computing the derivatives in (7.28), for different values of the parameter $\eta/\bar{k}$. We plot the values of $d_o/\tilde{\sigma}$ against $\eta/\bar{k}$ for measurements performed under both affine and relaxed (force equilibrium) conditions in Fig. 8.3. These measured values are well-approximated by the prediction (8.15) from the weakly nonlinear theory (8.14), confirming that these results are a manifestation of non-hyperelastic mechanical response. Importantly, these results show that experimental realisations of local average fibres with only passive components are impossible, because these fibres necessarily involve components that consume energy in some form.

We also consider networks containing a random mixture of active and passive fibres, such as the network shown in Fig. 8.4(a) in which active fibres are green and passive ones are

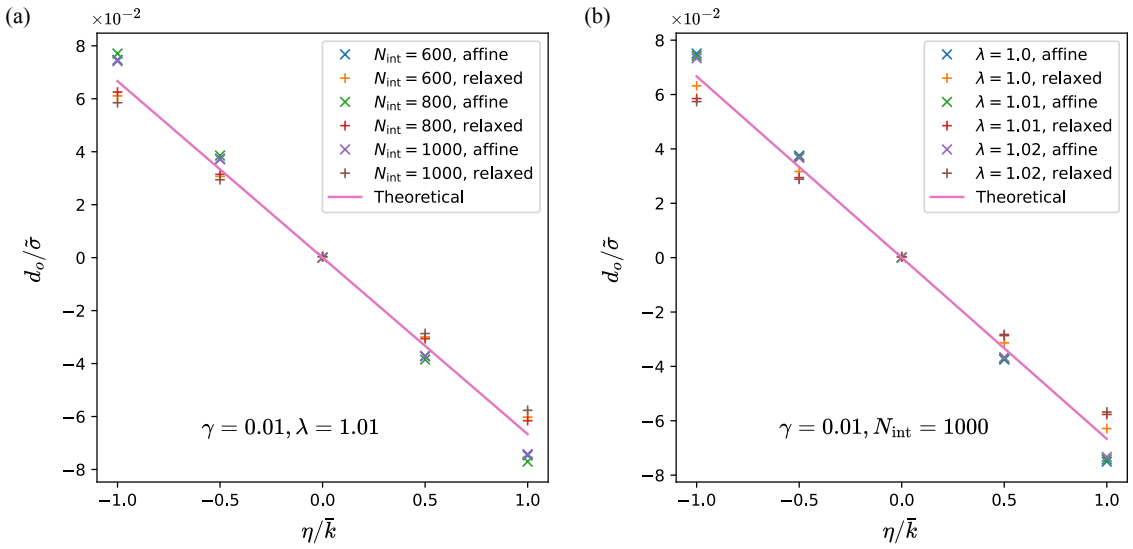


Fig. 8.3 Plots of the normalised odd modulus $d_o/\tilde{\sigma}$ against the parameter $\eta/\bar{k}$, shown with the theoretical prediction (8.15). (a) The odd modulus d_o is computed using (7.28) for three different random networks, with numbers of interior nodes $N_{int} = 600, 800, 1000$. For each network, the derivatives in (7.28) are evaluated at $\gamma = 0.01, \lambda = 1.01$. (b) The odd modulus is computed for a network with $N_{\text{int}} = 1000$, at $\gamma = 0.01$ and different stretches $\lambda = 1.0, 1.01, 1.02$.

black. The passive fibres behave as Hookean springs with stiffness $\bar{k}$, i.e. their force–length relation is obtained by setting $\eta = 0$ in (8.10). We vary the proportion p_{a} of active fibres in the network for fixed $\eta/\bar{k} = 0.5$ and plot the measured values of $d_o/\tilde{\sigma}$ on Fig. 8.4(b). If fibre end-to-end vectors transform affinely with the imposed deformation, then the network stress is given by (8.14) with η replaced by $p_{\text{a}}\eta$, i.e. by the rule of mixtures. Therefore, under the affine assumption, the odd modulus d_o is given by

$$\frac{d_o}{\tilde{\sigma}} = -\frac{1}{15}\frac{\eta}{\bar{k}}p_{\text{a}}, \tag{8.18}$$

which we plot on Fig. 8.4(b) as a solid line. We see that (8.18) is a reasonable approximation to the measured d_o values, consistent with the non-affine displacements of the network being relatively small. Therefore, active elasticity can be observed in local average networks even when not all fibres are active.

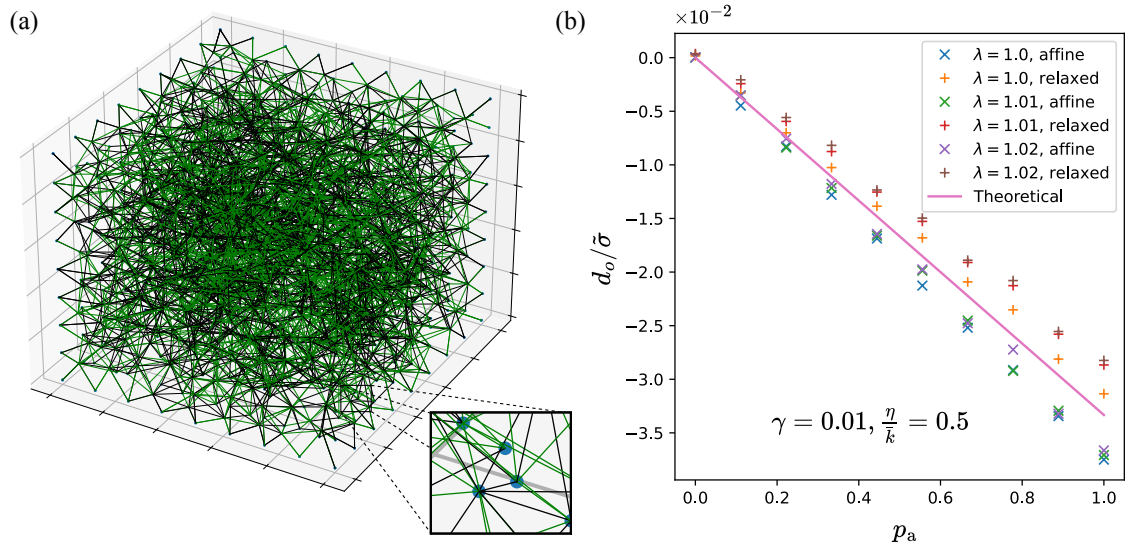


Fig. 8.4 The dependence of the odd modulus d_o on the fraction p_a of active fibres in a random network consisting of a mixture of active and passive fibres. (a) A three-dimensional network with a mixture of active and passive fibres, where the active fibres are shown in green and the passive in black. This network has number of interior nodes $N_{\text{int}} = 1000$ and proportion of active fibres $p_a = 0.5$. (b) Plot of the normalised odd modulus $d_o/\tilde{\sigma}$ against the proportion p_a of active fibres, where d_o is computed using (7.28) evaluated at $\gamma = 0.01, \frac{\eta}{\bar{k}} = 0.5$ and at different λ values. The theoretical prediction is given by (8.18).

8.4 Two-dimensional networks

Because the odd-elastic behaviour controlled by d_o in two-dimensional materials is similar to that in three-dimensional materials, here we focus on comparing two-dimensional local average networks to those consisting of fibres that exert transverse forces. These fibres have the force–length relation

$$\mathbf{f}(\mathbf{r}) = \bar{k}(|\mathbf{r}| - r_0)\hat{\mathbf{r}} + \tilde{\eta}(|\mathbf{r}| - r_0)\mathbf{W}\hat{\mathbf{r}}, \tag{8.19}$$

where $\mathbf{r}$ is the fibre end-to-end vector, r_0 is the unstretched length of the fibre, and $\bar{k}, \tilde{\eta}$ are the stiffnesses corresponding to the longitudinal and transverse forces exerted by the fibre, respectively. This fibre force–length relation is similar to the one used in Scheibner et al. [10] for lattice-based odd-elastic materials. We refer to random networks consisting of fibres with the force–length relation (8.19) as *transverse force networks*.

For two-dimensional local average networks, the weakly nonlinear theory gives the stress–strain relation

$$\tilde{\mathbf{S}}(\mathbf{I}+2\mathbf{E}) = \frac{n_0\bar{r}_0^2}{2}\left(\left(\frac{1}{4}\bar{k}+\left(\frac{1}{2}\eta+\frac{3}{16}\bar{k}\right)\operatorname{tr}\mathbf{E}\right)(\operatorname{tr}\mathbf{E})\mathbf{I}+\frac{1}{2}\bar{k}\mathbf{E}\right) + \frac{n_0\bar{r}_0^2}{8}\left(\eta-6\bar{k}\right)(\operatorname{tr}(\mathbf{E}^2)\mathbf{I}+2(\operatorname{tr}\mathbf{E})\mathbf{E}) - \frac{n_0\bar{r}_0^2}{8}\eta(\operatorname{tr}(\mathbf{E}^2)\mathbf{I}-2(\operatorname{tr}\mathbf{E})\mathbf{E})+O(||\mathbf{E}||^3), \tag{8.20}$$

where the integration over the unit sphere was performed using the identities (II.39) and (II.44). The analytical prediction for the odd modulus d_o takes the same form as for the three-dimensional case (8.15) but with a different numerical pre-factor:

$$\frac{d_o}{n_0\bar{r}_0^2\bar{k}} = -\frac{1}{8}\frac{\eta}{\bar{k}}. \tag{8.21}$$

8.4.1 Network generation and stress measurement

We generate two-dimensional random networks using the same algorithm as given for three-dimensional networks, with the additional constraint that fibres are not allowed to intersect. An example of a two-dimensional random network is shown in Fig. 8.5(a). Stresses are measured in the same way as used for three-dimensional networks, but with the velocity threshold set to $10^{-4}\frac{\bar{k}}{\zeta_d}L_{\text{net}}$ because network relaxation is less computationally expensive for two-dimensional networks.

8.4.2 Results and discussion

We impose a simple shear deformation on local average networks and compare the predictions of the weakly nonlinear theory (8.20) to the measured stresses. Fig. 8.5(b) shows the normalised shear stress $\tilde{\sigma}^{-1}T_{12}$ plotted against shear strain γ for local average networks containing different numbers $N_{\text{int}} = 80, 90, 100$ of interior nodes. The agreement between normalised stresses across the different local average networks and the predictions (8.20) shows that these numbers of interior nodes are sufficiently large for the continuum theory to be valid. As in the three-dimensional case, using the affine assumption overestimates stresses, but this assumption is a good approximation at small shear strains.

To show that the stress–strain behaviour of two-dimensional local average networks is non-hyperelastic, we compute the odd modulus d_o for several local average networks with different numbers of interior nodes using (7.38), with the chiral moduli $K^o, \bar{K}^o, A, \bar{A}, \tilde{A}$ set to zero, because chiral symmetry is not broken in local average networks. The measured values of $d_o/\tilde{\sigma}$ are plotted against $\eta/\bar{k}$ on Fig. 8.6(a). The analytical prediction (8.21) based on the

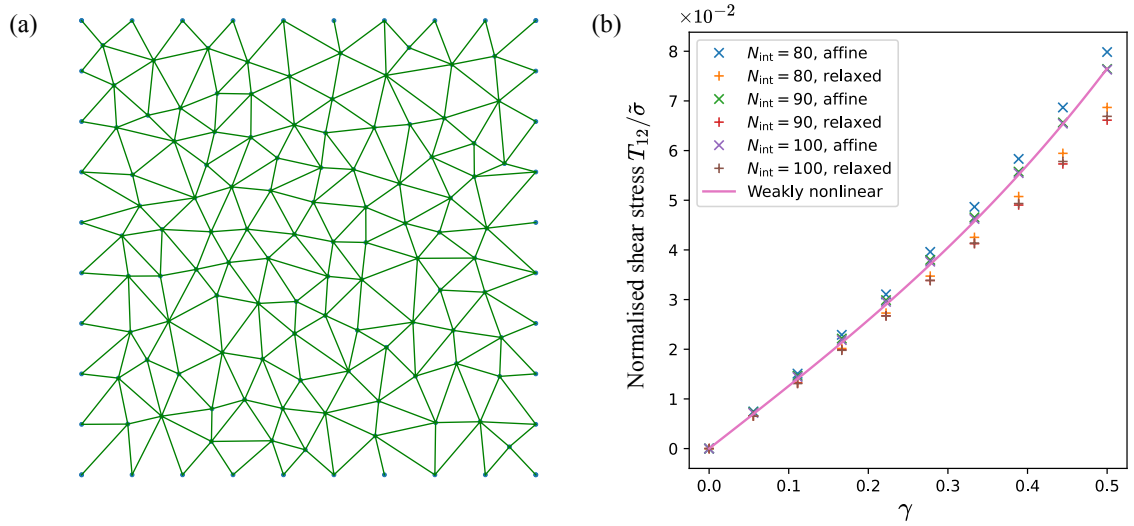


Fig. 8.5 Two-dimensional random networks and their shear responses. (a) A two-dimensional random network with $N_{\text{int}} = 100$ interior nodes, connected by fibres shown in green. (b) Plot of normalised shear stress $T_{12}/\tilde{\sigma}$ against shear strain γ when local average networks with $N_{\text{int}} = 80, 90, 100$ are subjected to simple shear. Stresses computed for affine displacements of the interior nodes are shown using $\times$ markers and stresses for the network relaxed to force equilibrium are shown using $+$ markers. The analytical predictions from the weakly nonlinear theory (8.20) are shown as a solid line.

weakly nonlinear theory is a good approximation to the measured values, confirming that two-dimensional local average networks exhibit non-hyperelastic mechanical responses.

We compare local average fibres (8.9), (8.10) and transverse force fibres (8.19) using the same network geometry by plotting $\tilde{\sigma}^{-1}\left(\frac{\partial T_{22}}{\partial \gamma} - \frac{\partial T_{12}}{\partial \lambda}\right)$ against shear strain γ in Fig. 8.6(b). For local average fibres, this difference in derivatives (7.38) is odd in γ, confirming that this network has $d_o \neq 0$ and zero chiral odd moduli $K^o, \bar{K}^o, A, \bar{A}, \tilde{A}$. In contrast, for transverse force fibres, this quantity (7.38) is even in γ with a dominant constant contribution, showing that $d_o = 0$ and that some chiral odd moduli are non-zero. These results show that the different forms of isotropic odd elasticity discussed in Section 7.3.2, i.e. achiral, deformation-induced odd elasticity controlled by d_o and chiral odd elasticity controlled by $K^o, \bar{K}^o, A, \bar{A}, \tilde{A}$, can be realised in random networks using local average and transverse force fibres, respectively.

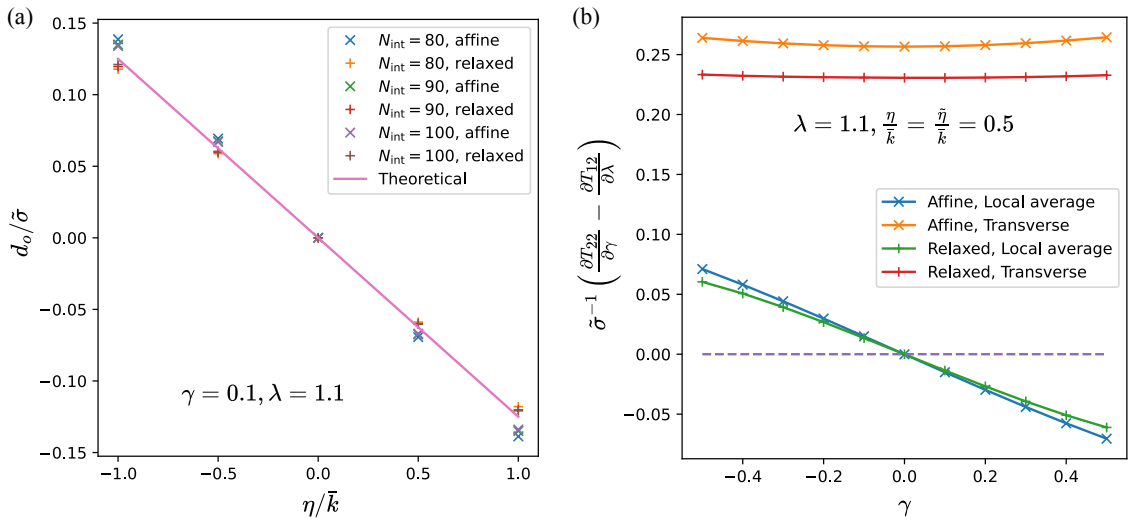


Fig. 8.6 Comparison of local average and transverse force networks. (a) The odd modulus d_o is computed using (7.38) for local average networks with different numbers of interior nodes N_{int}. The derivatives in (7.38) are evaluated at $\gamma = 0.1, \lambda = 1.1$. The analytical prediction (8.21) is shown as a solid line. (b) Plot of the difference in derivatives $\tilde{\sigma}^{-1}\left(\frac{\partial T_{22}}{\partial \gamma} - \frac{\partial T_{12}}{\partial \lambda}\right)$ against shear strain γ for a network with $N_{\text{int}} = 100$, using both local average and transverse force fibres. The network parameters are $\eta/\bar{k} = \tilde{\eta}/\bar{k} = 0.5$ and the derivatives are evaluated at $\lambda = 1.1$.

Chapter 9

Conclusions: nonlinear active elasticity

We identified a novel odd elastic modulus d_o that characterises the contributions of terms quadratic in strain to the stress response of a soft, isotropic active solid. Unlike the odd moduli in linear, isotropic odd elasticity, this modulus does not rely on broken chiral symmetry, so it is generically present in the elasticity of both two- and three-dimensional active solids. Because odd elasticity controlled by this modulus is induced by material deformation, our results point towards adaptive applications of active solids in which the strength of the odd response can be tuned by mechanical input. We also expanded the class of deformation cycles that can be used for work extraction, by showing that large deformations enable non-zero work to be extracted using cycles that would lead to zero work extraction in the linear regime. These results provide the theoretical foundations for future applications of soft active solids in energy harvesting.

Besides studying nonlinear active elasticity from a purely continuum perspective, we also presented specific microscopic realisations based on isotropic random networks of active fibres. These networks demonstrated a novel route towards odd elasticity: when fibres behave as springs whose stiffnesses depend on the average squared stretch of neighbouring fibres, the odd modulus d_o of the network is non-zero. This phenomenology is not chiral, so it is qualitatively distinct from that of networks consisting of fibres that exert transverse forces, which break chiral symmetry. These microscopic realisations are a step towards creating novel soft materials that enlarge the toolkit available to designers of active solids.

Our results motivate further work on the phenomenology of nonlinear active elasticity, such as studying the influence of the odd modulus d_o on the elastodynamics of soft solids [123], e.g. the propagation of nonlinear waves. Experimental realisations of local average networks, or similar materials, could be possible using robotic components, as has been done for realisations of chiral odd elasticity [100]. Although our local average networks were inspired by collagen networks containing fibroblasts, careful experimental

measurements are necessary to determine if these biological systems exhibit odd elasticity controlled by d_o.

Appendix II

II.1 Antisymmetric part of of the elasticity tensor

Here we differentiate (7.3) to express the elasticity tensor $\mathsf{C}_{\mathbf{F}_1}$ in terms of the stress function $\tilde{\mathbf{S}}$ in (6.11). Using (6.5), we obtain

$$\begin{aligned}\bar{\mathbf{P}}_1(\bar{\mathbf{F}}) &= (\det\bar{\mathbf{F}})\left(\frac{1}{\det\bar{\mathbf{F}}\mathbf{F}_1}(\bar{\mathbf{F}}\mathbf{F}_1)\,\hat{\mathbf{S}}\left(\bar{\mathbf{F}}\mathbf{F}_1\right)(\bar{\mathbf{F}}\mathbf{F}_1)^{\mathrm{T}}\right)\bar{\mathbf{F}}^{-\mathrm{T}}\\ &= \frac{1}{\det\mathbf{F}_1}\bar{\mathbf{F}}\mathbf{F}_1\,\tilde{\mathbf{S}}\left(\mathbf{F}_1^{\mathrm{T}}\bar{\mathbf{F}}^{\mathrm{T}}\bar{\mathbf{F}}\mathbf{F}_1\right)\mathbf{F}_1^{\mathrm{T}}.\end{aligned} \tag{II.1}$$

To compute the derivative of this function, let $\tilde{\mathbf{C}}(\mathbf{F}) = \mathbf{F}^{\mathrm{T}}\mathbf{F}$. Then,

$$\mathrm{D}_{\mathbf{F}}\tilde{\mathbf{C}}(\mathbf{F})[\mathbf{H}] = \mathbf{H}^{\mathrm{T}}\mathbf{F} + \mathbf{F}^{\mathrm{T}}\mathbf{H} = 2\,\mathrm{sym}\,(\mathbf{H}^{\mathrm{T}}\mathbf{F}). \tag{II.2}$$

By the chain rule, for $\mathbf{H} \in \mathtt{L}(\mathcal{V}_d)$,

$$\begin{aligned}\mathrm{D}_{\bar{\mathbf{F}}}\bar{\mathbf{P}}_1(\bar{\mathbf{F}})[\mathbf{H}] = {}& \frac{1}{\det\mathbf{F}_1}\mathbf{H}\mathbf{F}_1\,\tilde{\mathbf{S}}\left(\mathbf{F}_1^{\mathrm{T}}\bar{\mathbf{F}}^{\mathrm{T}}\bar{\mathbf{F}}\mathbf{F}_1\right)\mathbf{F}_1^{\mathrm{T}}\\ &+ \frac{1}{\det\mathbf{F}_1}\bar{\mathbf{F}}\mathbf{F}_1\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}\left(\mathbf{F}_1^{\mathrm{T}}\bar{\mathbf{F}}^{\mathrm{T}}\bar{\mathbf{F}}\mathbf{F}_1\right)\left[\mathbf{F}_1^{\mathrm{T}}(\mathrm{D}_{\mathbf{F}}\tilde{\mathbf{C}}(\bar{\mathbf{F}})[\mathbf{H}])\mathbf{F}_1\right]\mathbf{F}_1^{\mathrm{T}},\end{aligned} \tag{II.3}$$

which implies

$$\mathrm{D}_{\bar{\mathbf{F}}}\bar{\mathbf{P}}_1(\mathbf{I})[\mathbf{H}] = \mathbf{H}\hat{\mathbf{T}}(\mathbf{F}_1) + \frac{1}{\det\mathbf{F}_1}\mathbf{F}_1\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}\left(\mathbf{F}_1^{\mathrm{T}}\mathbf{F}_1\right)[\mathbf{F}_1^{\mathrm{T}}(2\,\mathrm{sym}\,\mathbf{H})\mathbf{F}_1]\mathbf{F}_1^{\mathrm{T}}. \tag{II.4}$$

The elasticity tensor $\mathsf{C}_{\mathbf{F}_1}$ can therefore be decomposed into $\mathsf{C}_{\mathbf{F}_1} = \mathsf{T}_{\mathbf{F}_1} + \mathsf{S}_{\mathbf{F}_1}$, where

$$\mathsf{T}_{\mathbf{F}_1}[\mathbf{H}] = \mathbf{H}\hat{\mathbf{T}}(\mathbf{F}_1), \quad \mathsf{S}_{\mathbf{F}_1}[\mathbf{H}] = \frac{1}{\det\mathbf{F}_1}\mathbf{F}_1\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}\left(\mathbf{F}_1^{\mathrm{T}}\mathbf{F}_1\right)[\mathbf{F}_1^{\mathrm{T}}(2\,\mathrm{sym}\,\mathbf{H})\mathbf{F}_1]\mathbf{F}_1^{\mathrm{T}}. \tag{II.5}$$

For any $\mathbf{G},\mathbf{H}\in \mathtt{L}(\mathcal{V}_d)$, using the symmetry of the Cauchy stress tensor $\hat{\mathbf{T}}(\mathbf{F}_1)$ gives

$$\begin{aligned}\mathbf{G}:\mathsf{T}_{\mathbf{F}_1}[\mathbf{H}] &= \operatorname{tr}\left(\mathbf{G}^{\mathrm{T}}\mathbf{H}\hat{\mathbf{T}}(\mathbf{F}_1)\right)=\operatorname{tr}\left(\hat{\mathbf{T}}(\mathbf{F}_1)\mathbf{G}^{\mathrm{T}}\mathbf{H}\right)=\operatorname{tr}\left(\hat{\mathbf{T}}(\mathbf{F}_1)^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{H}\right)=\operatorname{tr}\left((\mathbf{G}\hat{\mathbf{T}}(\mathbf{F}_1))^{\mathrm{T}}\mathbf{H}\right)\\ &=\mathbf{H}:\mathsf{T}_{\mathbf{F}_1}[\mathbf{G}],\end{aligned} \tag{II.6}$$

so the linear map $\mathsf{T}_{\mathbf{F}_1}$ is symmetric, i.e. this fourth-order tensor has major symmetry. Therefore, the antisymmetric part (odd elastic moduli) of the elasticity tensor $\mathsf{C}_{\mathbf{F}_1}$ comes from $\mathsf{S}_{\mathbf{F}_1}$. To compute $\operatorname{asym}\mathsf{S}_{\mathbf{F}_1}$, we derive the transpose of $\mathsf{S}_{\mathbf{F}_1}$. For any $\mathbf{G},\mathbf{H}\in \mathtt{L}(\mathcal{V}_d)$, and using the symmetry of $\tilde{\mathbf{S}}(\mathbf{C})$ for all $\mathbf{C}\in \operatorname{Psym}(d)$,

$$\begin{aligned}\mathbf{H}:\mathsf{S}^{\mathrm{T}}_{\mathbf{F}_1}[\mathbf{G}] &= \mathbf{G}:\mathsf{S}_{\mathbf{F}_1}[\mathbf{H}]=\mathbf{G}:\left(\frac{2}{\det\mathbf{F}_1}\mathbf{F}_1\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}\left(\mathbf{F}_1^{\mathrm{T}}\mathbf{F}_1\right)[\mathbf{F}_1^{\mathrm{T}}(\operatorname{sym}\mathbf{H})\mathbf{F}_1]\mathbf{F}_1^{\mathrm{T}}\right)\\ &=\operatorname{tr}\left(\mathbf{G}^{\mathrm{T}}\frac{2}{\det\mathbf{F}_1}\mathbf{F}_1\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}\left(\mathbf{F}_1^{\mathrm{T}}\mathbf{F}_1\right)[\mathbf{F}_1^{\mathrm{T}}(\operatorname{sym}\mathbf{H})\mathbf{F}_1]\mathbf{F}_1^{\mathrm{T}}\right)\\ &=\frac{2}{\det\mathbf{F}_1}\operatorname{tr}\left((\mathbf{F}_1^{\mathrm{T}}\mathbf{G}\mathbf{F}_1)^{\mathrm{T}}\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}\left(\mathbf{F}_1^{\mathrm{T}}\mathbf{F}_1\right)[\mathbf{F}_1^{\mathrm{T}}(\operatorname{sym}\mathbf{H})\mathbf{F}_1]\right)\\ &=\frac{2}{\det\mathbf{F}_1}\operatorname{tr}\left((\mathbf{F}_1^{\mathrm{T}}(\operatorname{sym}\mathbf{G})\mathbf{F}_1)^{\mathrm{T}}\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}\left(\mathbf{F}_1^{\mathrm{T}}\mathbf{F}_1\right)[\mathbf{F}_1^{\mathrm{T}}(\operatorname{sym}\mathbf{H})\mathbf{F}_1]\right)\\ &=\frac{2}{\det\mathbf{F}_1}(\mathbf{F}_1^{\mathrm{T}}(\operatorname{sym}\mathbf{G})\mathbf{F}_1):\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}\left(\mathbf{F}_1^{\mathrm{T}}\mathbf{F}_1\right)[\mathbf{F}_1^{\mathrm{T}}(\operatorname{sym}\mathbf{H})\mathbf{F}_1]\\ &=\frac{2}{\det\mathbf{F}_1}(\mathbf{F}_1^{\mathrm{T}}(\operatorname{sym}\mathbf{H})\mathbf{F}_1):\left(\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}\left(\mathbf{F}_1^{\mathrm{T}}\mathbf{F}_1\right)\right)^{\mathrm{T}}[\mathbf{F}_1^{\mathrm{T}}(\operatorname{sym}\mathbf{G})\mathbf{F}_1]\\ &=\mathbf{H}:\left(\frac{2}{\det\mathbf{F}_1}\mathbf{F}_1\left(\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}\left(\mathbf{F}_1^{\mathrm{T}}\mathbf{F}_1\right)\right)^{\mathrm{T}}[\mathbf{F}_1^{\mathrm{T}}(\operatorname{sym}\mathbf{G})\mathbf{F}_1]\mathbf{F}_1^{\mathrm{T}}\right),\end{aligned} \tag{II.7}$$

which implies that

$$\mathsf{S}^{\mathrm{T}}_{\mathbf{F}_1}[\mathbf{G}]=\frac{2}{\det\mathbf{F}_1}\mathbf{F}_1\left(\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}\left(\mathbf{F}_1^{\mathrm{T}}\mathbf{F}_1\right)\right)^{\mathrm{T}}[\mathbf{F}_1^{\mathrm{T}}(\operatorname{sym}\mathbf{G})\mathbf{F}_1]\mathbf{F}_1^{\mathrm{T}}, \tag{II.8}$$

from which (7.6) follows immediately.

II.2 Hyperelasticity and major symmetry

Here we establish that an elastic constitutive relation given in terms of the second Piola–Kirchhoff stress $\mathbf{S}=\tilde{\mathbf{S}}(\mathbf{C})$ is hyperelastic if and only if the derivative $\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}(\mathbf{C})$ has major symmetry for all positive-definite, symmetric second-order tensors $\mathbf{C}$. Although this result is a special case of the Poincaré Lemma [124], we provide an elementary proof that does not rely on manifold theory and differential forms.

We first suppose that the constitutive relation is hyperelastic, with first Piola–Kirchhoff stress given by $\hat{\mathbf{P}}(\mathbf{F}) = \frac{\mathrm{d}W}{\mathrm{d}\mathbf{F}}(\mathbf{F})$, and show that the derivative $\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}(\mathbf{C})$ has major symmetry for all positive-definite, symmetric second-order tensors $\mathbf{C}$. By the principle of material frame-indifference, the strain energy function W satisfies $W(\mathbf{QF}) = W(\mathbf{F})$ for all $\mathbf{Q} \in \mathrm{SO}(d)$. Therefore, by setting $\mathbf{Q} = \mathbf{R}^{\mathrm{T}}$, where $\mathbf{R}$ comes from the polar decomposition $\mathbf{F} = \mathbf{RU}$, we see that there exists a function $\widetilde{W}$ such that $W(\mathbf{F}) = \widetilde{W}(\mathbf{C})$, when $\mathbf{C} = \mathbf{F}^{\mathrm{T}}\mathbf{F}$. Using the same reasoning as employed in (I.2), the second Piola–Kirchhoff stress for a hyperelastic material is given by

$$\tilde{\mathbf{S}}(\mathbf{C}) = 2\frac{\mathrm{d}\widetilde{W}}{\mathrm{d}\mathbf{C}}(\mathbf{C}). \tag{II.9}$$

To show that hyperelasticity implies major symmetry, we define

$$\beta(s_1, s_2) = 2\widetilde{W}(\mathbf{C} + s_1\mathbf{C}_1 + s_2\mathbf{C}_2), \tag{II.10}$$

where $s_1, s_2 \in \mathbb{R}$ and $\mathbf{C}, \mathbf{C}_1, \mathbf{C}_2$ are positive-definite, symmetric second-order tensors. By the chain rule,

$$\begin{aligned} \frac{\partial^2\beta}{\partial s_2 \partial s_1}(s_1, s_2) &= \frac{\partial}{\partial s_2}\tilde{\mathbf{S}}(\mathbf{C} + s_1\mathbf{C}_1 + s_2\mathbf{C}_2) : \mathbf{C}_1 \\ &= \mathbf{C}_1 : \mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}(\mathbf{C} + s_1\mathbf{C}_1 + s_2\mathbf{C}_2)[\mathbf{C}_2]. \end{aligned} \tag{II.11}$$

By the symmetry of cross-derivatives $\frac{\partial^2\beta}{\partial s_2 \partial s_1}(s_1, s_2) = \frac{\partial^2\beta}{\partial s_1 \partial s_2}(s_1, s_2)$, we obtain

$$\mathbf{C}_1 : \mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}(\mathbf{C})[\mathbf{C}_2] = \mathbf{C}_2 : \mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}(\mathbf{C})[\mathbf{C}_1] \tag{II.12}$$

at $s_1 = s_2 = 0$. The arbitrariness of $\mathbf{C}, \mathbf{C}_1, \mathbf{C}_2$ implies that $\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}(\mathbf{C})$ has major symmetry for all positive-definite, symmetric second-order tensors $\mathbf{C}$.

We now prove the converse: if $\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}(\mathbf{C})$ has major symmetry for all $\mathbf{C} \in \mathrm{Psym}(d)$, then the elastic constitutive relation can be derived from a strain energy function. Let $\tilde{\mathbf{C}}(\xi, \mathbf{C})$ be a twice continuously differentiable function of ξ and $\mathbf{C}$, defining a path in $\mathrm{Psym}(d)$ parametrised by $\xi \in [\xi_0, \xi_1]$, connecting some fixed reference tensor $\mathbf{C}_0$ to $\mathbf{C}$, i.e. $\tilde{\mathbf{C}}(\xi_0, \mathbf{C}) = \mathbf{C}_0$ and $\tilde{\mathbf{C}}(\xi_1, \mathbf{C}) = \mathbf{C}$. We define

$$\widehat{W}(\mathbf{C}) = \frac{1}{2}\int_{\xi_0}^{\xi_1} \tilde{\mathbf{S}}(\tilde{\mathbf{C}}(\xi, \mathbf{C})) : \frac{\partial\tilde{\mathbf{C}}}{\partial\xi}(\xi, \mathbf{C})\,\mathrm{d}\xi, \tag{II.13}$$

and show that we can construct a strain energy function from $\widehat{W}$ if the major symmetry hypothesis is satisfied. Differentiating $\widehat{W}(\mathbf{C})$ gives

$$\begin{aligned} 2\mathrm{D}_{\mathbf{C}}\widehat{W}(\mathbf{C})[\mathbf{H}] = \int_{\xi_0}^{\xi_1} \mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}(\tilde{\mathbf{C}}(\xi,\mathbf{C}))\left[\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{C}}(\xi,\mathbf{C})[\mathbf{H}]\right] : \frac{\partial\tilde{\mathbf{C}}}{\partial\xi}(\xi,\mathbf{C}) \\ + \tilde{\mathbf{S}}(\tilde{\mathbf{C}}(\xi,\mathbf{C})) : \mathrm{D}_{\mathbf{C}}\left(\frac{\partial\tilde{\mathbf{C}}}{\partial\xi}(\xi,\mathbf{C})\right)[\mathbf{H}]\,\mathrm{d}\xi, \end{aligned} \tag{II.14}$$

for second-order tensors $\mathbf{H}$. Exchanging the order of derivatives in the second term and integrating by parts gives

$$\begin{aligned} &\int_{\xi_0}^{\xi_1} \tilde{\mathbf{S}}(\tilde{\mathbf{C}}(\xi,\mathbf{C})) : \frac{\partial}{\partial\xi}\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{C}}(\xi,\mathbf{C})[\mathbf{H}]\,\mathrm{d}\xi \\ &= \tilde{\mathbf{S}}(\tilde{\mathbf{C}}(\xi_1,\mathbf{C})) : \mathrm{D}_{\mathbf{C}}\tilde{\mathbf{C}}(\xi_1,\mathbf{C})[\mathbf{H}] - \tilde{\mathbf{S}}(\tilde{\mathbf{C}}(\xi_0,\mathbf{C})) : \mathrm{D}_{\mathbf{C}}\tilde{\mathbf{C}}(\xi_0,\mathbf{C})[\mathbf{H}] \\ &\qquad - \int_{\xi_0}^{\xi_1} \mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}(\tilde{\mathbf{C}}(\xi,\mathbf{C}))\left[\frac{\partial\tilde{\mathbf{C}}}{\partial\xi}(\xi,\mathbf{C})\right] : \mathrm{D}_{\mathbf{C}}\tilde{\mathbf{C}}(\xi,\mathbf{C})[\mathbf{H}]\,\mathrm{d}\xi. \end{aligned} \tag{II.15}$$

Substituting (II.15) into (II.14) gives

$$\begin{aligned} 2\mathrm{D}_{\mathbf{C}}\widehat{W}(\mathbf{C})[\mathbf{H}] = \tilde{\mathbf{S}}(\tilde{\mathbf{C}}(\xi_1,\mathbf{C})) : \mathrm{D}_{\mathbf{C}}\tilde{\mathbf{C}}(\xi_1,\mathbf{C})[\mathbf{H}] - \tilde{\mathbf{S}}(\tilde{\mathbf{C}}(\xi_0,\mathbf{C})) : \mathrm{D}_{\mathbf{C}}\tilde{\mathbf{C}}(\xi_0,\mathbf{C})[\mathbf{H}] \\ + \int_{\xi_0}^{\xi_1} \left(2\,\mathrm{asym}\,\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}(\tilde{\mathbf{C}}(\xi,\mathbf{C}))\right)\left[\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{C}}(\xi,\mathbf{C})[\mathbf{H}]\right] : \frac{\partial\tilde{\mathbf{C}}}{\partial\xi}(\xi,\mathbf{C})\,\mathrm{d}\xi, \end{aligned} \tag{II.16}$$

where $\mathrm{asym}\,\mathsf{C} = \frac{1}{2}(\mathsf{C} - \mathsf{C}^{\mathrm{T}})$ is the antisymmetric part of the fourth-order tensor C and the transpose of C satisfies $\mathbf{A}_1 : \mathsf{C}[\mathbf{A}_2] = \mathbf{A}_2 : \mathsf{C}^{\mathrm{T}}[\mathbf{A}_1]$ for all second-order tensors $\mathbf{A}_1, \mathbf{A}_2$. We now use the path $\tilde{\mathbf{C}}(\xi,\mathbf{C})$ defined by

$$\tilde{\mathbf{C}}(\xi,\mathbf{C}) = \frac{\xi-\xi_0}{\xi_1-\xi_0}(\mathbf{C}-\mathbf{C}_0) + \mathbf{C}_0, \tag{II.17}$$

so $\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{C}}(\xi,\mathbf{C})[\mathbf{H}] = \frac{\xi-\xi_0}{\xi_1-\xi_0}\mathbf{H}$. If $\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}(\mathbf{C})$ has major symmetry for all positive-definite, symmetric second-order tensors $\mathbf{C}$, then $\mathrm{asym}\,\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}(\mathbf{C}) = 0$, so (II.16) with the path (II.17) reduces to

$$(\tilde{\mathbf{S}}(\mathbf{C}) - \tilde{\mathbf{S}}(\mathbf{C}_0)) : \mathbf{H} = 2\frac{\mathrm{d}\widehat{W}}{\mathrm{d}\mathbf{C}}(\mathbf{C}) : \mathbf{H}, \tag{II.18}$$

which holds for all $\mathbf{H}$. Therefore,

$$\tilde{\mathbf{S}}(\mathbf{C}) = 2\frac{\mathrm{d}}{\mathrm{d}\mathbf{C}}\left(\widehat{W}(\mathbf{C}) + \frac{1}{2}\tilde{\mathbf{S}}(\mathbf{C}_0) : \mathbf{C}\right), \tag{II.19}$$

proving that the elastic constitutive relation defined by $\tilde{\mathbf{S}}$ is hyperelastic, because it takes the form (II.9).

II.3 Representation theorem for isotropic tensor functions in two dimensions

Here we prove that a function $\Gamma : \mathcal{A} \to \mathrm{L}(\mathcal{V}_2)$ defined on $\mathcal{A} \subseteq \mathrm{Sym}(2)$ is isotropic, i.e. it satisfies (7.11), if and only if it has the form

$$\Gamma(\mathbf{A}) = \beta_0(\mathcal{J}(\mathbf{A}))\mathbf{I} + \beta_1(\mathcal{J}(\mathbf{A}))\mathbf{A} + \beta_\mathrm{s}(\mathcal{J}(\mathbf{A}))\,\mathrm{sym}\,(\mathbf{WA}) + \beta_\mathrm{a}(\mathcal{J}(\mathbf{A}))\mathbf{W}, \tag{II.20}$$

where the scalar-valued functions $\beta_0, \beta_1, \beta_\mathrm{s}, \beta_\mathrm{a}$ depend only on the invariants $\mathcal{J}(\mathbf{A})$ of $\mathbf{A}$ given in (7.16), and $\mathbf{W}$ is the tensor representing the anticlockwise rotation by $\frac{1}{2}\pi$.

To show that any tensor function of the form (II.20) is isotropic, we first show that the function Υ defined by $\Upsilon(\mathbf{A}) = \mathrm{sym}\,(\mathbf{WA})$ is isotropic. Because all rotations in two-dimensional space commute, for any $\mathbf{Q} \in \mathrm{SO}(2)$ and $\mathbf{A} \in \mathcal{A} \subseteq \mathrm{Sym}(2)$,

$$\begin{aligned}\Upsilon(\mathbf{QAQ}^\mathrm{T}) &= \frac{1}{2}(\mathbf{WQAQ}^\mathrm{T} + \mathbf{QAQ}^\mathrm{T}\mathbf{W}^\mathrm{T}) = \frac{1}{2}(\mathbf{QWAQ}^\mathrm{T} + \mathbf{QAW}^\mathrm{T}\mathbf{Q}^\mathrm{T}) \\ &= \mathbf{Q}\Upsilon(\mathbf{A})\mathbf{Q}^\mathrm{T},\end{aligned}$$

so Υ is an isotropic tensor function. For all $\mathbf{Q} \in \mathrm{SO}(2)$, $\mathcal{J}(\mathbf{QAQ}^\mathrm{T}) = \mathcal{J}(\mathbf{A})$, so

$$\begin{aligned}\Gamma(\mathbf{QAQ}^\mathrm{T}) &= \beta_0(\mathcal{J}(\mathbf{A}))\mathbf{I} + \beta_1(\mathcal{J}(\mathbf{A}))\mathbf{QAQ}^\mathrm{T} + \beta_\mathrm{s}(\mathcal{J}(\mathbf{A}))\,\Upsilon(\mathbf{QAQ}^\mathrm{T}) + \beta_\mathrm{a}(\mathcal{J}(\mathbf{A}))\mathbf{W} \\ &= \mathbf{Q}\Gamma(\mathbf{A})\mathbf{Q}^\mathrm{T},\end{aligned}$$

which completes the proof in this direction.

To prove that any isotropic tensor function defined on $\mathcal{A} \subseteq \mathrm{Sym}(2)$ must take the form (II.20), we consider the symmetric and antisymmetric components of $\Gamma(\mathbf{A})$ separately. Since all antisymmetric tensors in two dimensions are scalar multiples of $\mathbf{W}$, $\mathrm{asym}\,(\Gamma(\mathbf{A})) = \tilde{\beta}_\mathrm{a}(\mathbf{A})\mathbf{W}$ for some scalar function $\tilde{\beta}_\mathrm{a} : \mathcal{A} \to \mathbb{R}$ satisfying

$$\tilde{\beta}_\mathrm{a}(\mathbf{A}) = \frac{\mathbf{W} : \Gamma(\mathbf{A})}{\mathbf{W} : \mathbf{W}} = \frac{1}{2}\mathbf{W} : \Gamma(\mathbf{A}). \tag{II.21}$$

For all $\mathbf{Q} \in \mathrm{SO}(2)$, by the isotropy of Γ,

$$\begin{aligned}\tilde{\beta}_{\mathrm{a}}(\mathbf{Q}\mathbf{A}\mathbf{Q}^{\mathrm{T}}) &= \frac{1}{2}\mathbf{W} : \Gamma(\mathbf{Q}\mathbf{A}\mathbf{Q}^{\mathrm{T}}) = \frac{1}{2}\mathbf{W} : (\mathbf{Q}\Gamma(\mathbf{A})\mathbf{Q}^{\mathrm{T}}) \\ &= \frac{1}{2}\mathrm{tr}(\mathbf{W}^{\mathrm{T}}\mathbf{Q}\Gamma(\mathbf{A})\mathbf{Q}^{\mathrm{T}}) = \frac{1}{2}\mathrm{tr}(\mathbf{W}^{\mathrm{T}}\Gamma(\mathbf{A})) = \tilde{\beta}_{\mathrm{a}}(\mathbf{A}),\end{aligned}$$

showing that $\tilde{\beta}_{\mathrm{a}}$ is an isotropic scalar-valued function. We adapt the proof in Section 37 of Gurtin [6] to show that $\tilde{\beta}_{\mathrm{a}}(\mathbf{A}) = \beta_{\mathrm{a}}(\mathcal{J}(\mathbf{A}))$, i.e. $\tilde{\beta}_{\mathrm{a}}(\mathbf{A})$ depends on $\mathbf{A}$ only through the invariants $\mathcal{J}(\mathbf{A})$. Suppose $\mathbf{A}$ and $\mathbf{A}'$ have the same invariants: $\mathcal{J}(\mathbf{A}) = \mathcal{J}(\mathbf{A}')$. This implies that they have the same eigenvalues μ_1, μ_2, so

$$\mathbf{A} = \mu_1\mathbf{m}_1 \otimes \mathbf{m}_1 + \mu_2\mathbf{m}_2 \otimes \mathbf{m}_2, \quad \mathbf{A}' = \mu_1\mathbf{m}_1' \otimes \mathbf{m}_1' + \mu_2\mathbf{m}_2' \otimes \mathbf{m}_2', \tag{II.22}$$

where $\{\mathbf{m}_1, \mathbf{m}_2\}$ and $\{\mathbf{m}_1', \mathbf{m}_2'\}$ are orthonormal sets of eigenvectors for $\mathbf{A}$ and $\mathbf{A}'$, respectively. Replacing $\mathbf{m}_2'$ with $-\mathbf{m}_2'$ if necessary, there exists a rotation tensor $\mathbf{Q} \in \mathrm{SO}(2)$ such that $\mathbf{m}_j' = \mathbf{Q}\mathbf{m}_j$ for $j = 1, 2$. Thus, $\mathbf{A}' = \mathbf{Q}\mathbf{A}\mathbf{Q}^{\mathrm{T}}$, which implies that $\tilde{\beta}_{\mathrm{a}}(\mathbf{A}') = \tilde{\beta}_{\mathrm{a}}(\mathbf{A})$, by the isotropy of $\tilde{\beta}_{\mathrm{a}}$. Because $\tilde{\beta}_{\mathrm{a}}$ is constant on subsets of $\mathcal{A}$ whose elements $\mathbf{A}$ have the same invariants $\mathcal{J}(\mathbf{A})$, we have shown the result $\tilde{\beta}_{\mathrm{a}}(\mathbf{A}) = \beta_{\mathrm{a}}(\mathcal{J}(\mathbf{A}))$.

Let $\Gamma_s(\mathbf{A}) = \mathrm{sym}\,\Gamma(\mathbf{A})$. Then, $\Gamma_s : \mathcal{A} \to \mathrm{L}(\mathcal{V}_2)$ is a $\mathrm{Sym}(2)$-valued isotropic tensor function. We show that $\Gamma_s(\mathbf{A}) = \beta_0(\mathcal{J}(\mathbf{A}))\mathbf{I} + \beta_1(\mathcal{J}(\mathbf{A}))\mathbf{A} + \beta_{\mathrm{s}}(\mathcal{J}(\mathbf{A}))\,\mathrm{sym}\,(\mathbf{W}\mathbf{A})$. Consider first the special case where $\mathbf{A}$ is a multiple of the identity tensor, i.e. $\mathbf{A} = \lambda\mathbf{I}$, where $\lambda \in \mathbb{R}$. Then for the isotropy condition (7.11) to be satisfied, $\Gamma_s(\lambda\mathbf{I}) = \mathbf{Q}\Gamma_s(\lambda\mathbf{I})\mathbf{Q}^{\mathrm{T}}$ for all $\mathbf{Q} \in \mathrm{SO}(2)$. Since $\Gamma_s(\lambda\mathbf{I})$ is symmetric, by the spectral theorem, it can be written as $\Gamma_s(\lambda\mathbf{I}) = \sigma_1\mathbf{n}_1 \otimes \mathbf{n}_1 + \sigma_2\mathbf{n}_2 \otimes \mathbf{n}_2$, where each term on the right-hand side is a function of λ and the eigenvectors $\mathbf{n}_j$ are chosen to be orthogonal unit vectors. Thus, $\mathbf{n}_2 = \pm\mathbf{W}\mathbf{n}_1$, where the sign depends on how the eigenvectors are chosen. Then the isotropy condition (7.11) with $\mathbf{Q} = \mathbf{W}$ leads to

$$\sigma_1\mathbf{n}_1 \otimes \mathbf{n}_1 + \sigma_2\mathbf{n}_2 \otimes \mathbf{n}_2 = \sigma_2\mathbf{n}_1 \otimes \mathbf{n}_1 + \sigma_1\mathbf{n}_2 \otimes \mathbf{n}_2,$$

which implies $\sigma_1 = \sigma_2$ because $\mathbf{n}_1 \otimes \mathbf{n}_1$ and $\mathbf{n}_2 \otimes \mathbf{n}_2$ are linearly independent. Therefore, $\Gamma_s(\lambda\mathbf{I}) = \hat{\beta}(\lambda)\mathbf{I}$ is a multiple of the identity, and $\lambda = \frac{1}{2}J_1(\mathbf{A})$ when $\mathbf{A} = \lambda\mathbf{I}$. We have therefore shown that the form of Γ_s is a special case of the claimed functional form when $\mathbf{A}$ is a multiple of the identity.

Consider now the generic case where $\mathbf{A}$ is not a multiple of the identity. Then, the set $\{\mathbf{I}, \mathbf{A}\}$ spans a two-dimensional subspace of $\mathrm{Sym}(2)$. We show that that $\mathrm{sym}\,(\mathbf{W}\mathbf{A})$ is non-zero and orthogonal to $\mathbf{I}$ and $\mathbf{A}$, so $\{\mathbf{I}, \mathbf{A}, \mathrm{sym}\,(\mathbf{W}\mathbf{A})\}$ is a basis for $\mathrm{Sym}(2)$, which is three-

dimensional. Because $\mathbf{A}$ and $\mathbf{A}^2$ are symmetric, they are orthogonal to any antisymmetric tensor, so

$$\begin{aligned}\mathbf{I} : \operatorname{sym}(\mathbf{WA}) &= \operatorname{tr}(\mathbf{WA}) = \mathbf{W} : \mathbf{A} = 0,\\ \mathbf{A} : \operatorname{sym}(\mathbf{WA}) &= \mathbf{A} : (\mathbf{WA}) = \operatorname{tr}(\mathbf{A}^2\mathbf{W}) = \mathbf{A}^2 : \mathbf{W} = 0.\end{aligned}$$

To show that $\operatorname{sym}(\mathbf{WA})$ is non-zero, we use the spectral decomposition

$$\mathbf{A} = \mu_1 \mathbf{m}_1 \otimes \mathbf{m}_1 + \mu_2 \mathbf{m}_2 \otimes \mathbf{m}_2$$

to obtain

$$\begin{aligned}||\operatorname{sym}(\mathbf{WA})||^2 &= \operatorname{sym}(\mathbf{WA}) : \operatorname{sym}(\mathbf{WA}) = \operatorname{sym}(\mathbf{WA}) : (\mathbf{WA})\\ &= \frac{1}{2}(\mathbf{WA} - \mathbf{AW}) : (\mathbf{WA}) = \frac{1}{2}\operatorname{tr}(\mathbf{AW}^{\mathrm{T}}(\mathbf{WA} - \mathbf{AW})) = \frac{1}{2}\operatorname{tr}(\mathbf{A}^2 - \mathbf{AW}^{\mathrm{T}}\mathbf{AW})\\ &= \frac{1}{2}\left(\mu_1^2 + \mu_2^2 - \operatorname{tr}((\mu_1 \mathbf{m}_1 \otimes \mathbf{m}_1 + \mu_2 \mathbf{m}_2 \otimes \mathbf{m}_2)(\mu_1 \mathbf{m}_2 \otimes \mathbf{m}_2 + \mu_2 \mathbf{m}_1 \otimes \mathbf{m}_1)\right)\\ &= \frac{1}{2}\left(\mu_1^2 + \mu_2^2 - 2\mu_1\mu_2\right) = \frac{1}{2}(\mu_1 - \mu_2)^2.\end{aligned}$$

Therefore, $\operatorname{sym}(\mathbf{WA}) \neq \mathbf{0}$ because $\mathbf{A}$ is not a multiple of the identity. Since $\{\mathbf{I}, \mathbf{A}, \operatorname{sym}(\mathbf{WA})\}$ is a basis for $\mathrm{Sym}(2)$, there exist scalar functions $\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\beta}_{\mathrm{s}}$ such that

$$\Gamma_s(\mathbf{A}) = \tilde{\beta}_0(\mathbf{A})\mathbf{I} + \tilde{\beta}_1(\mathbf{A})\mathbf{A} + \tilde{\beta}_{\mathrm{s}}(\mathbf{A})\operatorname{sym}(\mathbf{WA}). \tag{II.23}$$

We complete the proof by showing that $\tilde{\beta}_0, \tilde{\beta}_1, \tilde{\beta}_{\mathrm{s}}$ are isotropic scalar functions, i.e. they each satisfy $\tilde{\beta}_{(.)}(\mathbf{QAQ}^{\mathrm{T}}) = \tilde{\beta}_{(.)}(\mathbf{A})$ for all $\mathbf{Q} \in \mathrm{SO}(2)$. Taking the inner product of (II.23) with $\mathbf{I}$ and $\mathbf{A}$ gives

$$\mathbf{I} : \Gamma_s(\mathbf{A}) = 2\tilde{\beta}_0(\mathbf{A}) + (\operatorname{tr}\mathbf{A})\tilde{\beta}_1(\mathbf{A}), \qquad \mathbf{A} : \Gamma_s(\mathbf{A}) = \tilde{\beta}_0(\mathbf{A})(\operatorname{tr}\mathbf{A}) + \tilde{\beta}_1(\mathbf{A})\operatorname{tr}(\mathbf{A}^2),$$

which implies that

$$\tilde{\beta}_1(\mathbf{A}) = \frac{(\mathbf{A} - \frac{1}{2}(\operatorname{tr}\mathbf{A})\mathbf{I}) : \Gamma_s(\mathbf{A})}{\operatorname{tr}(\mathbf{A}^2) - \frac{1}{2}(\operatorname{tr}\mathbf{A})^2} = \frac{(\mathbf{A} - \frac{1}{2}(\operatorname{tr}\mathbf{A})\mathbf{I}) : \Gamma_s(\mathbf{A})}{||\mathbf{A} - \frac{1}{2}(\operatorname{tr}\mathbf{A})\mathbf{I}||^2},$$

which is well-defined whenever $\mathbf{A}$ is not a scalar multiple of $\mathbf{I}$. Because Γ_s is an isotropic tensor function, $\tilde{\beta}_1$ is a isotropic scalar function, which also implies that $\tilde{\beta}_0$ is isotropic.

Taking the inner product of (II.23) with sym $(\mathbf{WA})$ and rearranging gives

$$\tilde{\beta}_{\mathrm{s}}(\mathbf{A}) = \frac{\operatorname{sym}(\mathbf{WA}) : \Gamma_s(\mathbf{A})}{||\operatorname{sym}(\mathbf{WA})||^2},$$

which shows that $\tilde{\beta}_{\mathrm{s}}$ is an isotropic scalar function. To complete the proof, we apply the same reasoning used above in showing that $\tilde{\beta}_{\mathrm{a}}(\mathbf{A}) = \beta_{\mathrm{a}}(\mathcal{J}(\mathbf{A}))$ to prove that $\tilde{\beta}_0(\mathbf{A}) = \beta_0(\mathcal{J}(\mathbf{A}))$, $\tilde{\beta}_1(\mathbf{A}) = \beta_1(\mathcal{J}(\mathbf{A}))$, and $\tilde{\beta}_{\mathrm{s}}(\mathbf{A}) = \beta_{\mathrm{s}}(\mathcal{J}(\mathbf{A}))$.

II.4 Antisymmetric part of the elasticity tensor in the weakly nonlinear theory

To compute $\operatorname{asym}\mathsf{C}_{\mathbf{F}_1}$, we define $\mathbf{S}_o(\mathbf{E}) = \operatorname{tr}(\mathbf{E}^2)\mathbf{I} - 2(\operatorname{tr}\mathbf{E})\mathbf{E}$, whose derivative is

$$\begin{aligned} \mathrm{D}_{\mathbf{E}}\mathbf{S}_o(\mathbf{E})[\mathbf{H}] &= 2(\mathbf{E}:\mathbf{H})\mathbf{I} - 2(\mathbf{I}:\mathbf{H})\mathbf{E} - 2(\operatorname{tr}\mathbf{E})\mathbf{H} \\ &= 2(\mathbf{I}\otimes\mathbf{E} - \mathbf{E}\otimes\mathbf{I})[\mathbf{H}] - 2(\operatorname{tr}\mathbf{E})\mathbf{H}, \end{aligned} \tag{II.24}$$

where the tensor product $\otimes$ between two second-order tensors $\mathbf{A}, \mathbf{A}'$ is defined to be the fourth-order tensor $\mathbf{A}\otimes\mathbf{A}'$ that satisfies $(\mathbf{A}\otimes\mathbf{A}')[\mathbf{H}] = (\mathbf{A}':\mathbf{H})\mathbf{A}$ for all $\mathbf{H}\in\mathtt{L}(\mathcal{V}_d)$. Then,

$$\operatorname{asym}\mathrm{D}_{\mathbf{E}}\mathbf{S}_o(\mathbf{E}) = 2(\mathbf{I}\otimes\mathbf{E} - \mathbf{E}\otimes\mathbf{I}) = 4\operatorname{asym}(\mathbf{I}\otimes\mathbf{E}). \tag{II.25}$$

By the chain rule,

$$\operatorname{asym}\mathrm{D}_{\mathbf{C}}\tilde{\mathbf{S}}(\mathbf{I}+2\mathbf{E}) = \frac{1}{2}d_o\operatorname{asym}\mathrm{D}_{\mathbf{E}}\mathbf{S}_o(\mathbf{E}) = 2d_o\operatorname{asym}(\mathbf{I}\otimes\mathbf{E}). \tag{II.26}$$

Using (7.6) and setting $\mathbf{E} = \frac{1}{2}(\mathbf{F}_1^{\mathrm{T}}\mathbf{F}_1 - \mathbf{I})$ in (II.26), the antisymmetric component of the elasticity tensor $\mathsf{C}_{\mathbf{F}_1}$ satisfies

$$\begin{aligned} \operatorname{asym}\mathsf{C}_{\mathbf{F}_1}[\mathbf{H}] &= \frac{2}{\det\mathbf{F}_1}\mathbf{F}_1\left(d_o\operatorname{asym}(\mathbf{I}\otimes(\mathbf{F}_1^{\mathrm{T}}\mathbf{F}_1 - \mathbf{I}))\right)[\mathbf{F}_1{}^{\mathrm{T}}(\operatorname{sym}\mathbf{H})\mathbf{F}_1]\mathbf{F}_1{}^{\mathrm{T}} \\ &= \frac{d_o}{\det\mathbf{F}_1}\mathbf{F}_1(\mathbf{I}\otimes(\mathbf{F}_1^{\mathrm{T}}\mathbf{F}_1) - (\mathbf{F}_1^{\mathrm{T}}\mathbf{F}_1)\otimes\mathbf{I})[\mathbf{F}_1{}^{\mathrm{T}}(\operatorname{sym}\mathbf{H})\mathbf{F}_1]\mathbf{F}_1{}^{\mathrm{T}} \\ &= \frac{d_o}{\det\mathbf{F}_1}(\mathbf{B}_1\otimes\mathbf{B}_1^2 - \mathbf{B}_1^2\otimes\mathbf{B}_1)[\mathbf{H}], \end{aligned} \tag{II.27}$$

where $\mathbf{B}_1 = \mathbf{F}_1\mathbf{F}_1^{\mathrm{T}}$ is the left Cauchy–Green deformation tensor for the deformed reference state.

II.5 Derivatives of shear and normal stress for an isotropic hyperelastic material

An isotropic hyperelastic material has strain energy function W that satisfies $W(\mathbf{F}) = W(\mathbf{FQ})$ for all $\mathbf{Q} \in \mathrm{SO}(d)$, so by setting $\mathbf{Q} = \mathbf{R}^{\mathrm{T}}$, where $\mathbf{R}$ is the rotation tensor from the polar decomposition $\mathbf{F} = \mathbf{VR}$, we obtain $W(\mathbf{F}) = W(\mathbf{V})$. Defining $\tilde{W}(\mathbf{B}) = W(\sqrt{\mathbf{B}})$, we see that $W(\mathbf{F}) = \tilde{W}(\mathbf{B})$, where $\mathbf{B} = \mathbf{FF}^{\mathrm{T}}$. By the chain rule, the Cauchy stress for this hyperelastic material is

$$\tilde{\mathbf{T}}(\mathbf{B}) = \frac{2}{\sqrt{\det \mathbf{B}}} \frac{\mathrm{d}\tilde{W}}{\mathrm{d}\mathbf{B}}(\mathbf{B})\,\mathbf{B}. \tag{II.28}$$

For the deformation given by (7.27), the left Cauchy–Green deformation tensor has matrix representation

$$[\mathbf{B}] = \begin{bmatrix} 1+\gamma^2 & \gamma\lambda & 0 \\ \gamma\lambda & \lambda^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{II.29}$$

and $\det \mathbf{F} = \lambda$. By the definition given in Appendix I.1,

$$\left(\frac{\mathrm{d}\tilde{W}}{\mathrm{d}\mathbf{B}}(\mathbf{B})\right)_{ij} = \frac{\mathrm{d}\tilde{W}}{\mathrm{d}B_{ij}}(\mathbf{B}). \tag{II.30}$$

From (II.28), the shear stress is

$$T_{12} = \frac{2}{\lambda}\frac{\mathrm{d}\tilde{W}}{\mathrm{d}B_{1k}}B_{k2} = 2\left(\gamma\frac{\mathrm{d}\tilde{W}}{\mathrm{d}B_{11}} + \lambda\frac{\mathrm{d}\tilde{W}}{\mathrm{d}B_{12}}\right) \tag{II.31}$$

where the summation convention has been used and the explicit functional dependence on $\mathbf{B}$ has not been shown, for brevity. Differentiating with respect to λ gives

$$\begin{aligned}
\frac{1}{2}\frac{\partial T_{12}}{\partial \lambda} &= \gamma\frac{\partial^2 \tilde{W}}{\partial B_{ij}\partial B_{11}}\frac{\partial B_{ij}}{\partial \lambda} + \frac{\partial \tilde{W}}{\partial B_{12}} + \lambda\frac{\partial^2 \tilde{W}}{\partial B_{ij}\partial B_{12}}\frac{\partial B_{ij}}{\partial \lambda} \\
&= \frac{\partial \tilde{W}}{\partial B_{12}} + \left(\gamma\frac{\partial^2 \tilde{W}}{\partial B_{ij}\partial B_{11}} + \lambda\frac{\partial^2 \tilde{W}}{\partial B_{ij}\partial B_{12}}\right)\frac{\partial B_{ij}}{\partial \lambda} \\
&= \frac{\partial \tilde{W}}{\partial B_{12}} + \left(\gamma\frac{\partial^2 \tilde{W}}{\partial B_{12}\partial B_{11}} + \lambda\frac{\partial^2 \tilde{W}}{\partial B_{12}\partial B_{12}}\right)\gamma \\
&\quad + \left(\gamma\frac{\partial^2 \tilde{W}}{\partial B_{21}\partial B_{11}} + \lambda\frac{\partial^2 \tilde{W}}{\partial B_{21}\partial B_{12}}\right)\gamma + \left(\gamma\frac{\partial^2 \tilde{W}}{\partial B_{22}\partial B_{11}} + \lambda\frac{\partial^2 \tilde{W}}{\partial B_{22}\partial B_{12}}\right)2\lambda.
\end{aligned} \tag{II.32}$$

The normal stress in the shearing plane is

$$T_{22} = 2\left(\gamma\frac{\mathrm{d}\tilde{W}}{\mathrm{d}B_{21}} + \lambda\frac{\mathrm{d}\tilde{W}}{\mathrm{d}B_{22}}\right), \tag{II.33}$$

and differentiating with respect to γ gives

$$\begin{aligned}\frac{1}{2}\frac{\partial T_{22}}{\partial\gamma} &= \frac{\partial\tilde{W}}{\partial B_{21}} + \left(\gamma\frac{\partial^2\tilde{W}}{\partial B_{ij}\partial B_{21}} + \lambda\frac{\partial^2\tilde{W}}{\partial B_{ij}\partial B_{22}}\right)\frac{\partial B_{ij}}{\partial\gamma}\\ &= \frac{\partial\tilde{W}}{\partial B_{21}} + \left(\gamma\frac{\partial^2\tilde{W}}{\partial B_{11}\partial B_{21}} + \lambda\frac{\partial^2\tilde{W}}{\partial B_{11}\partial B_{22}}\right)2\gamma\\ &\quad + \left(\gamma\frac{\partial^2\tilde{W}}{\partial B_{12}\partial B_{21}} + \lambda\frac{\partial^2\tilde{W}}{\partial B_{12}\partial B_{22}}\right)\lambda + \left(\gamma\frac{\partial^2\tilde{W}}{\partial B_{21}\partial B_{21}} + \lambda\frac{\partial^2\tilde{W}}{\partial B_{21}\partial B_{22}}\right)\lambda.\end{aligned} \tag{II.34}$$

By the principle of material frame-indifference, $\tilde{W}(\mathbf{Q}\mathbf{B}\mathbf{Q}^{\mathrm{T}}) = \tilde{W}(\mathbf{B})$ for all $\mathbf{Q} \in \mathrm{SO}(d)$. Because $\tilde{W}$ is an isotropic scalar function, it depends on $\mathbf{B}$ only via the invariants of $\mathbf{B}$ [6]. Therefore, the value of its derivative $\frac{\mathrm{d}\tilde{W}}{\mathrm{d}\mathbf{B}}(\mathbf{B})$ is a tensor with the same eigenvectors as $\mathbf{B}$, i.e. it is a linear combination of $\mathbf{v}_j \otimes \mathbf{v}_j$ for $j = 1,\ldots,d$, where $\mathbf{v}_j$ are the eigenvectors of $\mathbf{B}$. In particular, $\frac{\mathrm{d}\tilde{W}}{\mathrm{d}\mathbf{B}}(\mathbf{B})$ is a symmetric tensor, so (II.30) implies that

$$\frac{\mathrm{d}\tilde{W}}{\mathrm{d}B_{ij}}(\mathbf{B}) = \frac{\mathrm{d}\tilde{W}}{\mathrm{d}B_{ji}}(\mathbf{B}) \tag{II.35}$$

for all $i, j \in \{1,\ldots,d\}$. Using this symmetry, we show that $\frac{\partial T_{22}}{\partial\gamma} - \frac{\partial T_{12}}{\partial\lambda} = 0$ for isotropic, hyperelastic materials:

$$\begin{aligned}\frac{1}{2}\left(\frac{\partial T_{22}}{\partial\gamma} - \frac{\partial T_{12}}{\partial\lambda}\right) &= \left(\frac{\partial\tilde{W}}{\partial B_{21}} - \frac{\partial\tilde{W}}{\partial B_{12}}\right) + \gamma^2\left(\frac{\partial^2\tilde{W}}{\partial B_{11}\partial B_{21}} - \frac{\partial^2\tilde{W}}{\partial B_{12}\partial B_{11}}\right)\\ &\quad + \gamma\lambda\left(\frac{\partial^2\tilde{W}}{\partial B_{21}\partial B_{21}} - \frac{\partial^2\tilde{W}}{\partial B_{12}\partial B_{12}}\right) + \lambda^2\left(\frac{\partial^2\tilde{W}}{\partial B_{12}\partial B_{22}} - \frac{\partial^2\tilde{W}}{\partial B_{21}\partial B_{22}}\right)\\ &= \left(\frac{\partial\tilde{W}}{\partial B_{21}} - \frac{\partial\tilde{W}}{\partial B_{12}}\right) + \gamma^2\frac{\partial}{\partial B_{11}}\left(\frac{\partial\tilde{W}}{\partial B_{21}} - \frac{\partial\tilde{W}}{\partial B_{12}}\right)\\ &\quad + \gamma\lambda\frac{\partial}{\partial B_{21}}\left(\frac{\partial\tilde{W}}{\partial B_{21}} - \frac{\partial\tilde{W}}{\partial B_{12}}\right) + \lambda^2\frac{\partial}{\partial B_{22}}\left(\frac{\partial\tilde{W}}{\partial B_{12}} - \frac{\partial\tilde{W}}{\partial B_{21}}\right) = 0.\end{aligned} \tag{II.36}$$

II.6 Integrals of tensors over the unit sphere

Here we prove useful identities to evaluate integrals over the unit sphere. We use the notation $\mathbb{B}^d = \{\mathbf{r} \in \mathcal{V}_d : |\mathbf{r}| \leq 1\}$ to refer to the d-dimensional unit ball, with volume

$$\mathtt{V}(d) = \begin{cases} \pi & d = 2 \\ \frac{4}{3}\pi & d = 3 \end{cases}. \tag{II.37}$$

II.6.1 Second-order tensors

We prove that

$$\int_{\mathcal{S}^{d-1}} \mathbf{x} \otimes \mathbf{x} \, \mathrm{d}A(\mathbf{x}) = \mathtt{V}(d)\mathbf{I}, \qquad \int_{\mathbb{B}^d} \mathbf{x} \otimes \mathbf{x} \, \mathrm{d}\mathbf{x} = \frac{\mathtt{V}(d)}{d+2}\mathbf{I}. \tag{II.38}$$

Let $\mathbf{J}_{\mathcal{S}^{d-1}} = \int_{\mathcal{S}^{d-1}} \mathbf{x} \otimes \mathbf{x} \, \mathrm{d}A(\mathbf{x})$. Then, for any fixed $\mathbf{a} \in \mathcal{V}_d$, because $\hat{\mathbf{n}}(\mathbf{x}) = \mathbf{x}$ is the unit normal to $\mathcal{S}^{d-1}$ at $\mathbf{x}$, the divergence theorem [6] can be applied to give

$$\begin{aligned} \mathbf{J}_{\mathcal{S}^{d-1}}\mathbf{a} &= \int_{\mathcal{S}^{d-1}} (\mathbf{x} \cdot \mathbf{a}) \, \mathbf{x} \, \mathrm{d}A(\mathbf{x}) = \int_{\mathcal{S}^{d-1}} (\hat{\mathbf{n}}(\mathbf{x}) \cdot \mathbf{a}) \, \mathbf{x} \, \mathrm{d}A(\mathbf{x}) = \int_{\mathcal{S}^{d-1}} (\mathbf{x} \otimes \mathbf{a}) \, \hat{\mathbf{n}}(\mathbf{x}) \, \mathrm{d}A(\mathbf{x}) \\ &= \int_{\mathbb{B}^d} \nabla \cdot (\mathbf{x} \otimes \mathbf{a}) \, \mathrm{d}\mathbf{x} = \int_{\mathbb{B}^d} ((\nabla \cdot \mathbf{a})\mathbf{x} + (\nabla \mathbf{x})\mathbf{a}) \, \mathrm{d}\mathbf{x} = \int_{\mathbb{B}^d} (\mathbf{0} + \mathbf{I}\mathbf{a}) \, \mathrm{d}\mathbf{x} = \mathtt{V}(d)\mathbf{I}\mathbf{a}, \end{aligned}$$

implying that $\mathbf{J}_{\mathcal{S}^{d-1}} = \mathtt{V}(d)\mathbf{I}$.

Let $\mathbf{J}_{\mathbb{B}^d} = \int_{\mathbb{B}^d} \mathbf{x} \otimes \mathbf{x} \, \mathrm{d}\mathbf{x}$. Then, letting $\mathcal{S}^{d-1}(x) = \{\mathbf{x} \in \mathcal{V}_d : |\mathbf{x}| = x\}$ be the unit sphere of radius x,

$$\begin{aligned} \mathbf{J}_{\mathbb{B}^d} &= \int_0^1 \int_{\mathcal{S}^{d-1}(x)} \mathbf{x} \otimes \mathbf{x} \, \mathrm{d}A(\mathbf{x}) \, \mathrm{d}x = \int_0^1 \int_{\mathcal{S}^{d-1}} \mathbf{x} \otimes \mathbf{x} \, x^{d-1} \, \mathrm{d}A(\hat{\mathbf{x}}) \, \mathrm{d}x \\ &= \int_0^1 x^{d+1} \int_{\mathcal{S}^{d-1}} \hat{\mathbf{x}} \otimes \hat{\mathbf{x}} \, \mathrm{d}A(\hat{\mathbf{x}}) \, \mathrm{d}x = \left(\int_0^1 x^{d+1} \, \mathrm{d}x \right) \mathbf{J}_{\mathcal{S}^{d-1}} = \frac{\mathtt{V}(d)}{d+2}\mathbf{I}, \end{aligned}$$

which proves the result (II.38).

II.6.2 Fourth-order tensors

We show that any $\mathbf{A} \in \mathtt{L}(\mathcal{V}_d)$ is mapped by the fourth order tensor

$$\mathtt{J}_{\mathcal{S}^{d-1}} = \int_{\mathcal{S}^{d-1}} \mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x} \, \mathrm{d}A(\mathbf{x})$$

to

$$\left(\int_{\mathcal{S}^{d-1}} \mathbf{x}\otimes\mathbf{x}\otimes\mathbf{x}\otimes\mathbf{x}\,\mathrm{d}A(\mathbf{x})\right)[\mathbf{A}] = \frac{\mathtt{V}(d)}{d+2}((\mathrm{tr}\,\mathbf{A})\mathbf{I}+2\,\mathrm{sym}\,\mathbf{A}). \tag{II.39}$$

For fixed $\mathbf{A}\in\mathtt{L}(\mathcal{V}_d)$ and $\mathbf{a}\in\mathcal{V}_d$, because $\hat{\mathbf{n}}(\mathbf{x})=\mathbf{x}$ is the unit normal to $\mathcal{S}^{d-1}$ at $\mathbf{x}$, the divergence theorem [6] can be applied to give

$$\begin{aligned}
\mathtt{J}_{\mathcal{S}^{d-1}}[\mathbf{A}]\mathbf{a} &= \int_{\mathcal{S}^{d-1}} (\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(\mathbf{x}\cdot\mathbf{a})\,\mathbf{x}\,\mathrm{d}A(\mathbf{x}) = \int_{\mathcal{S}^{d-1}} (\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(\mathbf{x}\otimes\mathbf{a})\,\hat{\mathbf{n}}(\mathbf{x})\,\mathrm{d}A(\mathbf{x})\\
&= \int_{\mathbb{B}^d} \nabla\cdot((\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(\mathbf{x}\otimes\mathbf{a}))\,\mathrm{d}A(\mathbf{x})\\
&= \int_{\mathbb{B}^d} (\mathbf{A}:\mathbf{x}\otimes\mathbf{x})\,\nabla\cdot(\mathbf{x}\otimes\mathbf{a}) + (\mathbf{x}\otimes\mathbf{a})(\nabla(\mathbf{A}:\mathbf{x}\otimes\mathbf{x}))\,\mathrm{d}A(\mathbf{x})\\
&= \int_{\mathbb{B}^d} (\mathbf{A}:\mathbf{x}\otimes\mathbf{x})\mathbf{a} + (\mathbf{x}\otimes\mathbf{a})(\mathbf{A}+\mathbf{A}^{\mathrm{T}})\mathbf{x}\,\mathrm{d}A(\mathbf{x})\\
&= \int_{\mathbb{B}^d} (\mathbf{A}:\mathbf{x}\otimes\mathbf{x})\mathbf{a} + (\mathbf{x}\otimes\mathbf{x})(\mathbf{A}+\mathbf{A}^{\mathrm{T}})\mathbf{a}\,\mathrm{d}A(\mathbf{x})\\
&= \frac{\mathtt{V}(d)}{d+2}((\mathrm{tr}\,\mathbf{A})\mathbf{I}+2\,\mathrm{sym}\,\mathbf{A})\mathbf{a},
\end{aligned}$$

where in the last equality (II.38) has been used.

The corresponding integral $\mathtt{J}_{\mathbb{B}^d}$ over the unit ball $\mathbb{B}^d$ satisfies, for each $\mathbf{A}\in\mathtt{L}(\mathcal{V}_d)$

$$\begin{aligned}
\left(\int_{\mathbb{B}^d} \mathbf{x}\otimes\mathbf{x}\otimes\mathbf{x}\otimes\mathbf{x}\,\mathrm{d}\mathbf{x}\right)[\mathbf{A}] &= \left(\int_0^1 x^{d-1}x^4\int_{\mathcal{S}^{d-1}} \hat{\mathbf{x}}\otimes\hat{\mathbf{x}}\otimes\hat{\mathbf{x}}\otimes\hat{\mathbf{x}}\,\mathrm{d}A(\hat{\mathbf{x}})\,\mathrm{d}x\right)[\mathbf{A}]\\
&= \frac{\mathtt{V}(d)}{(d+2)(d+4)}((\mathrm{tr}\,\mathbf{A})\mathbf{I}+2\,\mathrm{sym}\,\mathbf{A}).
\end{aligned} \tag{II.40}$$

II.6.3 Sixth-order tensors

We regard sixth-order tensors as bilinear maps from $\mathtt{L}(\mathcal{V}_d)\times\mathtt{L}(\mathcal{V}_d)$ to $\mathtt{L}(\mathcal{V}_d)$. Defining

$$\mathbb{J}_{\mathcal{S}^{d-1}} = \int_{\mathcal{S}^{d-1}} \mathbf{x}\otimes\mathbf{x}\otimes\mathbf{x}\otimes\mathbf{x}\otimes\mathbf{x}\otimes\mathbf{x}\,\mathrm{d}A(\mathbf{x}),$$

for any $\mathbf{A},\tilde{\mathbf{A}}\in\mathtt{L}(\mathcal{V}_d)$ and $\mathbf{a}\in\mathcal{V}_d$,

$$\begin{aligned}
\mathbb{J}_{\mathcal{S}^{d-1}}[\mathbf{A},\tilde{\mathbf{A}}]\mathbf{a} &= \int_{\mathcal{S}^{d-1}} (\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(\tilde{\mathbf{A}}:\mathbf{x}\otimes\mathbf{x})\mathbf{x}\otimes\mathbf{x}\,\mathrm{d}A(\mathbf{x})\,\mathbf{a}\\
&= \int_{\mathcal{S}^{d-1}} (\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(\tilde{\mathbf{A}}:\mathbf{x}\otimes\mathbf{x})(\mathbf{x}\cdot\mathbf{a})\mathbf{x}\,\mathrm{d}A(\mathbf{x})\\
&= \int_{\mathcal{S}^{d-1}} (\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(\tilde{\mathbf{A}}:\mathbf{x}\otimes\mathbf{x})(\mathbf{x}\otimes\mathbf{a})\hat{\mathbf{n}}(\mathbf{x})\,\mathrm{d}A(\mathbf{x}),
\end{aligned} \tag{II.41}$$

where we used $\hat{\mathbf{n}}(\mathbf{x})=\mathbf{x}$ for $\mathbf{x}\in\mathcal{S}^{d-1}$.

To apply the divergence theorem, we first evaluate

$$\begin{aligned}
&\nabla\cdot\big((\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(\tilde{\mathbf{A}}:\mathbf{x}\otimes\mathbf{x})(\mathbf{x}\otimes\mathbf{a})\big)\\
&=(\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(\tilde{\mathbf{A}}:\mathbf{x}\otimes\mathbf{x})\,\nabla\cdot(\mathbf{x}\otimes\mathbf{a})+(\mathbf{x}\otimes\mathbf{a})\,\nabla\big((\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(\tilde{\mathbf{A}}:\mathbf{x}\otimes\mathbf{x})\big)\\
&=(\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(\tilde{\mathbf{A}}:\mathbf{x}\otimes\mathbf{x})\mathbf{a}+(\mathbf{x}\otimes\mathbf{a})((\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(2\operatorname{sym}\tilde{\mathbf{A}})\mathbf{x}+(\tilde{\mathbf{A}}:\mathbf{x}\otimes\mathbf{x})(2\operatorname{sym}\mathbf{A})\mathbf{x})\\
&=(\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(\tilde{\mathbf{A}}:\mathbf{x}\otimes\mathbf{x})\mathbf{a}+((\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(\mathbf{a}\cdot(2\operatorname{sym}\tilde{\mathbf{A}})\mathbf{x})+(\tilde{\mathbf{A}}:\mathbf{x}\otimes\mathbf{x})(\mathbf{a}\cdot(2\operatorname{sym}\mathbf{A})\mathbf{x}))\mathbf{x}\\
&=(\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(\tilde{\mathbf{A}}:\mathbf{x}\otimes\mathbf{x})\mathbf{a}+((\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(\mathbf{x}\cdot(2\operatorname{sym}\tilde{\mathbf{A}})\mathbf{a})+(\tilde{\mathbf{A}}:\mathbf{x}\otimes\mathbf{x})(\mathbf{x}\cdot(2\operatorname{sym}\mathbf{A})\mathbf{a}))\mathbf{x}\\
&=\big((\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(\tilde{\mathbf{A}}:\mathbf{x}\otimes\mathbf{x})\mathbf{I}+(\mathbf{A}:\mathbf{x}\otimes\mathbf{x})(\mathbf{x}\otimes\mathbf{x})(2\operatorname{sym}\tilde{\mathbf{A}})+(\tilde{\mathbf{A}}:\mathbf{x}\otimes\mathbf{x})(\mathbf{x}\otimes\mathbf{x})(2\operatorname{sym}\mathbf{A})\big)\,\mathbf{a}.
\end{aligned}\tag{II.42}$$

Applying the divergence theorem to (II.41) and using (II.42) gives

$$\begin{aligned}
\mathbb{J}_{\mathcal{S}^{d-1}}[\mathbf{A},\tilde{\mathbf{A}}]&=(\tilde{\mathbf{A}}:\mathsf{J}_{\mathbb{B}^d}[\mathbf{A}])\mathbf{I}+2\mathsf{J}_{\mathbb{B}^d}[\mathbf{A}]\operatorname{sym}\tilde{\mathbf{A}}+2\mathsf{J}_{\mathbb{B}^d}[\tilde{\mathbf{A}}]\operatorname{sym}\mathbf{A}\\
&=\frac{\mathtt{V}(d)}{(d+2)(d+4)}\big((\tilde{\mathbf{A}}:((\operatorname{tr}\mathbf{A})\mathbf{I}+2\operatorname{sym}\mathbf{A}))\mathbf{I}\\
&\qquad+2((\operatorname{tr}\mathbf{A})\mathbf{I}+2\operatorname{sym}\mathbf{A})\operatorname{sym}\tilde{\mathbf{A}}+2((\operatorname{tr}\tilde{\mathbf{A}})\mathbf{I}+2\operatorname{sym}\tilde{\mathbf{A}})\operatorname{sym}\mathbf{A}\big)\\
&=\frac{\mathtt{V}(d)}{(d+2)(d+4)}\big(((\operatorname{tr}\mathbf{A})(\operatorname{tr}\tilde{\mathbf{A}})+(2\operatorname{sym}\mathbf{A}):(2\operatorname{sym}\tilde{\mathbf{A}}))\mathbf{I}\\
&\qquad+2((\operatorname{tr}\mathbf{A})\mathbf{I}+2\operatorname{sym}\mathbf{A})\operatorname{sym}\tilde{\mathbf{A}}+2((\operatorname{tr}\tilde{\mathbf{A}})\mathbf{I}+2\operatorname{sym}\tilde{\mathbf{A}})\operatorname{sym}\mathbf{A}\big).
\end{aligned}\tag{II.43}$$

Setting $\tilde{\mathbf{A}}=\mathbf{A}$ gives

$$\begin{aligned}
&\int_{\mathcal{S}^{d-1}}(\mathbf{A}:\mathbf{x}\otimes\mathbf{x})^2\mathbf{x}\otimes\mathbf{x}\,\mathrm{d}A(\mathbf{x})\\
&=\frac{\mathtt{V}(d)}{(d+2)(d+4)}\big(\big((\operatorname{tr}\mathbf{A})^2+2(\operatorname{sym}\mathbf{A}):(\operatorname{sym}\mathbf{A})\big)\,\mathbf{I}+4((\operatorname{tr}\mathbf{A})\mathbf{I}+2\operatorname{sym}\mathbf{A})\operatorname{sym}\mathbf{A}\big).
\end{aligned}\tag{II.44}$$

Part III

Topological mechanics in the continuum

Chapter 10

Introduction to topological mechanics

Metamaterials are structured materials whose properties are determined primarily by the arrangements and interactions of their structural components rather than by their chemical composition [125, 126]. Mechanical metamaterials have been designed to exhibit unusual mechanical responses, such as negative Poisson ratios [127, 128], negative effective masses [129], negative compressibilities [130], programmable shape transformations [131], and non-reciprocity [132], by exploiting geometry [127–132], instabilities [128] and nonlinearities [131, 132].

Topological mechanical metamaterials exhibit properties that are robust to both systematic and disordered perturbations, such as manufacturing defects [12, 133, 134]. These robust properties are characterised by topological invariants, which are integer-valued quantities often defined as winding numbers [12]. The use of topological invariants to characterise physical phenomena has already led to significant breakthroughs in quantum condensed matter physics, such as the description of the quantum hall effect [135] and topological insulators [136]. Efforts to apply these ideas to classical mechanics have given rise to the rich field of *topological mechanics*. Research in this field has used topological invariants to study a broad range of phenomena, from vibrational modes in models of microtubules [137] and zero-energy deformations in mechanical lattices [12], to chiral edge modes in coupled pendula [138] and gyroscopic metamaterials [139].

Here we study the topological mechanics of floppy modes [12–14], which are non-trivial zero-energy deformations that underpin experimentally realisable properties such as asymmetric surface softness in structurally homogeneous mechanical metamaterials [133, 140, 141]. Such metamaterials have potential applications ranging from robust cushioning and vibrational damping [133, 140, 141] to stress focussing [134] and fracture control [142], via the duality between floppy modes and states of self-stress [13]. The pioneering work by Kane and Lubensky [12] characterised topological floppy modes in mechanical lattices

of masses connected by springs, which are useful idealised models for mechanical metamaterials. They linked the presence of floppy modes localised on the boundaries of an isostatic lattice to topological invariants defined using the bulk vibrational bands of the lattice, a form of *bulk–boundary correspondence*. Further work in topological mechanics has not only discovered that isostatic lattices exhibit Weyl modes in the bulk [143, 144], but also shown that topological floppy modes arise in quasicrystals [145], disordered fibre networks [146], kirigami [147], origami [147], models for epithelial tissues [148] and geared metamaterials [149]. This broad range of settings for topological mechanics motivates the development of a continuum theory to capture universal features of topological floppy modes, independent of microscopic detail.

In this part of the thesis, we formulate a continuum theory for topological mechanics by augmenting standard linear elasticity with additional fields, and using the generalised elastic moduli of this theory to define topological invariants for both edge and bulk (Weyl) floppy modes. To put our results in context, we first review topological invariants for floppy modes in isostatic lattices, outlining the formalism for ball-and-spring lattices in Section 10.1 and the definitions of the invariants in Section 10.2. We then review previously proposed continuum theories for topological mechanics in Section 10.3 and outline how our approach differs from previous work. Due to the breadth of concepts introduced in this thesis, we need to re-assign symbols used in Parts I and II. For example, in this thesis part we use $\mathbf{C}$ for the compatibility matrix rather than the right Cauchy–Green deformation tensor. To avoid confusion, we define each symbol when introducing it, so that changes in notation from the previous thesis parts are made clear by the context.

10.1 Ball-and-spring networks

Networks of masses connected by central-force springs, also known as ball-and-spring networks, are useful models for studying the rigidity of mechanical metamaterials [13]. We first introduce general ball-and-spring networks, then focus on lattices consisting of periodic arrangements of masses and springs. Because topological invariants were first introduced to characterise the rigidity of ball-and-spring lattices [12], we give an introduction to the key ideas in the study of these lattices. These ideas motivate the continuum definitions of the Maxwell criterion and the topological invariants in Chapter 11.

We consider a ball-and-spring network in d-dimensional space with N_s sites occupied by masses. These masses are connected by N_b bonds. Let h_1 and h_2 be functions mapping bond index $k \in \{1,\dots,N_\mathrm{b}\}$ to site indices $h \in \{1,\dots,N_\mathrm{s}\}$ such that $h_1(k), h_2(k)$ are the sites connected by the kth bond. Suppose the sites are displaced so that the hth site undergoes a

change in position from $\mathbf{r}_h$ to $\mathbf{r}_h + \mathbf{u}_h$. The extension e_k of the kth bond, with unstretched length $\ell_k = |\mathbf{r}_{h_2(k)} - \mathbf{r}_{h_1(k)}|$, can be expressed in terms of the unit vector $\hat{\mathbf{s}}_k = \ell_k^{-1}(\mathbf{r}_{h_2(k)} - \mathbf{r}_{h_1(k)})$ pointing from site $h_1(k)$ to site $h_2(k)$ and the site displacements $\mathbf{u}_{h_1(k)}, \mathbf{u}_{h_2(k)}$ as follows:

$$\begin{aligned} e_k &= |\mathbf{r}_{h_2(k)} - \mathbf{r}_{h_1(k)} + \mathbf{u}_{h_2(k)} - \mathbf{u}_{h_1(k)}| - |\mathbf{r}_{h_2(k)} - \mathbf{r}_{h_1(k)}| \\ &= \sqrt{\ell_k^2 + 2\ell_k \hat{\mathbf{s}}_k \cdot (\mathbf{u}_{h_2(k)} - \mathbf{u}_{h_1(k)}) + |\mathbf{u}_{h_2(k)} - \mathbf{u}_{h_1(k)}|^2} - \ell_k \\ &= \hat{\mathbf{s}}_k \cdot (\mathbf{u}_{h_2(k)} - \mathbf{u}_{h_1(k)}) + O(|\mathbf{u}|^2) \\ &= \sum_{j=1}^{d} \sum_{h=1}^{N_\mathrm{s}} \hat{s}_{k,j} (\delta_{h,h_2(k)} - \delta_{h,h_2(k)}) u_{h,j} + O(|\mathbf{u}|^2), \end{aligned} \tag{10.1}$$

where $\delta_{h,l}$ is the Kronecker delta and $\hat{s}_{k,j}, u_{h,j}$ are the jth components of $\hat{\mathbf{s}}_k$ and $\mathbf{u}_h$ respectively. To linear order in the site displacements, the relation (10.1) can be expressed compactly as

$$\Xi = \mathbf{C}\Phi, \tag{10.2}$$

where $\Xi = [e_1, \ldots, e_{N_\mathrm{b}}]^\mathrm{T}$, $\Phi = [\mathbf{u}_1, \ldots, \mathbf{u}_{N_\mathrm{s}}]^\mathrm{T}$, and the $N_\mathrm{b} \times N_\mathrm{s} d$ matrix $\mathbf{C}$ with elements given by the coefficients of $u_{h,j}$ in (10.1) is known as the compatibility matrix. In this part of the thesis, we use capital boldface symbols for both matrices and second-order tensors, relying on definitions and context to distinguish them.

Let $\mathbf{f}_h$ be the net force acting on site h as a result of the bond tensions $\{\mathtt{t}_k\}_{k=1}^{N_\mathrm{b}}$. If the site displacements are sufficiently small for their effect on the network geometry to be neglected, then the kth bond can be assumed to be parallel to the unit vector $\hat{\mathbf{s}}_k$, giving the jth component of $\mathbf{f}_h$ as

$$f_{h,j} = \sum_{k=1}^{N_\mathrm{b}} \hat{s}_{k,j} (\delta_{h,h_1(k)} - \delta_{h,h_2(k)}) \mathtt{t}_k, \tag{10.3}$$

where $\mathtt{t}_k > 0$ represents a tensile force. This equation can be expressed compactly as

$$F_\mathrm{net} = -\mathbf{Q}T, \tag{10.4}$$

where $F_\mathrm{net} = [\mathbf{f}_1, \ldots, \mathbf{f}_{N_\mathrm{s}}]^\mathrm{T}$, $T = [\mathtt{t}_1, \ldots, \mathtt{t}_{N_\mathrm{b}}]^\mathrm{T}$, and the $N_\mathrm{s} d \times N_\mathrm{b}$ matrix $\mathbf{Q}$ with elements given by the coefficients of $\mathtt{t}_k$ in (10.4) is known as the equilibrium matrix. Comparing (10.1) and (10.3) shows that the compatibility and equilibrium matrices are related by

$$\mathbf{Q} = \mathbf{C}^\mathrm{T}. \tag{10.5}$$

Kane and Lubensky [12] formulated topological mechanics for periodic ball-and-spring networks, or lattices, formed by repeating unit cells. For each lattice in d-dimensional space,

we can define primitive lattice vectors $\{\mathbf{a}_1,\ldots,\mathbf{a}_d\}$ that generate the Bravais lattice of unit cell positions and span the d-dimensional space. The difference in position between equivalent points in different unit cells is always an integer linear combination of the primitive lattice vectors $m_1\mathbf{a}_1+\ldots+m_d\mathbf{a}_d$, where $m_1,\ldots,m_d$ belong to the set of integers $\mathbb{Z}$. Therefore, once an origin is fixed, unit cells can be indexed by integers $m_1,\ldots,m_d$. We consider a generic lattice whose unit cell contains n_s sites labelled by $h=1,\ldots,n_\mathrm{s}$. These sites are connected by n_b bonds. We first define $\Phi=[\mathbf{u}_1,\ldots,\mathbf{u}_{n_\mathrm{s}}]^\mathrm{T}$ to be the $n_\mathrm{s}d$-dimensional vector containing all of the site displacements for a single unit cell. With this definition, the origin of the unit cell indexed by $(m_1,\ldots,m_d)$ has position

$$\mathbf{x}=m_1\mathbf{a}_1+\ldots+m_d\mathbf{a}_d$$

and the site displacements in this unit cell are specified by

$$\Phi^{(m_1,\ldots,m_d)}=[\mathbf{u}_1^{(m_1,\ldots,m_d)},\ldots,\mathbf{u}_{n_\mathrm{s}}^{(m_1,\ldots,m_d)}]^\mathrm{T}. \tag{10.6}$$

Periodic lattices are conveniently studied in wavevector (reciprocal) space. This space is spanned by the primitive reciprocal lattice vectors $\{\mathbf{b}_1,\ldots,\mathbf{b}_d\}$, which are dual to the primitive lattice vectors $\{\mathbf{a}_1,\ldots,\mathbf{a}_d\}$, and are defined by $\mathbf{b}_j\cdot\mathbf{a}_k=2\pi\delta_{jk}$. We use the primitive reciprocal lattice vectors to express wavevectors,

$$\mathbf{q}=\frac{1}{2\pi}(\bar{q}_1\mathbf{b}_1+\ldots+\bar{q}_d\mathbf{b}_d),$$

where we adopt the convention that $\bar{q}_r$ are dimensionless and $\mathbf{b}_j$ have units of inverse length. The overbar $\bar{\cdot}$ is used to distinguish these dimensionless components $\bar{q}_r$ from the components q_j in $\mathbf{q}=q_1\mathbf{e}_1+\ldots+q_d\mathbf{e}_d$, where $\{\mathbf{e}_j\}_{j=1}^d$ is an orthonormal basis. These components q_j have units of inverse length. We consider (complex-valued) lattice displacements of the form

$$\begin{aligned}\mathbf{u}_h^{(m_1,\ldots,m_d)}(\mathbf{q})&=\tilde{\mathbf{u}}_h\exp(\mathrm{i}\mathbf{q}\cdot(m_1\mathbf{a}_1+\ldots+m_d\mathbf{a}_d))\\&=\tilde{\mathbf{u}}_h\exp(\mathrm{i}(\bar{q}_1m_1+\ldots+\bar{q}_dm_d)),\end{aligned} \tag{10.7}$$

where $\tilde{\mathbf{u}}_h$ is the displacement amplitude and the components $\bar{q}_j$ can be complex. The real parts of $\bar{q}_j$ represent sinusoidal variations across unit cells, and the imaginary parts represent exponentially growing or decaying variations. In this part of the thesis, complex values of physical quantities such as displacements, extensions or forces are used to represent amplitude and phase information conveniently; the actual physical quantities can be recovered by extracting the real or imaginary parts of these complex numbers. Wavevectors differing by reciprocal lattice vectors, i.e. integer linear combinations of primitive reciprocal lattice

vectors, are equivalent because, for any $n_1, \ldots, n_d \in \mathbb{Z}$ and wavevector $\mathbf{q}$,

$$e^{i(\mathbf{q}+n_1\mathbf{b}_1+\cdots+n_d\mathbf{b}_d)\cdot(m_1\mathbf{a}_1+\ldots+m_d\mathbf{a}_d)} = e^{i\mathbf{q}\cdot(m_1\mathbf{a}_1+\ldots+m_d\mathbf{a}_d)}e^{i2\pi(m_1n_1+\ldots+m_dn_d)} = e^{i\mathbf{q}\cdot(m_1\mathbf{a}_1+\ldots+m_d\mathbf{a}_d)}.$$

Therefore, the wavevectors $\mathbf{q}$ and $\mathbf{q}+n_1\mathbf{b}_1+\cdots+n_d\mathbf{b}_d$ describe the same spatial variation across a periodic lattice. Because of this equivalence, we only need to consider wavevectors in the (first) Brillouin zone of the lattice, which is the unit cell of the reciprocal lattice defined by the region closer to $\mathbf{q}=\mathbf{0}$ than any other point in the reciprocal lattice [150]. The topological invariants in Section 10.2 are defined by identifying each point $\mathbf{q}$ in the Brillouin zone with vectors $\mathbf{q}+n_1\mathbf{b}_1+\cdots+n_d\mathbf{b}_d$, for all $n_1, \ldots, n_d \in \mathbb{Z}$, to obtain a quotient space [151] isomorphic to a d-torus. Treating the Brillouin zone as a d-torus means that any path in wavevector space with endpoints $\mathbf{q}$ and $\mathbf{q}+n_1\mathbf{b}_1+\cdots+n_d\mathbf{b}_d$ can be regarded as a closed loop.

Because bonds in periodic lattices can connect sites assigned to different unit cells, we adopt the following convention for assigning bonds to unit cells: the bond k, connecting sites $h_1(k)$ and $h_2(k)$, belongs to the same unit cell as the site $h_1(k)$. For periodic lattices, each bond k not only has a direction defined by the unit vector $\hat{\mathbf{s}}_k$, but also has an associated lattice vector $\mathbf{n}_k$ pointing from the unit cell containing site $h_1(k)$ to that containing site $h_2(k)$. Using similar notation as for a general ball-and-spring network, the extension of the kth bond, where $k = 1, \ldots, n_\mathrm{b}$, in the unit cell indexed by $(m_1, \ldots, m_d)$ is given by

$$e_k^{(m_1,\ldots,m_d)} = \hat{\mathbf{s}}_k \cdot \left(\mathbf{u}_{h_2(k)}^{(m_1',\ldots,m_d')} - \mathbf{u}_{h_1(k)}^{(m_1,\ldots,m_d)}\right) + O(|\mathbf{u}|^2), \tag{10.8}$$

where $(m_1', \ldots, m_d')$ indexes the unit cell for $h_2(k)$, to account for bonds that connect sites in different unit cells. Substituting (10.7) into (10.8) and retaining only the lowest-order terms in displacements gives bond extensions

$$e_k^{(m_1,\ldots,m_d)}(\mathbf{q}) = \tilde{e}_k(\mathbf{q})e^{i\mathbf{q}\cdot(m_1\mathbf{a}_1+\ldots+m_d\mathbf{a}_d)}, \tag{10.9}$$

where the extension amplitudes $\tilde{e}_k(\mathbf{q})$ are given by

$$\tilde{e}_k(\mathbf{q}) = \hat{\mathbf{s}}_k \cdot \left(\tilde{\mathbf{u}}_{h_2(k)}e^{i\mathbf{q}\cdot\mathbf{n}_k} - \tilde{\mathbf{u}}_{h_1(k)}\right), \tag{10.10}$$

with $\mathbf{n}_k = (m_1' - m_1)\mathbf{a}_1 + \ldots + (m_d' - m_d)\mathbf{a}_d$ as the lattice vector associated with the bond k. In wavevector space, the compatibility matrix $\mathbf{C}(\mathbf{q})$ relates the amplitudes of site displacements $\tilde{\Phi} = [\tilde{\mathbf{u}}_1, \ldots, \tilde{\mathbf{u}}_{n_\mathrm{s}}]^\mathrm{T}$ to those of bond extensions $\tilde{\Xi}(\mathbf{q}) = [\tilde{e}_1(\mathbf{q}), \ldots, \tilde{e}_{n_\mathrm{b}}(\mathbf{q})]^\mathrm{T}$ via

$$\tilde{\Xi}(\mathbf{q}) = \mathbf{C}(\mathbf{q})\tilde{\Phi}, \tag{10.11}$$

the kth component of which is given by (10.10).

Forces on sites and bond tensions in periodic lattices can be expressed in wavevector space in an analogous way to displacements and extensions. The net force on the hth site of the unit cell indexed by $(m_1,\dots,m_d)$ is given by

$$\mathbf{f}_h^{(m_1,\dots,m_d)} = \sum_{k=1}^{n_\mathrm{b}} (\mathtt{t}_k^{(m_1,\dots,m_d)} \delta_{h,h_1(k)} - \mathtt{t}_k^{(m_1'',\dots,m_d'')} \delta_{h,h_2(k)}) \hat{\mathbf{s}}_k, \tag{10.12}$$

where $(m_1'',\dots,m_d'')$ indexes the unit cell connected to $(m_1,\dots,m_d)$ by the kth bond when $h = h_2(k)$. When the site h is equal to $h_2(k)$ for some bond k, then the lattice vector $\mathbf{n}_k$ pointing from the unit cell containing $h = h_2(k)$ to that containing $h_1(k)$, is $\mathbf{n}_k = \sum_{j=1}^d (m_j - m_j'')\mathbf{a}_j$. For bond tensions that are given by

$$\mathtt{t}_k^{(m_1,\dots,m_d)} = \tilde{\mathtt{t}}_k \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot(m_1\mathbf{a}_1+\dots+m_d\mathbf{a}_d)}, \tag{10.13}$$

the spatial variation of the forces on sites is

$$\mathbf{f}_h^{(m_1,\dots,m_d)} = \tilde{\mathbf{f}}_h(\mathbf{q}) \mathrm{e}^{\mathrm{i}\mathbf{q}\cdot(m_1\mathbf{a}_1+\dots+m_d\mathbf{a}_d)}, \tag{10.14}$$

where the force amplitudes are given by

$$\tilde{\mathbf{f}}_h(\mathbf{q}) = -\sum_{k=1}^{n_\mathrm{b}} (\delta_{h,h_2(k)} \mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{n}_k} - \delta_{h,h_1(k)}) \tilde{\mathtt{t}}_k \hat{\mathbf{s}}_k. \tag{10.15}$$

This equation can be expressed compactly as

$$\tilde{F}_\mathrm{net}(\mathbf{q}) = -\mathbf{Q}(\mathbf{q})\tilde{T}, \tag{10.16}$$

where $\tilde{F}_\mathrm{net}(\mathbf{q}) = [\tilde{\mathbf{f}}_1(\mathbf{q}),\dots,\tilde{\mathbf{f}}_{n_\mathrm{s}}(\mathbf{q})]^\mathrm{T}$, $\tilde{T} = [\tilde{\mathtt{t}}_1,\dots,\tilde{\mathtt{t}}_{n_\mathrm{b}}]^\mathrm{T}$, and the equilibrium matrix $\mathbf{Q}(\mathbf{q})$ is defined by the coefficients of $\tilde{\mathtt{t}}_k$ in (10.15). Comparing (10.10) and (10.15) shows that the wavevector-dependent compatibility and equilibrium matrices are related by

$$\mathbf{Q}(\mathbf{q})^\mathrm{T} = \mathbf{C}(-\mathbf{q}). \tag{10.17}$$

10.1.1 Zero and floppy modes, states of self-stress, and the Maxwell-Calladine rule

A *zero mode* of a ball-and-spring network refers to site displacements Φ that do not result in any bond extensions Ξ. By (10.1), a zero mode belongs to the null space of the compatibility

matrix $\mathbf{C}$. Rigid body rotations and uniform translations of the whole ball-and-spring network are trivial zero modes. *Floppy modes* are non-trivial zero modes. A *state of self-stress* refers to bond tensions T that do not result in net forces F_{net} acting on the masses, i.e. bond tensions that maintain force equilibrium in the absence of external forces. By (10.4), states of self-stress belong to the null space of the equilibrium matrix $\mathbf{Q}$.

Using (10.5), the Maxwell-Calladine counting rule [152, 153] relating the numbers of independent zero modes N_{ZM}, states of self stress N_{SS}, network degrees of freedom $N_{\text{s}}d$ and constraints N_{b} can be derived. Applying the rank–nullity theorem (e.g. Axler [42]) to the compatibility matrix $\mathbf{C}$ gives

$$\operatorname{rank}\mathbf{C} - N_{\text{ZM}} = N_{\text{s}}d, \tag{10.18}$$

where $\operatorname{rank}\mathbf{C}$ is the dimension of the range space of $\mathbf{C}$. Similarly, the rank–nullity theorem applied to the equilibrium matrix $\mathbf{Q}$ gives

$$\operatorname{rank}\mathbf{Q} - N_{\text{SS}} = N_{\text{b}}. \tag{10.19}$$

By (10.5), $\operatorname{rank}\mathbf{C} = \operatorname{rank}\mathbf{Q}$, giving the Maxwell-Calladine rule

$$N_{\text{ZM}} - N_{\text{SS}} = N_{\text{s}}d - N_{\text{b}}. \tag{10.20}$$

This rule characterises how the balance between the number of degrees of freedom and the number of constraints in a ball-and-spring network determines the rigidity of the network. Networks with $N_{\text{s}}d = N_{\text{b}}$ are on the edge of rigidity and are known as isostatic, or Maxwell, networks.

The aforementioned concepts can be formulated for periodic lattices with spatial variations specified by a (possibly complex) wavevector $\mathbf{q}$, as in (10.7). A zero mode at wavevector $\mathbf{q}$ refers to site displacements given by (10.7) where the displacement amplitudes $\tilde{\Phi}$ are elements of the null space of $\mathbf{C}(\mathbf{q})$. Zero modes at $\mathbf{q} = \mathbf{0}$ are uniform translations of the lattice, so floppy modes correspond to displacements (10.7) with $\mathbf{q} \neq \mathbf{0}$.

For periodic lattices, the Maxwell-Calladine rule is expressed in terms of the numbers of sites and bonds per unit cell and the numbers of independent zero modes $n_{\text{ZM}}(\mathbf{q})$ and states of self-stress $n_{\text{SS}}(\mathbf{q})$ at a given wavevector $\mathbf{q}$. Using similar reasoning as that used to arrive at (10.20),

$$n_{\text{ZM}}(\mathbf{q}) - n_{\text{SS}}(-\mathbf{q}) = n_{\text{s}}d - n_{\text{b}} \tag{10.21}$$

holds for each wavevector $\mathbf{q}$. Lattices satisfying $n_{\text{s}}d = n_{\text{b}}$ are on the edge of rigidity and are isostatic, or Maxwell, lattices. For such lattices, the compatibility matrix $\mathbf{C}(\mathbf{q})$ is square, so

the existence of floppy modes at wavevector $\mathbf{q}$ is equivalent to the condition

$$\det \mathbf{C}(\mathbf{q}) = 0. \tag{10.22}$$

Bulk floppy modes vary sinusoidally across the lattice with no growth or decay in amplitude, corresponding to $\mathbf{q}$ with real components. In contrast, localised floppy modes correspond to $\mathbf{q}$ with at least one complex component. A floppy mode localised on an edge with normal vector $\mathbf{b}$ has wavevector

$$\mathbf{q} = \mathbf{q}_{||} + q_{\perp} \frac{\mathbf{b}}{|\mathbf{b}|}, \tag{10.23}$$

where $\mathbf{q}_{||}$, the surface wavevector, is parallel to the edge and has real components, whereas $q_{\perp}$ is complex and its imaginary part is equal to the inverse decay length of the localised floppy mode. When the equation for floppy mode wavevectors (10.22) satisfies the conditions of the implicit function theorem (e.g. [77]), there exists a function $f_{\perp}$ such that $q_{\perp} = f_{\perp}(\mathbf{q}_{||})$, i.e. $\det \mathbf{C}(\mathbf{q}_{||} + f_{\perp}(\mathbf{q}_{||})\,\mathbf{b}/|\mathbf{b}|) = 0$ for all $\mathbf{q}_{||}$ in some open subset of the set of surface wavevectors. This function $f_{\perp}$ characterises how the floppy mode decay length varies with the surface wavevector $\mathbf{q}_{||}$.

10.2 Topological invariants

Topological invariants for Maxwell lattices are defined as the winding numbers of the function $\mathbf{q} \mapsto \det \mathbf{C}(\mathbf{q})$, which maps the Brillouin zone of the lattice, regarded as a d-torus, to the complex plane $\mathbb{C}$. Closed loops $\mathcal{C}_1$ and $\mathcal{C}_2$ in the Brillouin zone are homotopy equivalent [151] if $\mathcal{C}_1$ can be continuously deformed into $\mathcal{C}_2$, and vice versa. Thus, the set of closed loops in the Brillouin zone can be partitioned into equivalence classes based on homotopy. Each of these homotopy classes can be assigned a winding number: the number of times $\det \mathbf{C}(\mathbf{q})$ winds around the origin of the complex plane as $\mathbf{q}$ is varied along a closed loop belonging to that class. The winding number associated with a homotopy class C can be computed using the integral

$$n_{\mathrm{C}} = \frac{1}{2\pi \mathrm{i}} \int_{\mathcal{C}} \frac{\nabla_{\mathbf{q}} \det \mathbf{C}(\mathbf{q})}{\det \mathbf{C}(\mathbf{q})} \cdot \mathrm{d}\mathbf{q}, \tag{10.24}$$

where the path of integration $\mathcal{C}$ is any closed loop belonging to the class C. These winding numbers are robust to perturbations because they are unchanged by continuous variations in the geometrical parameters defining the lattice, provided $\det \mathbf{C}(\mathbf{q}) \neq 0$ throughout these variations. Because the loops used to calculate these topological invariants involve only $\mathbf{q}$ with real components, the invariants are computed using only information about the bulk displacements of the lattice sites.

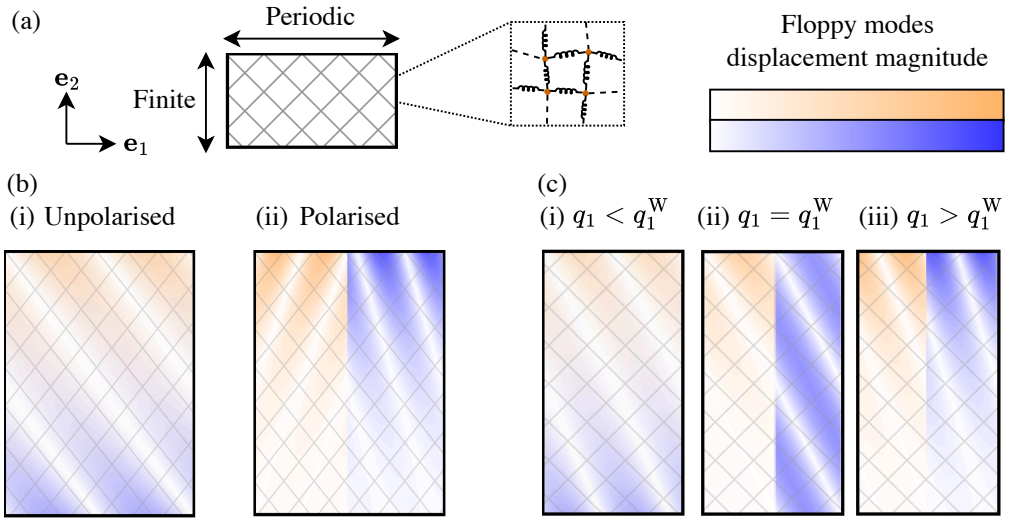


Fig. 10.1 Floppy modes on a strip cut from a two-dimensional lattice. (a) The strip considered is periodic along the $\mathbf{e}_1$-direction and finite along the $\mathbf{e}_2$-direction. The floppy modes are coloured orange and blue, and the associated displacement magnitudes are indicated by the colour intensity. (b) Edge floppy modes can be localised (i) symmetrically between opposite edges, or (ii) asymmetrically, where floppy modes on the same edge are displayed side-by-side for clarity. (c) A Weyl point with component q_1^{W} along the edge parallel to $\mathbf{e}_1$ corresponds to the blue bulk floppy mode shown on (c)(ii). The blue floppy mode, which is localised on (c)(i) the bottom edge for wavevectors with $q_1 < q_1^{\mathrm{W}}$, becomes a bulk Weyl mode at (c)(ii) $q_1 = q_1^{\mathrm{W}}$ and switches to (c)(iii) the top edge, where it is localised for $q_1 > q_1^{\mathrm{W}}$.

In Maxwell lattices, bulk–boundary correspondence relates these bulk topological invariants to floppy modes localised at the boundaries (edges) of a finite lattice, which correspond to solutions $\mathbf{q}$ of (10.22) that have complex components. This correspondence is established using Cauchy's argument principle in the form of the meromorphic counting theorem (e.g. Asmar and Grafakos [154]) applied to the function $\det\mathbf{C}(\mathbf{q})$, which determines the spatial variation of floppy mode displacements via the condition (10.22). A proof of this correspondence is outlined in the Supplemental Information of Sun and Mao [15]. Given an infinite lattice, a boundary can be created by cutting bonds along an edge normal to a primitive reciprocal lattice vector $\mathbf{b}$. If the lattice is gapped, i.e. there are no bulk floppy modes, then $\det\mathbf{C}(\mathbf{q}) \neq 0$ for all real $\mathbf{q} \neq 0$, so the integral in (10.24) is well-defined. Provided the choice of unit cell is appropriate for the edge chosen [14], the number of floppy modes localised on the edge with *inwards* normal $\mathbf{b}$ is equal to the invariant n_{C} in (10.24) with the loop $\mathcal{C}$ given by the path

$$\tilde{\mathbf{q}}(\tau) = \mathbf{q}_{||} + \tau\mathbf{b}, \qquad \tau \in [0,1], \tag{10.25}$$

where $\mathbf{q}_{||}$ is any vector in the Brillouin zone that is not parallel to $\mathbf{b}$. This path defines a closed loop in wavevector space because the points $\mathbf{q}_{||}$ and $\mathbf{q}_{||}+\mathbf{b}$ are equivalent for any reciprocal lattice vector $\mathbf{b}$. Although the path of integration (10.25) depends on $\mathbf{q}_{||}$, the value of the invariant n_{C} is *independent of* $\mathbf{q}_{||}$ because the integral in (10.24) is a continuous function of $\mathbf{q}_{||}$ that takes discrete values, so it must take the same value for all $\mathbf{q}_{||}$ not parallel to $\mathbf{b}$ [155]. We have outlined bulk–boundary correspondence in Maxwell lattices: the bulk topological invariant n_{C} in (10.24) with path of integration (10.25) equals the number of floppy modes localised on the edge with inward normal $\mathbf{b}$.

Importantly, in a strip cut from a Maxwell lattice such that it is finite along the $\mathbf{b}$-direction, localised floppy modes can be symmetrically or asymmetrically distributed between opposite edges, as shown on Fig. 10.1(b). Because the localisation of floppy modes is characterised using a topological invariant, the presence of an asymmetric distribution of edge floppy modes is often referred to as *topological polarisation* [12–14]. To characterise this polarisation, previous work on topological floppy modes [12, 14] defined a vector $\mathbf{r}_{\mathrm{T}}$ in terms of the primitive lattice vectors $\{\mathbf{a}_1,\ldots,\mathbf{a}_d\}$. This vector points towards the edges on which floppy modes are localised and takes the form [14]

$$\mathbf{r}_{\mathrm{T}} = -\sum_{j=1}^{d} n_j \mathbf{a}_j, \tag{10.26}$$

where the integers n_j are special cases of the winding numbers (10.24), with the path $\mathcal{C}$ obtained by setting $\mathbf{b}=\mathbf{b}_j$ in (10.25).

In addition to edge floppy modes, the invariant (10.24) also characterises bulk floppy modes, corresponding to wavevectors $\mathbf{q}\neq\mathbf{0}$ that have purely real components satisfying (10.22). For two-dimensional Maxwell lattices, Rocklin et al. [143] defined the winding number for a bulk floppy mode at an isolated wavevector $\mathbf{q}^{\mathrm{W}}$ by taking the path of integration $\mathcal{C}$ in (10.24) to be a loop enclosing the point $\mathbf{q}^{\mathrm{W}}$. Such floppy mode wavevectors that are associated with non-trivial winding numbers are known as *Weyl points*. An example of a Weyl point is illustrated by Fig. 10.1(c). For a lattice with Weyl points, the number of floppy modes localised on a given edge depends on the surface wavevector $\mathbf{q}_{||}$ (cf. (10.23)), unlike the number of edge modes in a fully gapped Maxwell lattice, i.e. a lattice for which (10.22) has no real solutions other than $\mathbf{q}=\mathbf{0}$. To illustrate this dependence on the surface wavevector, let $\mathbf{e}_{||}$ be the unit vector parallel to the edge of the two-dimensional lattice, so that the surface wavevector is $\mathbf{q}_{||}=q_{||}\mathbf{e}_{||}$. The Weyl winding number is the number of floppy modes that switch from one edge to the opposite edge as $q_{||}$ passes through the projection $\mathbf{q}^{\mathrm{W}}\cdot\mathbf{e}_{||}$ of the Weyl point onto the subspace of wavevectors parallel to the edge, as illustrated in Fig. 10.1(c). For three-dimensional Maxwell lattices, the wavevectors of bulk floppy

modes with non-trivial winding numbers necessarily form lines or curves in wavevector space, known as Weyl lines [144].

10.3 Continuum theories for topological mechanics

Continuum descriptions of physical systems, such as elasticity theory and hydrodynamics, have long been powerful tools for characterising large-scale phenomena that exhibit universal features across systems with significantly different microscopic interactions. The variety of systems exhibiting topological mechanical behaviour, ranging from origami [147] and models of epithelial tissue [148] to disordered fibre networks [146], quasicrystals [145] and ball-and-spring lattices [12], motivates using a continuum theory to capture universal features of topological mechanics that are independent of microscopic detail. Furthermore, such a theory would enable efficient computation of large-scale deformations due to topological floppy modes, facilitating the design of mechanical metamaterials with robust properties.

Recently proposed continuum models [15–17] have captured some of the rich phenomenology of topological mechanics. Because topological polarisation involves an asymmetric distribution of floppy modes between opposite edges, as shown on Fig. 10.1(b)(ii), any continuum theory that captures this phenomenon must break spatial inversion symmetry.

Both Sun and Mao [15] and Saremi and Rocklin [16] propose continuum theories in which the elastic energy depends on strain gradients in addition to the usual strain tensor, to break this symmetry. Sun and Mao [15] consider a generic expansion of elastic energy in terms of strain and its gradients, retaining the first gradient of strain to obtain their continuum theory. Saremi and Rocklin [16] motivate their theory by expressing the extension of a spring in a generic lattice in terms of strain and its gradient, to obtain a continuum compatibility matrix. Both approaches characterise topological polarisation by defining invariants that do not rely on a Brillouin zone. These invariants count the difference between the numbers of floppy modes localised on opposite edges.

Taking a different approach, Nassar et al. [17] consider a weakly-distorted kagome lattice to derive elasticity augmented by one additional field, which they call microtwist elasticity. In this approach, spatial inversion symmetry is broken by the dependence of stress on both the additional field and its gradient. Microtwist elasticity was also adapted to obtain a continuum model for weakly-distorted three-dimensional pyrochlore lattices in Xia et al. [156]. Although microtwist elasticity is able to capture asymmetrically localised floppy modes, it is formulated in terms of specific lattices (kagome and pyrochlore), and the polarisation vector defined by Nassar et al. [17] only applies to kagome lattices.

None of the previously proposed continuum theories capture Weyl modes, which appear generically in both theoretical models [143, 144] and experimental realisations [140] of discrete systems. We therefore aim to capture both topological edge floppy modes and Weyl modes in a continuum theory that is independent of any underlying lattice structure.

Chapter 11

Augmented elasticity for topological mechanics

Standard linear elasticity is governed by equations that are symmetric under spatial inversion for homogenous materials [17]. Consequently, for any finite strip of such a material, the existence of a floppy mode localised on one edge implies the existence of another floppy mode localised on the opposite edge. Thus, standard linear elasticity is unable to capture the asymmetric distribution of edge modes characteristic of topological polarisation. In this chapter, we construct a generalised elasticity theory that breaks spatial inversion symmetry, which we use in Chapter 12 to explore topological phenomena in the continuum. We begin by constructing a strain energy function (elastic energy density) and use the Euler–Lagrange equations to derive equations of motion in Section 11.1. Then, we define a Maxwell criterion and winding numbers in the continuum using the generalised elasticity moduli of our continuum theory. Although we construct this theory without reference to any underlying microscopic structure, we show in Section 11.2 that our theory arises naturally as the continuum limit of a ball-and-spring lattice.

11.1 Generalised elasticity and the Maxwell criterion

Standard linear elasticity is determined by the displacement field $\mathbf{u}(\mathbf{x},t)$ at position $\mathbf{x}$ and time t. We consider n_w additional continuum fields $\{\varphi_m(\mathbf{x},t)\}_{m=1}^{n_w}$, which represent displacements associated with internal degrees of freedom. Although Nassar et al. [17] and Xia et al. [156] previously used specific cases of augmented elasticity with $n_w = 1$ and $n_w = 3$ to model kagome and pyrochlore lattices, respectively, our approach uses symmetry considerations to construct a theory without reference to a microscopic lattice.

Because elasticity arises from the energetic costs of spatial variations in the kinematic fields $\mathbf{u}, \varphi_1, \ldots \varphi_{n_w}$, the strain energy function W, which is the elastic energy per unit volume, depends on $\nabla\mathbf{u}$, $\nabla\varphi_1, \ldots, \nabla\varphi_{n_w}$, where ∇ is the spatial gradient. We refer to these kinematic quantities as the *generalised strain measures*. To consider only variations over the longest length scales, we do not include any higher-order gradients. However, the dependence of the strain energy function on these generalised strain measures is not sufficient to break spatial inversion symmetry because the transformation $\mathbf{x} \mapsto -\mathbf{x}$, $\mathbf{u} \mapsto -\mathbf{u}$, $\varphi_m \mapsto -\varphi_m$ for $m = 1, \ldots, n_w$, which reverses the signs of position vectors and generalised displacements, leaves the gradients $\nabla\mathbf{u}$, $\nabla\varphi_1, \ldots, \nabla\varphi_{n_w}$ unchanged. Therefore, to break spatial inversion symmetry, we need a strain energy function that depends on the fields $\varphi_1, \ldots, \varphi_{n_w}$ alongside their gradients.

To construct this strain energy function, we define a vector of generalised strain measures that includes $\varphi_1, \ldots, \varphi_{n_w}$:

$$\Lambda = \left[\nabla^s\mathbf{u}\ \nabla\varphi_1 \ldots \nabla\varphi_{n_w}\ \varphi_1 \ldots \varphi_{n_w}\right]^{\mathrm{T}}, \tag{11.1}$$

where we have replaced $\nabla\mathbf{u}$ by its symmetrisation $\nabla^s\mathbf{u} = \frac{1}{2}\left(\nabla\mathbf{u} + (\nabla\mathbf{u})^{\mathrm{T}}\right)$, because the isotropy of space requires that the strain energy be invariant under infinitesimal rigid-body rotations of the system, represented by the antisymmetric part of $\nabla\mathbf{u}$. For stress to depend linearly on the generalised strain measures, the strain energy function is the quadratic form

$$W = \frac{1}{2}\Lambda^{\mathrm{T}}\hat{\mathbf{K}}\Lambda, \tag{11.2}$$

where $\hat{\mathbf{K}}$ is a matrix of generalised elastic moduli, which is symmetric and positive semi-definite to ensure stability. By including dependence on $\varphi_1, \ldots, \varphi_{n_w}$, the strain energy function is not symmetric under spatial inversion due to the presence of terms in W with only a single gradient, such as $\varphi_m\nabla\mathbf{u}$ and $\varphi_m\nabla\varphi_k$. Because the additional fields $\varphi_1, \ldots, \varphi_{n_w}$ are treated as displacements, the dependence of W on $\nabla^s\mathbf{u}$ requires that the gradients $\nabla\varphi_1, \ldots, \nabla\varphi_{n_w}$ be included for consistency.

We have introduced a fundamental difference between $\varphi_1, \ldots, \varphi_{n_w}$ and $\mathbf{u}$ as kinematic fields: $\mathbf{u}$ affects W only through its gradient $\nabla\mathbf{u}$ because of the elastic system's uniform translational symmetry, whereas no such symmetries exist for $\varphi_1, \ldots, \varphi_{n_w}$. We show in Section 11.2.4 that when the continuum theory is obtained from homogenising a lattice, the energetic cost of non-zero $\varphi_1, \ldots, \varphi_{n_w}$ is associated with the lattice being gapped at sufficiently small wavenumbers, a necessary condition for the presence of topological polarisation.

To obtain the equations of motion of a continuum system with our strain energy function, we define a kinetic energy

$$E_K = \frac{1}{2}(\partial_t \Psi)^{\mathrm{T}} \hat{\mathbf{M}} \partial_t \Psi, \tag{11.3}$$

where $\Psi = [\mathbf{u}\ \varphi_1, \ldots, \varphi_{n_w}]^{\mathrm{T}}$ are the generalised displacements and $\hat{\mathbf{M}}$ is the generalised mass matrix, which is symmetric and positive-definite. For completeness, we define the total potential energy density of the system to be $W + V_{\mathrm{ext}}$, where $V_{\mathrm{ext}}(\mathbf{u}, \varphi_1, \ldots, \varphi_{n_w})$ accounts for the conservative body force density $\mathbf{f}(\mathbf{x},t)$ and the generalised body torque density $\boldsymbol{\tau}_k(\mathbf{x},t)$. Applying the Euler–Lagrange equations to the Lagrangian density $\mathcal{L} = E_K - (W + V_{\mathrm{ext}})$ results in the equations of motion

$$\rho \partial_t^2 \mathbf{u}(\mathbf{x},t) + \sum_{m=1}^{n_w} \partial_t^2 \varphi_m(\mathbf{x},t) \mathbf{p}_m = \mathbf{f}(\mathbf{x},t) + \nabla \cdot \boldsymbol{\sigma}(\mathbf{x},t), \tag{11.4a}$$

and for $k = 1, \ldots, n_w$,

$$\mathbf{p}_k \cdot \partial_t^2 \mathbf{u}(\mathbf{x},t) + \sum_{m=1}^{n_w} \mu_{km} \partial_t^2 \varphi_m(\mathbf{x},t) = \boldsymbol{\tau}_k(\mathbf{x},t) - \eta_k(\mathbf{x},t) + \nabla \cdot \mathbf{c}_k(\mathbf{x},t), \tag{11.4b}$$

where ρ is the mass density, and $\mathbf{p}_k$, μ_{km} are generalised inertia and moment of inertia densities, respectively. These equations contain the stress measures

$$\boldsymbol{\sigma}(\mathbf{x},t),\ \{\mathbf{c}_k(\mathbf{x},t)\}_{k=1}^{n_w} \text{ and } \{-\eta_k(\mathbf{x},t)\}_{k=1}^{n_w}$$

conjugate to the strain measures

$$\nabla^s \mathbf{u}(\mathbf{x},t),\ \{\nabla \varphi_m(\mathbf{x},t)\}_{m=1}^{n_w} \text{ and } \{\varphi_m(\mathbf{x},t)\}_{m=1}^{n_w},$$

respectively. The form of (11.4a) shows that $\boldsymbol{\sigma}(\mathbf{x},t)$ is the first Piola–Kirchhoff stress from standard elasticity, which in the linear-elastic regime is equal to the Cauchy stress [6]. To express the constitutive relation between these stress and strain measures, we first introduce our notation for tensors in this part of the thesis.

In Section 2.1.1, we introduced second-order tensors as linear maps on the space of vectors. These tensors are elements $\mathbf{M}$ of the linear space spanned by the basis of tensor products $\{\mathbf{e}_i \otimes \mathbf{e}_j\}_{i,j=1}^d$, where $\{\mathbf{e}_j\}_{j=1}^d$ is a basis for the space of d-dimensional vectors and the tensor product $\otimes$ is as defined in Section 2.1.1. Here we define third-order tensors to be elements $\mathfrak{B}$ of the linear space spanned by the basis $\{\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k\}_{i,j,k=1}^d$, and fourth-order tensors to be elements C of the linear space spanned by $\{\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_l\}_{i,j,k,l=1}^d$. Unlike in Part I, where we use $\mathbf{M}\mathbf{v}$ to denote the image of $\mathbf{v}$ under the linear map defined by the second-

order tensor $\mathbf{M}$, in this part, we use $\mathbf{M}\cdot\mathbf{v}$ to make explicit the contraction of indices involved, i.e. $(\mathbf{M}\cdot\mathbf{v})_i = M_{ij}v_j$, where the summation convention has been used. This notation clarifies how third-order tensors act on vectors and second-order tensors: the third-order tensor $\mathfrak{B}$ can be used to map a vector $\mathbf{v}$ to a second order tensor $\mathfrak{B}\cdot\mathbf{v}$ that has (i,j)th component $B_{ijk}v_k$ by contracting one pair of indices, or to map a second-order tensor $\mathbf{M}$ to a vector $\mathfrak{B}:\mathbf{M}$ that has ith component $B_{ijk}M_{jk}$ by contracting two pairs of indices. The symbols $\cdot$ and : indicate single and double contraction, respectively, and their positions indicate which pairs of indices are contracted, e.g. $\mathbf{v}\cdot\mathfrak{B}$ has (i,j)th component v_kB_{kij} and $\mathbf{M}:\mathfrak{B}$ has ith component $M_{jk}B_{jki}$. We use the same notation for the action of fourth-order tensors as linear maps on the space of second-order tensors.

The stress and strain measures are related by the constitutive relations

$$\boldsymbol{\sigma} = \mathsf{C}:\nabla^s\mathbf{u} + \sum_{m=1}^{n_w}\left(\mathfrak{B}_m\cdot\nabla\varphi_m + \mathbf{N}_m\varphi_m\right) \tag{11.5a}$$

$$\mathbf{c}_k = \mathfrak{B}_k^{\mathrm{T}}:\nabla^s\mathbf{u} + \sum_{m=1}^{n_w}\left(\mathbf{M}_{km}\cdot\nabla\varphi_m + \mathbf{h}_{km}\varphi_m\right) \tag{11.5b}$$

$$\eta_k = \mathbf{N}_k:\nabla^s\mathbf{u} + \sum_{m=1}^{n_w}\left(\mathbf{h}_{mk}\cdot\nabla\varphi_m + J_{km}\varphi_m\right), \tag{11.5c}$$

where C is the fourth-order elasticity tensor from standard elasticity, and $\mathfrak{B}_k$, $\mathbf{N}_k$ are third-order and second-order tensors, respectively. The coefficients $\mathbf{h}_{km}$ and J_{km} are vectors and scalars, respectively. The transpose operation on third-order tensors is defined to satisfy $(\mathfrak{B}^{\mathrm{T}})_{ijk} = (\mathfrak{B})_{kij}$, motivated by treating the third-order tensor as a map from the space of vectors to the space of second-order tensors. We refer the coefficients in (11.5) as *generalised elastic moduli*. Defining the generalised stresses $\Sigma = [\boldsymbol{\sigma}\ \mathbf{c}_1\ldots\mathbf{c}_{n_w}\ \eta_1\ldots\eta_{n_w}]^{\mathrm{T}}$, these constitutive relations (11.5) can be expressed compactly as

$$\Sigma = \frac{\mathrm{d}W}{\mathrm{d}\Lambda} = \hat{\mathbf{K}}\Lambda, \tag{11.6}$$

where $\hat{\mathbf{K}}$ is the matrix of generalised elastic moduli from (11.2).

11.1.1 Maxwell criterion and topological floppy modes

Having constructed a generalised elasticity theory that breaks spatial inversion symmetry, we proceed to show how our theory enables the definition of topological invariants. In Chapter 12, we show how these invariants are linked to topological polarisation and Weyl modes in the continuum.

We formulate a Maxwell criterion for the continuum by defining an effective compatibility matrix from our generalised elasticity theory, motivated by the concepts introduced in Chapter 10. To begin, we recall that the generalised displacements $\Psi = [\mathbf{u}\ \varphi_1 \dots \varphi_{n_w}]^{\mathrm{T}}$ represent our $d+n_w$ continuum degrees of freedom. As in the discrete case [13], we consider generalised displacements of the form $\Psi = \tilde{\Psi}\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{x}}$, where the complex components of $\mathbf{q}$ represent spatially growing and decaying modes and the tilde over vectors and scalars indicates that they represent amplitudes. In this representation, the (complex) generalised strain measures in Λ are

$$\begin{aligned} \nabla^s\mathbf{u} &= \operatorname{sym}\left(\left(\tilde{\mathbf{u}}\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{x}}\right)\otimes(\mathrm{i}\mathbf{q})\right) \\ \nabla\varphi_m &= \mathrm{i}\mathbf{q}\,\tilde{\varphi}_m\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{x}} \\ \varphi_m &= \tilde{\varphi}_m\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{x}}, \end{aligned} \tag{11.7}$$

where $\operatorname{sym}(\mathbf{A}) = \frac{1}{2}\left(\mathbf{A}+\mathbf{A}^\dagger\right)$ is the symmetrisation of $\mathbf{A}$, defined using the conjugate transpose (Hermitian transpose) $\mathbf{A}^\dagger = (\mathbf{A}^{\mathrm{T}})^*$, involving the complex conjugate $(\cdot)^*$. By linearity, (11.7) can be written as

$$\Lambda = \mathbf{L}(\mathbf{q})\tilde{\Psi}\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{x}}, \tag{11.8}$$

where $\mathbf{L}(\mathbf{q})$ is a $\mathbf{q}$-dependent linear map from the generalised displacements Ψ to the generalised strains Λ, which can be represented as a matrix with respect to some orthonormal basis $\{\mathbf{e}_j\}_{j=1}^d$. Using (11.6), the generalised stresses are given by $\Sigma = \hat{\mathbf{K}}\mathbf{L}(\mathbf{q})\tilde{\Psi}\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{x}}$. Let $n_K = \operatorname{rank}\hat{\mathbf{K}}$ be the dimension of the range space of $\hat{\mathbf{K}}$, and $\{y_1,\dots,y_{n_K}\}$ be an orthonormal basis for this range space, which is easily obtained from the set of eigenvectors of $\hat{\mathbf{K}}$ corresponding to non-zero eigenvalues. The integer n_K is the continuum version of the number of bonds n_{b} in the discrete case, because it represents the number of constraints imposed by the generalised elastic moduli. We define $\mathbf{P}_K$ to be the matrix with columns $y_1,\dots,y_{n_K}$, so $\mathbf{P}_K\mathbf{P}_K^{\mathrm{T}}$ is the projection operator onto the range space of $\hat{\mathbf{K}}$.

In a discrete lattice, pre-multiplying the compatibility matrix by the diagonal matrix containing bond spring constants results in a linear map from site displacements to bond forces. The $n_K \times (d+n_w)$ matrix $\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q})$ is a continuum analogue to this linear map, because it maps generalised displacements Ψ to the n_K independent components of generalised stress Σ. Recalling that the discrete Maxwell criterion is satisfied for ball-and-spring networks in which the number of constraints equals the number of degrees of freedom, a natural definition for the Maxwell criterion in the continuum is

$$n_K = d+n_w, \tag{11.9}$$

which balances the number of constraints n_K against the number of degrees of freedom $d+n_w$. The matrix $\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q})$ is square when the continuum Maxwell criterion (11.9) is satisfied, so $\det\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q})$ exists in this case. Our criterion (11.9) generalises the condition for Maxwell media in Sun and Mao [15] by accounting for the constraints corresponding to all the generalised stresses (11.5), whereas they consider only the $d(d+1)/2$ constraints corresponding to the usual stress $\boldsymbol{\sigma}$.

To use the continuum Maxwell criterion (11.9) to define topological invariants, we first relate zero modes to the matrix $\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q})$ mapping generalised displacements to stress components. Zero modes at wavevector $\mathbf{q}$ exist if and only if there are generalised strains Λ of the form (11.8) that result in $V=\frac{1}{2}\Lambda^{\mathrm{T}}\hat{\mathbf{K}}\Lambda=0$, i.e. $\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q})$ has a non-trivial null space. Given the Maxwell criterion (11.9), zero modes exist at wavevector $\mathbf{q}$ if and only if $\det\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q})=0$. Comparing this condition with $\det\mathbf{C}(\mathbf{q})=0$ for zero mode wavevectors in the discrete case, we define topological invariants in the continuum by analogy with (10.24), which uses the integrand $(\det\mathbf{C}(\mathbf{q}))^{-1}\nabla_{\mathbf{q}}\det\mathbf{C}(\mathbf{q})$. Because such invariants would depend on $\det\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q})$ only via the ratio $(\det\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q}))^{-1}\nabla_{\mathbf{q}}\det\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q})$, any matrix whose determinant is proportional to $\det\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q})$ could be chosen to define the topological invariant. We exploit this gauge choice to obtain a slightly simpler expression for the invariant by defining an effective continuum compatibility matrix $\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q})$, which we use to define a topological invariant via $\det\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q})$. Because $\hat{\mathbf{K}}$ is symmetric, $\mathbf{P}_K\mathbf{P}_K^{\mathrm{T}}$ projects onto the orthogonal complement of the null space of $\hat{\mathbf{K}}$. Therefore, we have $\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q})=\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{P}_K\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q})$, which implies

$$\det\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q})=\det\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{P}_K\det\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}),$$

where $\det\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{P}_K>0$ is a constant. Thus,

$$\frac{\nabla_{\mathbf{q}}\det\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q})}{\det\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q})}=\frac{\nabla_{\mathbf{q}}\det\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q})}{\det\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q})},$$

so we define our topological invariants via

$$n[\mathbf{q}(\tau)]=\frac{1}{2\pi\mathrm{i}}\int_{\tau_0}^{\tau_1}\frac{\frac{\mathrm{d}}{\mathrm{d}\tau}\det\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}(\tau))}{\det\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}(\tau))}\,\mathrm{d}\tau, \tag{11.10}$$

where $\mathbf{q}(\tau)$ represents a parametrisation of a path in $\mathbf{q}$-space by the parameter $\tau\in[\tau_0,\tau_1]$. If the path is closed, $\mathbf{q}(\tau_0)=\mathbf{q}(\tau_1)$. For an appropriate choice of path $\mathbf{q}(\tau)$, the invariant $n[\mathbf{q}(\tau)]$ counts the number of zero mode wavevectors enclosed by the path, by the meromorphic counting theorem based on Cauchy's argument principle (e.g. Asmar and Grafakos [154]).

When considering topological edge modes in Section 12.1, we modify this form of invariant to account for the absence of a Brillouin zone in the continuum, adapting an approach applied to higher strain gradient theories in Saremi and Rocklin [16]. In this approach, we consider non-closed paths $\mathbf{q}(\tau)$ for which $\mathbf{q}$ is real, to retain the bulk–boundary correspondence seen in discrete topological mechanics: the invariant is computed using bulk modes but contains information about the edge modes.

We have defined an effective compatibility matrix, from which we will compute topological invariants for edge modes and Weyl points in Chapter 12. Importantly, our results are completely independent of microscopic detail and rely only on the continuum model. Although we have used periodic lattices to motivate our definitions, none of our results require the existence of an underlying lattice, only that the generalised elastic moduli satisfy our Maxwell criterion (11.9). Therefore, our approach might also apply to non-periodic topological systems, such as fibre networks [146], quasicrystals [145], and in formulating model-free topological mechanics [157].

11.1.2 Relation to higher strain gradient theories

Here we show that continuum theories involving the gradient of the linearised strain $\nabla(\nabla^s\mathbf{u})$, such as those in Sun and Mao [15] and Saremi and Rocklin [16], can be obtained as special cases of our generalised elasticity theory. Following an approach similar to that in Forest and Sab [158], we constrain the additional fields $\varphi_1,\dots,\varphi_{n_w}$ so that they are linear functions of $\nabla^s\mathbf{u}$. To do this, we set $n_w=\frac{1}{2}d(d+1)$, which is the number of independent components of $\nabla^s\mathbf{u}$, and impose the constraint

$$\varphi_m=\mathbf{D}_m:\nabla^s\mathbf{u}, \tag{11.11}$$

for $m=1,\dots,n_w$, where $\mathbf{D}_m$ is a second-order tensor. Using (11.11), the strain energy function W becomes dependent on the gradient of linearised strain $\nabla(\nabla^s\mathbf{u})$ via the generalised strain measures $\nabla\varphi_m$, which are linear combinations of the components of $\nabla(\nabla^s\mathbf{u})$. Continuum theories involving higher gradients of linearised strain can be similarly recovered by including more additional fields φ_m and expressing these fields in terms of higher strain gradients.

To formulate topological mechanics in the continuum, our approach and previous approaches [15, 16] use theories which have more elastic moduli than standard linear elasticity. A possible advantage of our approach is that our generalised continuum theory with strain energy function (11.2) introduces fewer additional elastic moduli to capture topological phenomena than theories involving strain gradients such as $\nabla(\nabla^s\mathbf{u})$. The number of elastic moduli in a generalised linear theory with n_Λ independent components of generalised

strain is $n_\Lambda(n_\Lambda+1)/2$, because the matrix of elastic moduli defining the mapping from generalised strain to generalised stress is symmetric. Since this number is a monotonically increasing function of n_Λ, to compare the numbers of elastic moduli in linear continuum theories, we can compare the values of n_Λ for each theory. In linear elasticity that depends on both the usual strain $\nabla^s\mathbf{u}$, which has $d(d+1)/2$ independent components, and the first strain gradient $\nabla(\nabla^s\mathbf{u})$, which has $d^2(d+1)/2$ independent components, there are $n_\Lambda = d(d+1)/2 + d^2(d+1)/2 = d(d+1)^2/2$ components of generalised strain. In our generalised elasticity theory, where stress depends on n_w additional fields and their gradients, besides the usual strain tensor $\nabla^s\mathbf{u}$, there are $n_\Lambda = d(d+1)/2 + n_w d + n_w = (d+1)(d/2+n_w)$ components of generalised strain. Comparing these values of n_Λ shows that the number of elastic moduli in our theory is smaller than that in a first strain gradient theory when

$$n_w < \frac{1}{2}d^2. \tag{11.12}$$

Substituting $d=2$ and $d=3$ into (11.12), we see that generalised elasticity theories with $n_w = 1$ have fewer coefficients than lowest-order strain gradient theories. As we show in Section 12.1, topological polarisation can be captured when $n_w = 1$, so our approach is able to capture the same topological phenomenology as the first strain gradient theories in Sun and Mao [15] and Saremi and Rocklin [16] with fewer additional elastic moduli. Additionally, we show in Section 12.2 that Weyl points can be captured using $n_w = 2$, so our approach captures a class of topological phenomena not previously seen in higher strain gradient theories.

11.2 Continuum limit of a ball-and-spring lattice

We have augmented linear elasticity with additional fields to capture topological phenomena, based on the requirement that the theory be able to break spatial inversion symmetry. Here we take the continuum limit of a generic ball-and-spring lattice and show that our augmented continuum theory emerges naturally. This homogenisation procedure shows that the additional fields $\varphi_1,\dots,\varphi_{n_w}$ correspond to local soft modes in the lattice. This procedure also links the generalised elastic moduli to the compatibility matrix of the underlying lattice.

Our continuum theory describes physical phenomena on a macroscopic length scale L much larger than the lattice unit cell size a, i.e. in the limit $\delta = a/L \to 0$. To take this limit, we fix L and consider arbitrarily small a. When referring to the *order* of a quantity in δ, we refer to its asymptotic behaviour as $\delta \to 0$, i.e. the statement that g is of order m in δ, also written as $g(\delta) = O(\delta^m)$, means that $\lim_{\delta\to 0} \frac{g(\delta)}{\delta^m}$ is a constant. This asymptotic order should

not be confused with the order of a tensor, as defined in Section 11.1, and the meaning of this word will be clear from its context.

We require a homogenisation procedure to obtain continuum equations of motion systematically from the lattice equations in the limit $\delta \to 0$. Such a procedure should have the following features.

1. The procedure applies to generic lattices, and captures the lattice geometry through a few parameters, the generalised elastic moduli.

2. The kinematic fields, including the displacement field, correspond to physical displacements of the lattice. In particular, the continuum displacement field $\mathbf{u}(\mathbf{x},t)$ is a physically meaningful function of lattice displacements.

3. Only terms of the lowest order in δ are retained in the equations of motion, because these are the dominant terms in the continuum limit $\delta \to 0$.

11.2.1 Lattice site displacements and the compatibility matrix

We obtain a continuum description $\Phi(\mathbf{x},t)$ of the lattice degrees of freedom $\Phi^{(m_1,\ldots,m_d)}$ in (10.6) by replacing the site displacements $\mathbf{u}_h^{(m_1,\ldots,m_d)}$ with smooth fields $\mathbf{u}_h(\mathbf{x},t)$ that satisfy $\mathbf{u}_h^{(m_1,\ldots,m_d)}(t) = \mathbf{u}_h(\mathbf{x},t)$ when $\mathbf{x} = m_1\mathbf{a}_1 + \ldots + m_d\mathbf{a}_d$. Using these site displacement fields, (10.8) becomes

$$e_k(\mathbf{x},t) = \hat{\mathbf{s}}_k \cdot (\mathbf{u}_{h_2(k)}(\mathbf{x}+\mathbf{n}_k,t) - \mathbf{u}_{h_1(k)}(\mathbf{x},t)) + O(|\mathbf{u}|^2), \tag{11.13}$$

where $\mathbf{n}_k = (m_1' - m_1)\mathbf{a}_1 + \ldots + (m_d' - m_d)\mathbf{a}_d$ is the lattice vector associated with the kth bond, pointing from the unit cell containing the site $h_1(k)$ to that containing site $h_2(k)$, as defined before (10.8). Because the lattice vectors have magnitudes that scale with unit cell size, $|\mathbf{a}_j| \propto a$, where $a = L\delta$, implying that $|\mathbf{n}_k| = O(\delta)$. Retaining the terms in (11.13) of the lowest orders in $\mathbf{u}$ and δ gives

$$\begin{aligned} e_k(\mathbf{x},t) &= \hat{\mathbf{s}}_k \cdot ((1+\mathbf{n}_k\cdot\nabla)\mathbf{u}_{h_2(k)}(\mathbf{x},t) - \mathbf{u}_{h_1(k)}(\mathbf{x},t)) \\ &= \hat{\mathbf{s}}_k \cdot (\mathbf{u}_{h_2(k)}(\mathbf{x},t) - \mathbf{u}_{h_1(k)}(\mathbf{x},t)) + \hat{\mathbf{s}}_k \cdot (\mathbf{n}_k \cdot \nabla \mathbf{u}_{h_2(k)}(\mathbf{x},t)), \end{aligned} \tag{11.14}$$

where ∇ represents differentiation with respect to $\mathbf{x}$. To obtain a continuum version of (10.11), we re-write (11.14) as

$$\Xi(\mathbf{x},t) = (\mathbf{C}_0 + \mathbf{C}_\mathbf{x})\Phi(\mathbf{x},t), \tag{11.15}$$

where $\Xi(\mathbf{x},t) = [e_1(\mathbf{x},t)\dots e_{n_\mathrm{b}}(\mathbf{x},t)]^\mathrm{T}$, $\mathbf{C}_0\Phi(\mathbf{x},t)$ has kth component given by the first term in the second line of (11.14), and $\mathbf{C}_\mathbf{x}\Phi(\mathbf{x},t)$ has kth component given by the second term. The continuum compatibility matrix $\mathbf{C}_0 + \mathbf{C}_\mathbf{x}$ has been decomposed into the matrix $\mathbf{C}_0$ that does not contain differential operators from ∇, and the differential operator $\mathbf{C}_\mathbf{x}$. These terms are related to the wavevector-space compatibility matrix $\mathbf{C}(\mathbf{q})$ introduced in (10.11) by $\mathbf{C}_0 = \mathbf{C}(\mathbf{0})$ and

$$\mathbf{C}_\mathbf{x} = \sum_{j,r=1}^{d} (-\mathrm{i})(\partial_{\bar{q}_r}\mathbf{C}(\mathbf{q})|_{\mathbf{q}=\mathbf{0}})(\mathbf{a}_r\cdot\mathbf{e}_j)\partial_{x_j},$$

as can be shown by comparing (11.14) and (10.10). We show that $\mathbf{C}_\mathbf{x}\Phi = O(\delta)$ by expressing it in terms of non-dimensional variables. We use the non-dimensional spatial variable $\mathbf{x}^* = L^{-1}\mathbf{x}$, obtained by normalising with respect to system size L. By our assumption that L is the length scale over which spatial variations in Φ occur, the derivatives $\partial_{x_j^*}\Phi = O(1)$. Therefore,

$$\mathbf{C}_\mathbf{x}\Phi = \sum_{j,r=1}^{d} (-\mathrm{i})(\partial_{\bar{q}_r}\mathbf{C}(\mathbf{q})|_{\mathbf{q}=\mathbf{0}})(L^{-1}\mathbf{a}_r\cdot\mathbf{e}_j)\partial_{x_j^*}\Phi$$

is $O(\delta)$, because $|L^{-1}\mathbf{a}_r| = O(\delta)$, and all the other factors are $O(1)$. This decomposition of the compatibility matrix shows how spatial gradients enter the continuum equations of motion for $\Phi(\mathbf{x},t)$.

11.2.2 Perturbing the compatibility matrix

Although topologically polarised lattices are not spatial-inversion symmetric, this broken symmetry might not be visible on large length scales. We demonstrate that perturbing a lattice configuration about a gapless state ensures that the homogenised theory preserves the symmetry breaking terms. Thus, we generalise the geometrical perturbations employed in references [17, 156] from specific lattices to a generic lattice, and justify the form of the perturbation.

The compatibility matrix $\mathbf{C}(\mathbf{q})$ of a generic lattice depends on the wavevector components $\bar{q}_r$ $(r = 1,\dots,d)$ via complex phases of the form $\mathrm{e}^{\mathrm{i}\bar{q}_r}$, as shown in (10.10). Therefore, a generic form for the determinant is

$$\det\mathbf{C}(\mathbf{q}) = \sum_\ell c_\ell \exp(\mathrm{i}\mathbf{n}_\ell\cdot\mathbf{q}) = \sum_\ell c_\ell \sum_{m=0}^{\infty} \frac{(\mathrm{i}\mathbf{n}_\ell\cdot\mathbf{q})^m}{m!} = \sum_{m=0}^{\infty} \frac{\mathrm{i}^m}{m!} P_m(\mathbf{q}), \tag{11.16}$$

where the sum over ℓ is a finite sum over the permutations of the matrix elements and we have defined $P_m(\mathbf{q}) = \sum_\ell c_\ell(\mathbf{n}_\ell\cdot\mathbf{q})^m$, which is a homogeneous polynomial of degree m in the components $\bar{q}_r$, with real coefficients. We obtain the first equality in (11.16) using the

expression for the determinant of a matrix in terms of permutations of its elements, and the second equality using the Taylor series for the exponential function. In our notation, $\bar{q}_r^{-1}$ is the characteristic number of unit cells over which spatial variations occur. For these variations to be visible in the continuum, $\bar{q}_r^{-1}$ should scale as L/a, which is the size of the system expressed in terms of numbers of unit cells, so $\bar{q}_r = O(\delta)$. This observation tells us which terms in (11.16) are of lowest order in δ and are therefore retained in the continuum theory.

The continuum limit reproduces topological mechanics only if the coefficients c_ℓ depend on δ. Suppose instead that the coefficients c_ℓ are independent of δ. Then, (11.16) gives

$$\det \mathbf{C}(\mathbf{q}) = \frac{\mathrm{i}^{m_0}}{m_0!} P_{m_0}(\mathbf{q}) + O(\delta^{m_0+1}),$$

where m_0 is the smallest m for which P_m is not identically zero. Therefore, in the continuum limit, the terms lowest order in δ come from just one P_m, implying that the zero modes in this limit correspond to the solutions of $P_{m_0}(\mathbf{q}) = 0$, a polynomial with real coefficients. Consider a strip of a Maxwell lattice with edges normal to the dth primitive reciprocal lattice vector. Spatial variations along the strip edges are specified by real values of $\bar{q}_1, \ldots, \bar{q}_{d-1}$, corresponding to a wavevector in the surface Brillouin zone of the strip. The zero mode corresponding to a given surface wavevector $\mathbf{q}_{||}$ has spatial variation normal to the edges given by $\bar{q}_d$, which satisfies a polynomial equation with real coefficients. Therefore, the zero modes of this strip have $\bar{q}_d$ values that are either real or occur in complex conjugate pairs. The former case corresponds to bulk modes, and the latter to edge modes symmetrically distributed between opposite edges. Thus, topological polarisation, if present in the discrete lattice, will not be visible in the continuum limit unless the coefficients c_ℓ depend on δ.

To capture topological polarisation at large length scales, the polynomial equation satisfied by $\bar{q}_d$ must have complex coefficients, to lowest order in δ. To achieve this, we begin with a gapless lattice, i.e. one for which bulk floppy modes exist at $\mathbf{q} = \mathbf{0}$, and perturb the positions of the sites to arrive at a gapped configuration, in which the only zero modes at $\mathbf{q} = \mathbf{0}$ are the uniform translations. This perturbation turns the floppy modes of the gapless lattice into *soft modes*, which have low but non-zero energy. The unperturbed gapless lattice has wavevector space compatibility matrix $\mathbf{C}(\mathbf{q})$ with n_w floppy modes at $\mathbf{q} = \mathbf{0}$, so $\mathbf{C}_0 = \mathbf{C}(\mathbf{0})$ has a $(d + n_w)$-dimensional null space. Under the geometrical perturbation, the compatibility matrix becomes

$$\mathbf{C}_p(\mathbf{q}; \varepsilon) = \mathbf{C}(\mathbf{q}) + \varepsilon \mathbf{C}_w(\mathbf{q}) + O(\varepsilon^2), \tag{11.17}$$

where ε parametrises the perturbation to site positions. Letting $\mathbf{C}_w = \mathbf{C}_w(\mathbf{0})$, the perturbations to site positions are chosen so that $\mathbf{C}_0 + \varepsilon \mathbf{C}_w$ has null space of dimension d, consisting of only

the translational zero modes. The n_w soft modes of the perturbed lattice at $\mathbf{q} = \mathbf{0}$ correspond to bond extensions of magnitude $O(\varepsilon)$. By the same reasoning used in Section 11.2.1 to obtain (11.15), we deduce that the continuum counterpart to $\mathbf{C}_p(\mathbf{q};\varepsilon)$ is

$$\mathbf{C}_{p,\mathbf{x}} = \mathbf{C}_0 + \mathbf{C}_\mathbf{x} + \varepsilon(\mathbf{C}_w + \mathbf{C}_{\mathbf{x},w}) + O(\varepsilon^2), \tag{11.18}$$

where

$$\mathbf{C}_{\mathbf{x},w} = \sum_{j,r=1}^{d} (-\mathrm{i})(\partial_{\bar{q}_r}\mathbf{C}_w(\mathbf{0}))(\mathbf{a}_r \cdot \mathbf{e}_j)\partial_{x_j},$$

following our decomposition of the unperturbed compatibility matrix into $\mathbf{C} + \mathbf{C}_\mathbf{x}$ in (11.15). As in Section 11.2.1, $\mathbf{C}_{\mathbf{x},w}\Phi = O(\delta)$.

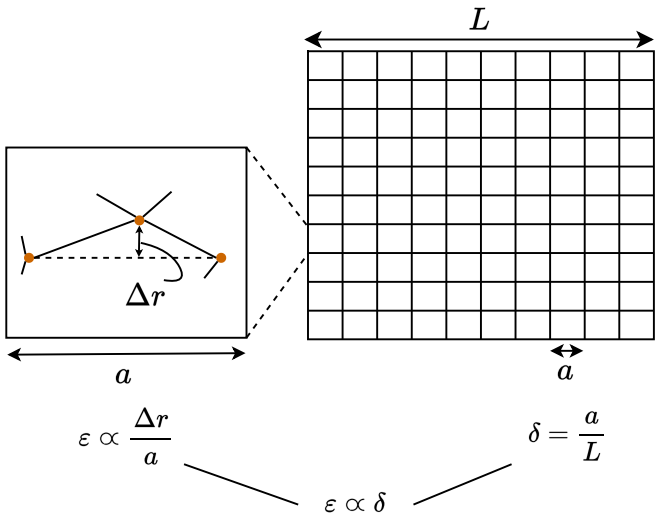


Fig. 11.1 Physical interpretation of the scaling relationship $\varepsilon = O(\delta)$. This choice of scaling relationship links two ratios. The grid on the right represents the many unit cells present in a system with macroscopic length scale L. The box on the left represents a single unit cell and the perturbation $\Delta r \propto \varepsilon a$ applied to the site positions in order to obtain a gapped lattice from a gapless configuration. The relationship $\varepsilon = O(\delta)$ connects a length scale associated with the geometry of a single unit cell to the characteristic number of unit cells over which spatial variations occur in the continuum. In the continuum theory, this scaling relationship ensures that the terms breaking spatial inversion symmetry are relevant in the limit $\delta \to 0$.

To show how ε-perturbing a lattice leads to symmetry-breaking terms being retained in the continuum limit, we need to specify how ε scales with δ. Suppose $\varepsilon = O(\delta)$, which is illustrated in Fig. 11.1. Physically, this choice of scaling for ε means that the geometric perturbation to the site positions *within* a unit cell relative to unit cell size a has magnitude

inversely proportional to the number of unit cells L/a over which spatial variations need to occur to be visible at the continuum level.

In a lattice ε-perturbed away from a configuration with n_w floppy modes, the coefficients $c_\ell(\varepsilon)$ of $\bar{q}_r$ in the expansion of $\det \mathbf{C}_p(\mathbf{q};\varepsilon)$ are polynomials in ε. In general, as shown rigorously in Appendix III.1,

$$\det \mathbf{C}_p(\mathbf{q};\varepsilon) = \sum_{m=d}^{d+n_w} \frac{\mathrm{i}^m}{m!} \varepsilon^{d+n_w-m} \bar{P}_m(\bar{q}_1,\ldots,\bar{q}_d) + O(\delta^{d+n_w+1}), \tag{11.19}$$

where $\bar{P}_m(\bar{q}_1,\ldots,\bar{q}_d) = \sum_{k_1+\ldots+k_d=m} \bar{A}_{k_1,\ldots,k_d} \bar{q}_1^{k_1} \ldots \bar{q}_d^{k_d}$ for real coefficients $\bar{A}_{k_1,\ldots,k_d}$. In the expansion of $\det \mathbf{C}_p(\mathbf{q};\varepsilon)$, the terms that are of order d in dimensionless wavevector components $\bar{q}_r = O(\delta)$ (the lowest possible order of $\bar{q}_r$) contain a factor of ε^{n_w}, raising their overall order in δ to $d+n_w$, by $\varepsilon = O(\delta)$. However, these are not the only terms of this order in δ, because terms of order m in $\bar{q}_r$ have order $d+n_w-m$ in ε, so these terms also have order $d+n_w$ in δ, the lowest order present in the expansion of $\det \mathbf{C}_p(\mathbf{q};\varepsilon)$. These terms are shown in the sum over $m = d,\ldots,d+n_w$ in (11.19). Given real values of $\bar{q}_1,\ldots,\bar{q}_{d-1}$, solving for zero modes using the terms lowest order in δ is equivalent to obtaining the roots of a polynomial in $\bar{q}_d$ with complex coefficients. Therefore, an asymmetric distribution of floppy modes across opposite edges is possible, because complex roots are not constrained to occur in complex conjugate pairs. This motivates the study of ε-perturbed lattices to observe topological phenomena in the continuum limit.

We have shown that choosing $\varepsilon = O(\delta)$ is sufficient to capture asymmetric edge mode localisation at small wavenumbers $\bar{q}_r$, when the polarisation is present in the discrete lattice. To show that $\varepsilon = O(\delta)$ is necessary for modelling topological mechanics in the continuum, we write $\varepsilon = O(\delta^\nu)$, where $\nu \geq 0$ because we require ε to remain bounded in the continuum limit $\delta \to 0$. The possibility that ε is independent of δ is accounted for by the case $\nu = 0$. We show in Appendix III.2.3 that $\varepsilon = O(\delta)$ is the only choice that captures topological polarisation because $\nu \neq 1$ results in continuum equations of motion that are spatial inversion symmetric.

11.2.3 The continuum displacement field and other kinematic fields

To relate the continuum fields $\mathbf{u}(\mathbf{x},t)$, $\varphi_1(\mathbf{x},t),\ldots,\varphi_{n_w}(\mathbf{x},t)$ to the lattice site displacements contained in $\Phi(\mathbf{x},t)$, we first define $\mathbf{u}(\mathbf{x},t)$ to be the average lattice site displacement:

$$\mathbf{u}(\mathbf{x},t) = \frac{1}{n_\mathrm{s}} \sum_{h=1}^{n_\mathrm{s}} \mathbf{u}_h(\mathbf{x},t). \tag{11.20}$$

The components of $\mathbf{u}(\mathbf{x},t)$ are proportional to the projections of $\Phi(\mathbf{x},t)$ along the vectors representing uniform translations $\{U_j = [\mathbf{e}_j \dots \mathbf{e}_j]^{\mathrm{T}}\}_{j=1}^{d}$ of the whole unit cell, which are eigenvectors of the dynamical matrix evaluated at $\mathbf{q} = \mathbf{0}$, $\varepsilon = 0$, i.e. $\mathbf{C}_0^{\mathrm{T}}\mathbf{K}\mathbf{C}_0$. Therefore, a natural definition for the other continuum degrees of freedom involves considering the other eigenvectors of the symmetric matrix $\mathbf{C}_0^{\mathrm{T}}\mathbf{K}\mathbf{C}_0$, which form an orthogonal basis for the space of lattice site displacements. This orthogonality means that the eigenvectors can be treated as representing independent lattice degrees of freedom.

The eigenvectors of $\mathbf{C}_0^{\mathrm{T}}\mathbf{K}\mathbf{C}_0$ can be categorised into those corresponding to zero eigenvalues, which include the uniform translations $\{U_j\}_{j=1}^{d}$ and the n_w (non-translational) floppy modes $\{W_m\}_{m=1}^{n_w}$, and those corresponding to non-zero eigenvalues, which we refer to as the high-frequency modes $\{V_k\}_{k=1}^{n_s d-(d+n_w)}$. The sets of eigenvectors $\{U_j\}_{j=1}^{d}$, $\{W_m\}_{m=1}^{n_w}$,and $\{V_k\}_{k=1}^{n_s d-(d+n_w)}$ each span a subspace with corresponding orthogonal projection operator $\mathbf{P}_u$, $\mathbf{P}_w$ and $\mathbf{P}_v$, with $\mathbf{P}_u+\mathbf{P}_w+\mathbf{P}_v$ equal to the identity matrix $\mathbf{I}$. Therefore, we can decompose the continuum degrees of freedom into the subspaces spanned by these three sets of eigenvectors:

$$\Phi(\mathbf{x},t) = \Phi_u(\mathbf{x},t) + \Phi_w(\mathbf{x},t) + \Phi_v(\mathbf{x},t), \tag{11.21}$$

where $\Phi_u(\mathbf{x},t) = \mathbf{P}_u\Phi(\mathbf{x},t)$, $\Phi_w(\mathbf{x},t) = \mathbf{P}_w\Phi(\mathbf{x},t)$ and $\Phi_v(\mathbf{x},t) = \mathbf{P}_v\Phi(\mathbf{x},t)$. We show in Section 11.2.4 that the degrees of freedom associated with the high-frequency modes do not appear explicitly in the continuum equations of motion, because they can be eliminated in the continuum limit. Therefore, the remaining degrees of freedom are associated with the floppy modes $\{W_m\}_{m=1}^{n_w}$, so we define

$$\varphi_m(\mathbf{x},t) = W_m^{\mathrm{T}}\Phi(\mathbf{x},t) = W_m^{\mathrm{T}}\Phi_w(\mathbf{x},t), \tag{11.22}$$

for $m = 1,\dots,n_w$, which are the projections of $\Phi(\mathbf{x},t)$ along the vectors representing floppy modes. Although the vectors $\{U_j\}_{j=1}^{d}$ and $\{W_m\}_{m=1}^{n_w}$ are all in the null space of $\mathbf{C}$, for the ε-perturbed lattice, only $\{U_j\}_{j=1}^{d}$ are in the null space of $\mathbf{C}_0+\varepsilon\mathbf{C}_w$, whereas the vectors $\{W_m\}_{m=1}^{n_w}$ are soft modes with low but non-zero energy. We show in Section 11.2.4 that the projections in (11.22) give rise to the additional fields in the generalised continuum theory presented in Section 11.1. Our definitions (11.20) and (11.22) satisfy Item 2 listed in Section 11.2, because they are physically meaningful definitions of continuum fields in terms of lattice site displacements.

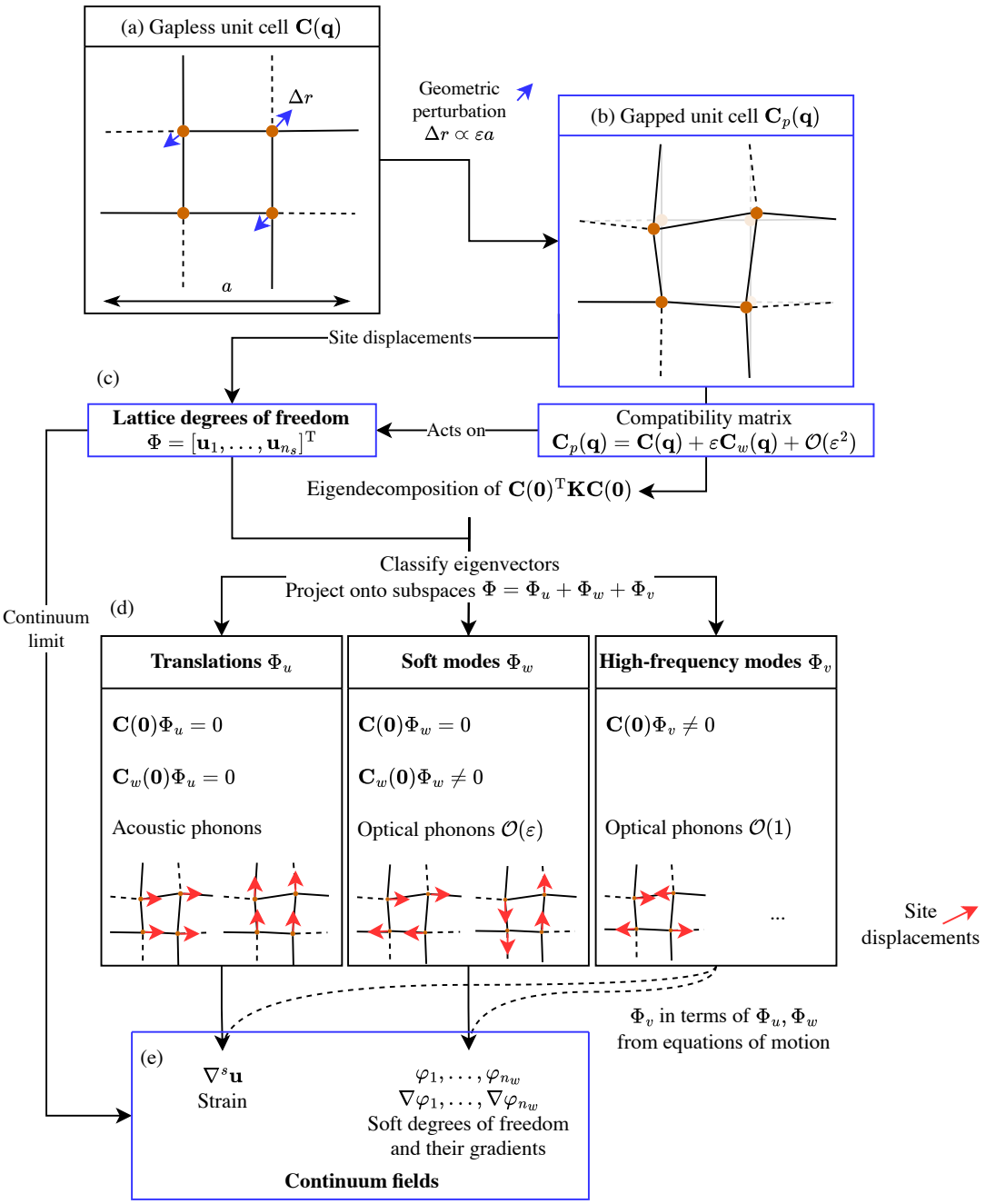


Fig. 11.2 Obtaining continuum fields from a discrete lattice. (a) A gapless lattice is geometrically perturbed, shifting the site positions by distances proportional to εa, where a is the unit cell size. (b) The resulting gapped lattice has a compatibility matrix given by (11.17). (c) The lattice degrees of freedom are the $n_s d$ site displacements, which are projected onto three subspaces, as shown in (11.21). (d) The subspaces are spanned by the eigenvectors of $\mathbf{C}(\mathbf{0})^{\mathrm{T}}\mathbf{K}\mathbf{C}(\mathbf{0})$, classified according to whether they are in the null spaces of both $\mathbf{C}(\mathbf{0})$ and $\mathbf{C}_w(\mathbf{0})$ (translations Φ_u), in only the null space of $\mathbf{C}(\mathbf{0})$ (soft modes Φ_w), or in neither of these null spaces (high-frequency modes Φ_v). The equations of motion are used to eliminate Φ_v, as shown in (11.27). (e) The continuum fields arise naturally from the degrees of freedom in (d), satisfying the equations of motion (11.4) and the constitutive relations (11.5).

11.2.4 The homogenisation procedure

Here we obtain the continuum equations of motion (11.4) systematically from the discrete lattice equations of motion, with details given in Appendix III.2.1. Fig. 11.2 outlines the overall approach, bringing together concepts introduced in the previous subsections.

Before introducing the lattice equations of motion, we define the $n_s d \times n_s d$ mass matrix $\tilde{\mathbf{M}}_a$ as the diagonal matrix whose (h,h)-element is the mass M_h of the particle at the hth site. These sites are connected by n_b bonds, modelled as Hookean springs with spring constants $\{K_k\}_{k=1}^{n_b}$. The $n_b \times n_b$ spring constant matrix is a diagonal matrix whose (k,k)-element is K_k. We now consider how the masses $\tilde{\mathbf{M}}_a$ and spring constants $\mathbf{K}_a$ scale with the lattice constant a. In the continuum limit, the mass density of the system should be preserved, so $\tilde{\mathbf{M}}_a = a^d \tilde{\mathbf{M}}$, where $\tilde{\mathbf{M}}$ is independent of a. The scaling of spring constants $\mathbf{K}_a$ with lattice constant a depends on the nature of the elastic bonds connecting the sites. For $d = 3$, we model the bonds as linearly elastic rods. The spring constant for the extension of a rod is given by $K = EA/L_0$, where A is the rod's cross-sectional area, L_0 is its rest length, and E is the Young modulus. We consider the limit $a \to 0$ while preserving the shape and material of the rods, so that $A \propto a^2$ and $L_0 \propto a$, leading to $K \propto a$. For $d = 2$, we treat the bonds as linearly elastic ribbons, for which the spring constants are independent of a, consistent with a physical quasi-two-dimensional realisation, in which out-of-plane thickness is kept constant and in-plane dimensions scale as a. In this case, $A \propto a$ and $L \propto a$, so K is independent of a. These two cases are summarised together as $\mathbf{K}_a = a^{d-2}\mathbf{K}$, where $\mathbf{K}$ is independent of the microscopic length scale a.

We now turn to the equations of motion for the degrees of freedom of the discrete lattice $\Phi^{(m_1,\ldots,m_d)}$, which we replace with the continuum fields $\Phi(\mathbf{x},t)$. To relate these fields to the forces acting on lattice sites, we need a continuum equivalent of the equilibrium matrix. Replacing the quantities in (10.12) with their continuum equivalents, we obtain

$$\begin{aligned}\mathbf{f}_h(\mathbf{x},t) &= \sum_{k=1}^{n_b} \left(\mathtt{t}_k(\mathbf{x},t)\delta_{h,h_1(k)} - \mathtt{t}_k(\mathbf{x}-\mathbf{n}_k,t)\delta_{h,h_2(k)}\right)\hat{\mathbf{s}}_k \\ &= -\sum_{k=1}^{n_b} (\delta_{h,h_2(k)} - \delta_{h,h_1(k)})\mathtt{t}_k(\mathbf{x},t)\,\hat{\mathbf{s}}_k + \sum_{k=1}^{n_b} \delta_{h,h_2(k)}(\mathbf{n}_k \cdot \nabla \mathtt{t}_k(\mathbf{x},t))\,\hat{\mathbf{s}}_k,\end{aligned} \tag{11.23}$$

where $\mathbf{n}_k$ is the lattice vector defined after (10.12). Comparing (11.23) and (11.14) shows that the continuum equilibrium matrix is related to the continuum compatibility matrix by replacing all spatial derivatives ∂_{x_j} with $-\partial_{x_j}$ and taking the matrix transpose. Therefore, the continuum equilibrium matrix for the perturbed lattice is $\mathbf{Q}_{p,\mathbf{x}} = (\mathbf{C}_0 + \varepsilon\mathbf{C}_w - \mathbf{C}_\mathbf{x} - \varepsilon\mathbf{C}_{\mathbf{x},w} + O(\varepsilon^2))^{\mathrm{T}}$.

Using the equilibrium matrix, we can write Newton's second law for the n_s degrees of freedom as

$$\tilde{\mathbf{M}}_a \partial_t^2 \Phi(\mathbf{x},t) = -\mathbf{Q}_{p,\mathbf{x}} \mathbf{K}_a \mathbf{C}_{p,\mathbf{x}} \Phi(\mathbf{x},t) + F_a(\mathbf{x},t), \tag{11.24}$$

where $F_{\text{ext},a} = a^d F_{\text{ext}}$ is a vector of external forces whose scaling with a preserves the force density F_{ext}.

To reduce the number of degrees of freedom in (11.24), we first use the decomposition in (11.21) to obtain three sets of equations, one for each projection $\mathbf{P}_u, \mathbf{P}_w, \mathbf{P}_v$, given as (III.2) in Appendix III.2.1. We use the $\mathbf{P}_v$ projection (III.2c) to eliminate the high-frequency degrees of freedom Φ_v. Retaining only terms lowest order in δ (i.e. neglecting $O(\delta^2)$ terms) in the $\mathbf{P}_v$ projection, we find

$$\mathbf{P}_v \mathbf{C}_0^{\mathrm{T}} \mathbf{K} \mathbf{C}_0 \mathbf{P}_v \Phi_v(\mathbf{x},t) \approx -\mathbf{P}_v \mathbf{C}_0^{\mathrm{T}} \mathbf{K} (\mathbf{C}_{\mathbf{x}} \Phi_u(\mathbf{x},t) + (\mathbf{C}_{\mathbf{x}} + \varepsilon \mathbf{C}_w) \Phi_w(\mathbf{x},t)), \tag{11.25}$$

where, significantly, the inertial terms are neglected because they are $O(\delta^2)$, so (11.25) is a constraint that uniquely determines Φ_v.

We now solve (11.25) for Φ_v and substitute the result into the other projections of the equations of motion. We define $\mathbf{P}_v \mathbf{C}_0^{\mathrm{T}} \mathbf{K} \mathbf{C}_0 \mathbf{P}_v|_{\mathcal{U}_v}$ to be the restriction of $\mathbf{P}_v \mathbf{C}_0^{\mathrm{T}} \mathbf{K} \mathbf{C}_0 \mathbf{P}_v$ to the subspace $\mathcal{U}_v$ spanned by the high-frequency modes $\{V_k\}_{k=1}^{n_s d-(d+n_w)}$. Since $\Phi_v(\mathbf{x},t)$ is a linear combination of eigenvectors of $\mathbf{C}_0^{\mathrm{T}} \mathbf{K} \mathbf{C}_0$ with non-zero eigenvalues, this restriction is invertible. We then define

$$\mathbf{X} = \left(\mathbf{P}_v \mathbf{C}_0^{\mathrm{T}} \mathbf{K} \mathbf{C}_0 \mathbf{P}_v|_{\mathcal{U}_v}\right)^{-1} \mathbf{P}_v, \tag{11.26}$$

and solve (11.25) for Φ_v in terms of Φ_u, Φ_w:

$$\Phi_v \approx -\mathbf{X} \mathbf{C}_0^{\mathrm{T}} \mathbf{K} (\mathbf{C}_{\mathbf{x}} \Phi_u + (\mathbf{C}_{\mathbf{x}} + \varepsilon \mathbf{C}_w) \Phi_w). \tag{11.27}$$

This result shows that $|\Phi_v(\mathbf{x},t)|$ is $O(\delta)$. The matrix $\mathbf{X}$ is also known as the Moore–Penrose pseudoinverse of $\mathbf{C}_0^{\mathrm{T}} \mathbf{K} \mathbf{C}_0$ [42].

We now use (11.27) to eliminate $\Phi_v(\mathbf{x},t)$ from the equations of motion. Substituting (11.27) into the projections of (11.24) onto the subspaces spanned by the uniform translations (III.2a) and soft modes (III.2b) results in the continuum equations of motion

$$\begin{aligned}
a^2 \mathbf{P}_u (\tilde{\mathbf{M}} \partial_t^2 (\Phi_u + \Phi_w) - F) &= \mathbf{P}_u \mathbf{C}_{\mathbf{x}}^{\mathrm{T}} \tilde{\mathbf{K}} (\varepsilon \mathbf{C}_w \Phi_w + \mathbf{C}_{\mathbf{x}} (\Phi_u + \Phi_w)) \\
a^2 \mathbf{P}_w (\tilde{\mathbf{M}} \partial_t^2 (\Phi_u + \Phi_w) - F) &= \mathbf{P}_w \mathbf{C}_{\mathbf{x}}^{\mathrm{T}} \tilde{\mathbf{K}} (\varepsilon \mathbf{C}_w \Phi_w + \mathbf{C}_{\mathbf{x}} (\Phi_u + \Phi_w)) \\
&\quad - \varepsilon \mathbf{P}_w \mathbf{C}_w^{\mathrm{T}} \tilde{\mathbf{K}} (\varepsilon \mathbf{C}_w \Phi_w + \mathbf{C}_{\mathbf{x}} (\Phi_u + \Phi_w)),
\end{aligned} \tag{11.28}$$

where

$$\tilde{\mathbf{K}} = \mathbf{K} - \mathbf{K}\mathbf{C}_0\mathbf{P}_v\mathbf{X}\mathbf{P}_v\mathbf{C}_0^{\mathrm{T}}\mathbf{K} \tag{11.29}$$

is the effective spring constant matrix. The expression (11.29) shows that the relaxation of the high-frequency modes in the continuum limit reduces the spring constant matrix from $\mathbf{K}$ to $\tilde{\mathbf{K}}$, as detailed in Appendix III.2.1. In Appendix III.2.2, we re-cast (11.28) into the form (11.4) with constitutive relations (11.5), which we previously introduced based on symmetry considerations, giving the generalised elastic moduli in terms of the compatibility and spring constant matrices of the underlying discrete lattice in (III.5).

A key feature of the homogenisation procedure resulting in (11.4), (11.5) is that all the terms retained in the equations of motion have the same order in δ as $\delta \to 0$, to discard terms that are irrelevant to phenomena at continuum length scales. This is consistent with Item 3 in Section 11.2, and enables a reduction in the number of degrees of freedom, from $n_s d$ components in the discrete lattice equations (11.24) to $d + n_w$ components in the continuum equations (11.4). For a kagome lattice, these numbers are $n_s d = 6$ and $d + n_w = 3$. Because higher order terms in δ are discarded, we expect that solving for zero modes using the homogenised continuum theory is equivalent to solving for zero modes using $\det \mathbf{C}_p(\mathbf{q};\varepsilon) = 0$, where only the lowest-order terms in δ are retained. This equivalence is proven in Appendix III.2.4, where we show that $\det \mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q})$ of the continuum theory is proportional to the lowest-order terms in $\det \mathbf{C}_p(\mathbf{q};\varepsilon)$ when both the discrete and the continuum Maxwell criteria are satisfied.

In summary, this homogenisation procedure links our generalised continuum theory (11.4), (11.5) to microscopic lattice-based realisations. With this connection, we see that the continuum displacement field $\mathbf{u}$ represents the average site displacement within a unit cell and the additional fields $\varphi_1, \ldots, \varphi_{n_w}$ represent local soft modes of the lattice. Gapping the lattice by perturbing its configuration away from a gapless state gives rise to dependence of the strain energy function on the fields $\varphi_1, \ldots, \varphi_{n_w}$ in addition to their gradients. Our results therefore provide design principles for realising topological phenomena in the continuum using mechanical metamaterials. These principles dictate how the underlying lattice geometry gives rise to the desired number of local soft modes n_w and generalised elastic moduli.

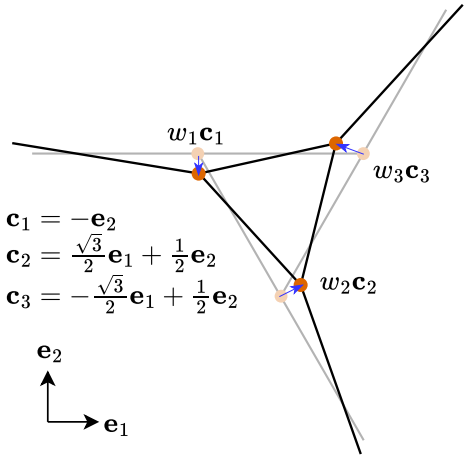


Fig. 11.3 Constructing distorted kagome lattices. The positions of sites in the kagome unit cell are shifted along directions parallel to unit vectors $\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3$, indicated using blue arrows, by distances $w_1 = 1, w_2 = 1.5, w_3 = 1.75$ respectively. The undistorted unit cell has three sites placed at the vertices of an equilateral triangle.

Chapter 12

Topological floppy modes in the continuum

We study floppy modes within our augmented elasticity theory, which has n_w additional fields. By showing that $n_w \geq 1$ is required to capture topological edge modes, and that $n_w \geq 2$ is necessary for the presence of bulk Weyl points, we classify topological floppy modes in the continuum according to the number of additional fields required to capture them. The topological invariants in the continuum take the form (11.10), and depend on the effective compatibility matrix $\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q})$ via its determinant. We establish the results of this chapter using the following properties of the effective compatibility matrix. The quantity $\det\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q})$ depends on $\mathbf{q}$ only via i$\mathbf{q}$, and is a polynomial in the components of i$\mathbf{q}$. By collecting terms with the same number of factors of q_j, the components of $\mathbf{q}$, we obtain

$$\det\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}) = \sum_{m=d}^{d+n_w} \frac{\mathrm{i}^m}{m!} P_m(\mathbf{q}), \tag{12.1}$$

where $P_m(\mathbf{q})$ is a homogeneous polynomial of degree m in the components of $\mathbf{q}$, with real coefficients. The highest degree present is $d+n_w$ because $\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q})$ has dimensions $(d+n_w)\times(d+n_w)$ and each element in the matrix has at most one factor of the form q_j. Since $\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q})$ has a null space of at least d dimensions, corresponding to the d uniform translations, we conclude that $P_m(\mathbf{q}) = 0$ for $0 \leq m < d$, using an argument similar to that in Appendix A.4 of Baardink [155].

12.1 Topological edge modes

Because setting $n_w = 0$ in the continuum equations of motion (11.4), (11.5) gives standard linear elasticity, which cannot capture topological polarisation, we conclude that $n_w \geq 1$ is necessary for topological polarisation in the continuum. We use (11.10) to define an invariant for edge modes in the continuum. Furthermore, we show that in two-dimensional elasticity, the presence of shear-dominant Guest–Hutchinson modes [159], i.e. those associated with a positive Poisson ratio, is sufficient for topological polarisation in the continuum.

12.1.1 Topological invariant for edge modes

Here we define an invariant based on (11.10) that counts only edge modes that are visible on large length scales. To define these *continuum edge modes*, consider an edge normal to $\mathbf{e}_k$, where $\{\mathbf{e}_j\}_{j=1}^d$ is an orthonormal basis. A floppy mode localised on this edge corresponds to a complex wavevector $\mathbf{q} = \sum_{j=1}^d q_j \mathbf{e}_j$ parametrised by expressing the complex component q_k normal to the edge as a function of the real components $\{q_j\}_{j\neq k}$ along the edge, i.e. $q_k = F\left(\{q_j\}_{j\neq k}\right)$. Because spatial variations over large length scales correspond to wavevectors near $\mathbf{q} = \mathbf{0}$, we define continuum edge modes to be edge floppy modes whose wavevectors satisfy

$$\lim_{\substack{q_j\to 0\\ j\neq k}} q_k = \lim_{\substack{q_j\to 0\\ j\neq k}} F\left(\{q_j\}_{j\neq k}\right) = 0. \tag{12.2}$$

Edge floppy modes for which this limit is non-zero correspond to spatial variations normal to the edge with length scale given by the limit of q_k^{-1} as $q_j \to 0$ for $j \neq k$, so they would not be visible as edge modes at macroscopic length scales L much larger than this value.

To define a topological invariant to count these continuum edge modes, let $N_{+,k}$ and $N_{-,k}$ be the numbers of continuum edge modes decaying along the $+\mathbf{e}_k$ and $-\mathbf{e}_k$ directions, respectively. Therefore, $N_{+,k} - N_{-,k}$ is the difference between the numbers of continuum edge modes localised on opposite edges normal to $\mathbf{e}_k$, characterising topological polarisation. Using the approach of Saremi and Rocklin [16] to exclude edge modes irrelevant to the continuum limit, we show in Appendix III.3 that the difference $N_{+,k} - N_{-,k}$ is given by the invariant

$$N_{+,k} - N_{-,k} = \lim_{\alpha\to 0^+} \frac{1}{\pi \mathrm{i}} \int_{-\alpha^c}^{\alpha^c} \frac{\frac{\mathrm{d}}{\mathrm{d}\tau} \det \mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q}_\alpha^k(\tau))}{\det \mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q}_\alpha^k(\tau))} \mathrm{d}\tau, \tag{12.3}$$

where the integration path is given by

$$\mathbf{q}_\alpha^k(\tau) = \tau \mathbf{e}_k + \alpha \sum_{j=1, j\neq k}^{d} \mathbf{e}_j$$

for real $\tau \in [-\alpha^c, \alpha^c]$, and exponent c satisfying $0 < c < 1/(d+n_w)$. Along this path, τ ranges over the q_k component and all other components of $\mathbf{q}$ are set to α. The path of integration in (12.3) is not a closed path because there is no Brillouin zone in the continuum, i.e. there are no non-contractible loops in $\mathbf{q}$-space for which the components of $\mathbf{q}$ are real.

However, we can choose loops in the complex plane that enclose the complex roots $\tau = q_k$ of the polynomial $\det \mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q}_\alpha^k(\tau))$. Cauchy's argument principle (e.g. Asmar and Grafakos [154]) applied to such loops counts the numbers of edge modes $N_{+,k}, N_{-,k}$. However, when we use loops involving complex components of $\mathbf{q}$, this topological invariant is not calculated solely from bulk information. To compute the numbers of edge modes using an integral involving only real components of $\mathbf{q}$, we take the limit $\alpha \to 0^+$ in (12.3), for which contributions from complex $\mathbf{q}$ components are negligible, as proven in Appendix III.3. Taking the limit $\alpha \to 0^+$ also ensures that $N_{+,k} - N_{-,k}$ counts only floppy modes that satisfy the defining property (12.2) of continuum edge modes. We have therefore defined a topological invariant for edge modes that relies only on information in continuum compatibility matrix $\mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q})$ at small real values of q_j, as reflected by (12.3).

The requirement that the exponent c satisfy $c < 1/(d+n_w)$ ensures that the domain of integration $[-\alpha^c, \alpha^c]$ is sufficiently large to capture the roots $\tau = q_k$ of $\det \mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q}_\alpha^k(\tau))$ corresponding to the continuum edge modes. We show in Appendix III.3 that this requirement can be relaxed to $0 < c < 1$ when the components q_k for all of the edge modes are analytic functions of α.

12.1.2 Topological polarisation in two-dimensional elasticity

Here we show how topological polarisation in two-dimensional elasticity is fully characterised by the generalised elastic moduli in (11.5). We show that a strip of two-dimensional material whose generalised elastic moduli satisfy the continuum Maxwell criterion (11.9) has exactly two continuum edge floppy modes, so we define topological polarisation to mean that a strip can be cut such that both floppy modes are localised on the same edge. By contrast, in the unpolarised case, every strip has floppy modes localised on opposite edges. We show that topological polarisation is determined by the sign of a discriminant Δ. When $\Delta > 0$, we define polarisation directions $\mathbf{p}_1$ and $\mathbf{p}_2$ that characterise the edge orientations $\mathbf{e}_2$ for which floppy modes are asymmetrically localised, as shown in Fig. 12.1. These polarisation directions have a role in the continuum analogous to that of the polarisation lattice vector $\mathbf{r}_{\mathrm{T}}$ introduced in Kane and Lubensky [12] for discrete lattices. Our key result is that systems with the discriminant $\Delta > 0$ are topologically polarised, whereas those with $\Delta < 0$ are not. These two cases correspond to systems with shear-dominant (positive Poisson ratio) and dilation-dominant (negative Poisson ratio) Guest–Hutchinson modes [159, 160], respectively.

Rocklin et al. [159] showed that the Guest–Hutchinson mode, which is a zero-energy uniform strain, in a two-dimensional topologically polarised Maxwell lattice is necessarily shear-dominant. Here we show that the converse is true in the continuum: the condition that the Guest–Hutchinson mode is shear-dominant is *sufficient* for topological polarisation.

To define the discriminant Δ, we first consider a two-dimensional strip of material described by our augmented elasticity theory (11.4), (11.5) and an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2\}$, where $\mathbf{e}_1$ is parallel to the strip edges. We consider only systems with no bulk floppy modes, so there are no purely real non-zero solutions $\mathbf{q}$ to $\det \mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q}) = 0$, where $\mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q})$ is the effective compatibility matrix introduced in Section 11.1.1. For a zero mode, the wavevector $\mathbf{q} = q_1 \mathbf{e}_1 + q_2 \mathbf{e}_2$ is given by $q_2 = F_j(q_1)$, where $j = 1, \dots, 2 + n_w$ indexes the zero modes for a given value of q_1, which specifies the spatial variation along the strip edges. Here, each F_j is a function that satisfies $\det \mathbf{P}_K^{\mathrm{T}} \mathbf{L}(q_1 \mathbf{e}_1 + F_j(q_1) \mathbf{e}_2) = 0$ for all q_1 on a sufficiently small neighbourhood of zero. To study the properties of these zero modes, we write

$$P_m(\mathbf{q}) = P_m(q_1 \mathbf{e}_1 + q_2 \mathbf{e}_2) = \sum_{j=0}^{m} A_{j,m-j} q_1^j q_2^{m-j} \tag{12.4}$$

in (12.1), where the polynomial coefficients $A_{j,k}$ are real. Near $\mathbf{q} = \mathbf{0}$, $\det \mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q})$ is controlled by the terms lowest order in $\mathbf{q}$, corresponding to $m = 2$. We define the discriminant to classify these terms. The polynomial $P_2(\mathbf{q})$ is a quadratic form,

$$P_2(q_1 \mathbf{e}_1 + q_2 \mathbf{e}_2) = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} A_{2,0} & \frac{1}{2} A_{1,1} \\ \frac{1}{2} A_{1,1} & A_{0,2} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}^{\mathrm{T}} \mathbf{A}_2 \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}, \tag{12.5}$$

where we defined $\mathbf{A}_2$ to be the matrix shown after the first equality. We then define the discriminant

$$\Delta = A_{1,1}^2 - 4 A_{2,0} A_{0,2} = -4 \det \mathbf{A}_2, \tag{12.6}$$

which is independent of our choice of basis $\{\mathbf{e}_1, \mathbf{e}_2\}$, because under a change of basis corresponding to the orthogonal tensor $\mathbf{R}$, $\mathbf{A}_2$ transforms to $\mathbf{R} \mathbf{A}_2 \mathbf{R}^{\mathrm{T}}$. Since $\det \mathbf{R} \mathbf{A}_2 \mathbf{R}^{\mathrm{T}} = \det \mathbf{A}_2$, Δ is a quantity associated with the material under consideration, independent of coordinate system. In Appendix III.4, we show that $\Delta \neq 0$ implies the existence of exactly one uniform strain associated with zero stress, i.e. a zero-energy deformation. This uniform strain is known as a Guest–Hutchinson mode [14, 159, 160], and is shear-dominant (dilation-dominant) when its associated strain tensor $\mathbf{E}_{\mathrm{GH}}$ has principal strains, i.e. eigenvalues, of opposite (equal) signs, equivalent to $\det \mathbf{E}_{\mathrm{GH}} < 0$ ($\det \mathbf{E}_{\mathrm{GH}} > 0$). We show in Appendix III.4 that $\Delta > 0$ and $\Delta < 0$ are equivalent to the Guest–Hutchinson mode being shear-dominant and dilation-dominant, respectively, and that these two cases correspond to positive and

negative Poisson ratios, respectively, with respect to the principal strain directions of the Guest–Hutchinson mode.

We now show that there are two continuum edge modes, which by definition satisfy $\lim_{q_1\to 0} F_j(q_1) = 0$, i.e. (12.2) for $d = 2$. Because the coefficients $A_{j,k}$ in $P_m(\mathbf{q})$ depend on the basis $\{\mathbf{e}_1, \mathbf{e}_2\}$, and $A_{0,2} = 0$ only for specific, non-generic choices of basis, we consider the generic case $A_{0,2} \neq 0$. We can always choose edge directions such that the orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ with $\mathbf{e}_1$ parallel to the edges corresponds to $A_{0,2} \neq 0$. In this case, setting $q_1 = 0$ in $\det \mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}) = 0$ results in

$$q_2^2\left(A_{0,2} + \sum_{m=3}^{2+n_w} A_{0,m} q_2^{m-2}\right) = 0,$$

which has $q_2 = 0$ as a solution with multiplicity 2. Therefore, there are exactly two continuum edge modes when $A_{0,2} \neq 0$, and we represent these modes using the functions F_1 and F_2. Using the theory of Puiseux series [161], we show in Appendix III.5 that F_1 and F_2 are analytic functions on a neighbourhood of zero. Therefore, the continuum floppy modes can be studied using their power series expansions

$$F_j(q_1) = \sum_{k=0}^{\infty} \frac{1}{k!} F_j^{(k)}(0) q_1^k. \tag{12.7}$$

Setting $q_2 = F_j(q_1)$ in $\det \mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}) = 0$ and differentiating with respect to q_1 twice shows that the first derivatives evaluated at zero, $F_j'(0)$, satisfy the quadratic equation $P_2\Big(1, F_j'(0)\Big) = 0$, giving

$$F_j'(0) = -\frac{A_{1,1} + (-1)^j\sqrt{\Delta}}{2A_{0,2}}. \tag{12.8}$$

Since Δ is the discriminant for the quadratic equation satisfied by $F_j'(0)$, if $\Delta < 0$, $F_1'(0)$ and $F_2'(0)$ are a complex conjugate pair. Setting $\beta_1 = \mathrm{Im}\, F_1(0) = -\mathrm{Im}\, F_2(0)$, where $\mathrm{Im}\, z$ is the imaginary part of the complex number z, we see from (12.7) that the decay constants are $\mathrm{Im}\, F_1(q_1) = \beta_1 q_1 + O(q_1^2)$ and $\mathrm{Im}\, F_2(q_1) = -\beta_1 q_1 + O(q_1^2)$. The number of modes localised on each edge must remain the same as q_1 is varied, otherwise there would be bulk zero modes corresponding to q_1 values at which edge modes switch sides. Therefore, the number of modes localised on each edge is completely determined by the terms lowest order in q_1 present in $\mathrm{Im}\, F_j(q_1)$, i.e. $\pm\beta_1 q_1$. Because the decay constants have opposite signs in this case, $\Delta < 0$ corresponds to the *unpolarised* case in which floppy modes are localised on opposite edges.

When $\Delta > 0$, $F_1'(0)$ and $F_2'(0)$ are both real. The *soft directions* in the bulk are directions parallel to $\mathbf{e}_1 + F_j'(0)\mathbf{e}_2$ for $j = 1,2$. These directions, first introduced in Saremi and Rocklin [16], correspond to solutions of $P_2(\mathbf{q}) = A_{2,0}q_1^2 + A_{1,1}q_1q_2 + A_{0,2}q_2^2 = 0$, i.e. solutions of $\det \mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}) = 0$ to lowest order in q_1 and q_2. If a solution to $P_2(\mathbf{q}) = 0$ corresponds to wavevectors normal to the edge, then it takes the form $q_1 = 0$, $q_2 = \tilde{q}$ for any $\tilde{q} \in \mathbb{R}$. From this, we see that considering the generic case $A_{0,2} \neq 0$, which excludes such solutions, is equivalent to choosing an edge that is not normal to a soft direction. Here we prove that the material is topologically polarised when $\Delta > 0$. This proof proceeds by the following steps:

1. Polarisation directions $\mathbf{p}_1$ and $\mathbf{p}_2$ are constructed as vectors orthogonal to the soft directions. We show that these polarisation directions indicate the edge on which the continuum floppy modes are localised.

2. Strips with edges normal to $\hat{\mathbf{n}}$ have both floppy modes localised on the edge to which $\hat{\mathbf{n}}$ points if and only if $\hat{\mathbf{n}}$ makes an acute angle with both polarisation directions, as shown on Fig. 12.1.

3. Because $\Delta > 0$, the soft directions are distinct. Therefore, there always exist normals $\hat{\mathbf{n}}$ such that the corresponding strips are topologically polarised, as shown by the thick black arcs on Fig. 12.1.

We represent the two soft directions S_1 and S_2 using the unit vectors

$$\mathbf{e}_1^{S_j} = \frac{\mathbf{e}_1 + F_j'(0)\mathbf{e}_2}{|\mathbf{e}_1 + F_j'(0)\mathbf{e}_2|}, \tag{12.9}$$

for $j = 1,2$. Then, we construct two distinct orthonormal bases $\{\mathbf{e}_1^{S_1}, \mathbf{e}_2^{S_1}\}$ and $\{\mathbf{e}_1^{S_2}, \mathbf{e}_2^{S_2}\}$. Each polarisation direction $\mathbf{p}_j$ is given by $\pm\mathbf{e}_2^{S_j}$, where the sign is chosen such that $\mathbf{p}_j$ points toward the soft edge. As we argued previously, the edges on which floppy modes are localised are indicated by the signs of the terms of lowest order in q_1 present in $\mathrm{Im}\, F_j(q_1)$. When $\Delta > 0$, the terms involving $F_j'(0)$ are real, so the imaginary part of q_2 is

$$\mathrm{Im}\, F_j(q_1) = \frac{1}{2}\mathrm{Im}\, F_j''(0)q_1^2 + O(q_1^3),$$

showing that $F_j''(0)$ determines the direction of decay of the floppy mode associated with F_j. To compute $F_j''(0)$, we set $q_2 = F_j(q_1)$ in $\det \mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}) = 0$ and differentiate with respect to q_1 three times to obtain

$$F_j''(0) = \mathrm{i}\frac{2(-1)^j \sum_{k=0}^{3} A_{3-k,k}\left(F_j'(0)\right)^k}{3\sqrt{\Delta}}, \tag{12.10}$$

where (12.8) has been used. Because $F_j'(0)$ is real, $F_j''(0)$ is purely imaginary. We now show that each continuum floppy mode can be associated with a soft direction. The continuum floppy modes take the form

$$\begin{aligned}\mathbf{q} = q_1\mathbf{e}_1 + F_j(q_1)\mathbf{e}_2 &= q_1\big(\mathbf{e}_1 + F_j'(0)\mathbf{e}_2\big) + \mathrm{i}\frac{1}{2}\big(\mathrm{Im}\,F_j''(0)\big)q_1^2\mathbf{e}_2 + O(q_1^3) \\ &= q_1\sqrt{1+\Big(F_j'(0)\Big)^2}\,\mathbf{e}_1^{S_j} + \mathrm{i}\frac{1}{2}\big(\mathrm{Im}\,F_j''(0)\big)\,q_1^2\mathbf{e}_2 + O(q_1^3),\end{aligned} \tag{12.11}$$

using the expansion (12.7) and the definition (12.9). For sufficiently small q_1, each continuum edge floppy mode therefore has sinusoidal wavefronts normal to a soft direction $\mathbf{e}_1^{S_j}$, which we associate with the floppy mode.

We identify the polarisation direction $\mathbf{p}_j$ by considering the case $\mathbf{e}_1 = \mathbf{e}_1^{S_j}$, i.e. when the strip edges are parallel to a soft direction. In this basis, we adopt the notation $q_2 = F_{S_j}(q_1)$, so the wavevector of the associated floppy mode is $\mathbf{q} = q_1\mathbf{e}_1^{S_j} + F_{S_j}(q_1)\mathbf{e}_2^{S_j}$. Because $\mathbf{e}_1^{S_j}$ is a soft direction, $F_{S_j}'(0) = 0$, implying that $A_{0,2} = 0$ with respect to the basis $\{\mathbf{e}_1^{S_j}, \mathbf{e}_2^{S_j}\}$. Using (12.10), we find $F_{S_j}''(0) = -\mathrm{i}\frac{2}{3}\frac{A_{3,0}}{A_{1,1}}$, where $A_{j,k}$ here are computed with respect to the basis $\{\mathbf{e}_1^{S_j}, \mathbf{e}_2^{S_j}\}$. For $j = 1, 2$, we define the polarisation directions

$$\mathbf{p}_j = -\mathrm{sgn}(\mathrm{Im}\,F_{S_j}''(0))\mathbf{e}_2^{S_j} = \mathrm{sgn}\left(\frac{A_{3,0}}{A_{1,1}}\right)\mathbf{e}_2^{S_j}, \tag{12.12}$$

where $\mathrm{sgn}(x)$ is the sign of x, so that $\mathbf{p}_j$ points towards the edge on which the floppy mode is localised. Because we assumed that $A_{0,2} \neq 0$ to derive this result, which is equivalent to requiring that the normal vector $\mathbf{e}_2^{S_j}$ not point along a soft direction, (12.12) is valid whenever the soft directions are not orthogonal. In Appendix III.6, we generalise this procedure to include the special case of orthogonal soft directions, provided $1 \leq n_w \leq 2$.

These polarisation directions $\mathbf{p}_j$ determine the floppy mode localisation for any edges not normal to a soft direction. To show this, we consider a strip with edges parallel to $\mathbf{e}_1$ and normal to $\mathbf{e}_2$. We write $\hat{\mathbf{n}} = \mathbf{e}_2$ to emphasise its role as the normal vector to the strip edges. Provided $\hat{\mathbf{n}} \cdot \mathbf{p}_j \neq 0$, we show in Appendix III.7 that the floppy mode associated with the soft direction $\mathbf{e}_1^{S_j}$ has wavevector

$$\mathbf{q} = -\frac{\mathrm{sgn}(\mathrm{Im}\,F_{S_j}''(0))}{\hat{\mathbf{n}} \cdot \mathbf{p}_j} q_1\mathbf{e}_1^{S_j} - \mathrm{i}\frac{|\mathrm{Im}\,F_{S_j}''(0)|}{2(\hat{\mathbf{n}} \cdot \mathbf{p}_j)^3} q_1^2\hat{\mathbf{n}} + O(q_1^3). \tag{12.13}$$

A version of (12.13) has previously been derived as (26) of Saremi and Rocklin [16] without using the polarisation directions we introduce. Fig. 12.1 shows the importance of considering polarisation directions in addition to soft directions. The two systems shown in Fig. 12.1(a,b)

have the same soft directions but have opposite polarisation vectors $\mathbf{p}_2$, leading to significantly different edge localisations.

The continuum edge mode associated with soft direction $\mathbf{e}_1^{S_j}$ is localised on the edge to which $\hat{\mathbf{n}}$ points if and only if $\hat{\mathbf{n}}$ makes an acute angle with $\mathbf{p}_j$. This is the case because $\hat{\mathbf{n}} \cdot \mathbf{p}_j > 0$ in (12.13) is equivalent to $\mathrm{Im}(\mathbf{q} \cdot \mathbf{e}_2) < 0$, so the mode grows exponentially as $\mathrm{e}^{\mathrm{i}\mathbf{q}\cdot\mathbf{x}}$ in the $\mathbf{e}_2$-direction, where $\mathbf{e}_2 = \hat{\mathbf{n}}$. Therefore, the polarisation directions fully determine the distribution of localised floppy modes between the edges of any strip, provided the edges are not orthogonal to a soft direction. When the Guest–Hutchinson mode is shear-dominant ($\Delta > 0$, positive Poisson ratio), it is always possible to choose edge orientations for which the same edge hosts both floppy modes, i.e. to choose a normal vector $\hat{\mathbf{n}}$ that makes an acute angle with both polarisation directions, as shown in Fig. 12.1. This reasoning enables us to compute the topological invariant (12.3) in terms of the polarisation directions as

$$N_{+,2} - N_{-,2} = -\mathrm{sgn}(\mathbf{e}_2 \cdot \mathbf{p}_1) - \mathrm{sgn}(\mathbf{e}_2 \cdot \mathbf{p}_2). \tag{12.14}$$

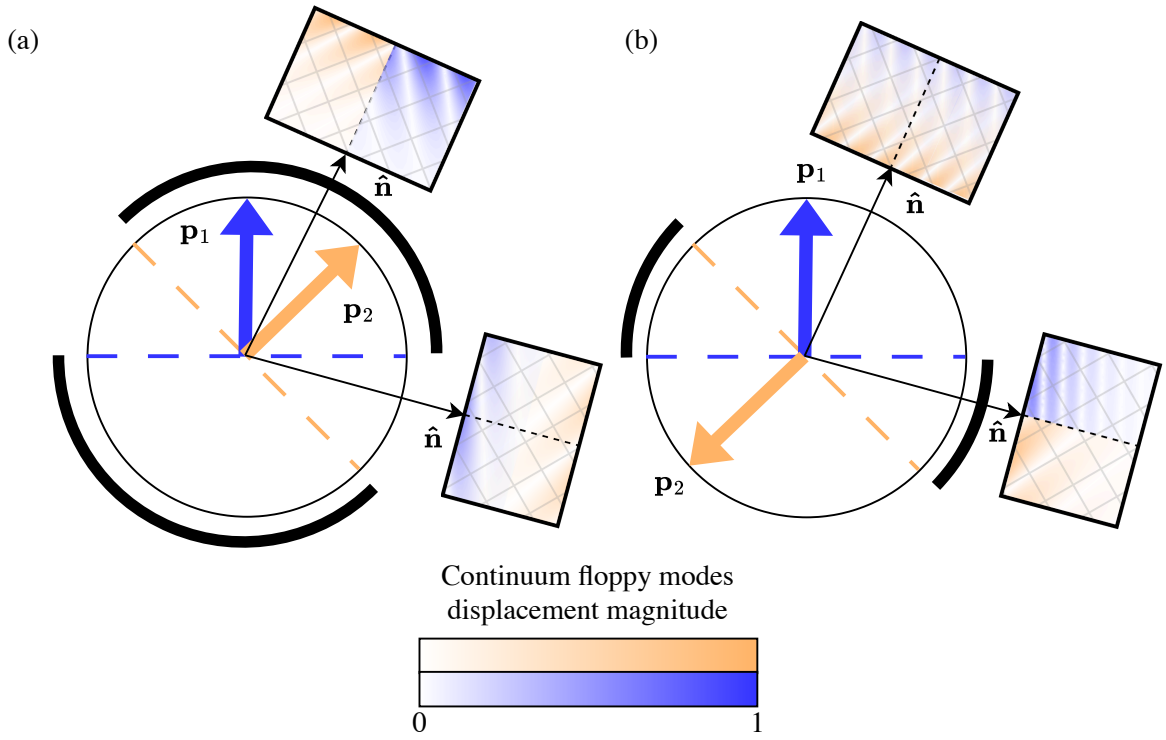


Fig. 12.1 Shear-dominant (positive Poisson ratio) Guest–Hutchinson modes imply topological polarisation in the continuum. Each of the two continuum floppy modes is associated with a polarisation direction (12.12) that can be computed from the generalised elastic moduli when $\Delta > 0$. The continuum floppy mode distribution between the edges of a strip is a function of the normal $\hat{\mathbf{n}}$ to the strip edges, and is completely determined by the polarisation directions $\mathbf{p}_1$, $\mathbf{p}_2$. The rectangles represent strips of different edge orientations and the displacement magnitudes of the two localised floppy modes are indicated in the legend. The thick black arcs indicate normal directions that make acute angles with the both polarisation directions, implying both floppy modes are localised on the same edge. The soft directions are indicated by the dashed lines. The blue and orange colours distinguish the two continuum floppy modes shown, and indicate the associated soft direction, polarisation direction, and floppy mode displacement magnitude. (a) and (b) have the same soft directions but opposite polarisation vectors $\mathbf{p}_2$, resulting in different floppy mode distributions for the same strip orientation.

We illustrate our analytical results by solving the partial differential equations (11.5) with the stress measures $\Sigma = 0$ to obtain continuum floppy modes. Provided the continuum Maxwell criterion (11.9) is satisfied, the system of equations (11.5) consists of $d + n_w$ independent equations for $d + n_w$ dependent variables. We obtain generalised elastic moduli corresponding to unpolarised and polarised cases by homogenising distorted kagome lattices, and show the solutions to $\Sigma = 0$ in Fig. 12.2(a) and (b), respectively. In both cases, $n_w = 1$, and we give details of the boundary conditions in Appendix III.9. We find a symmetric distribution of the edge modes in the unpolarised case and an asymmetric distribution in

the polarised case. The green double-headed arrows on Fig. 12.2(b) are parallel to the soft directions and indicate the wavelengths corresponding to the term parallel to the soft direction $\mathbf{e}_1^{S_j}$ in (12.13). These wavelengths are given by $2\pi|\hat{\mathbf{n}}\cdot\mathbf{p}_j|/|q_1|$, where $\hat{\mathbf{n}}$ is the normal to the horizontal boundaries, and $q_1 = 2\pi/\lambda$ is set by the wavelength λ of the Dirichlet boundary conditions on the horizontal boundaries. These arrows show that the term parallel to the soft direction $\mathbf{e}_1^{S_j}$ in (12.13) is a good approximation to the wavefronts in the continuum theory.

This completes our characterisation of topological polarisation in two-dimensional elasticity, which is independent of any underlying lattice. This characterisation is based on the soft and polarisation directions, which are computed from the generalised elastic moduli for any $n_w \geq 1$ via $A_{j,k}$. Our characterisation can be applied to discrete lattices with two edge floppy modes in the continuum, including the distorted kagome lattice [12, 159] at small distortions. Although Rocklin et al. [159] show that a distorted kagome lattice with shear-dominant Guest–Hutchinson mode can have a Kane–Lubensky polarisation vector (cf. Section 10.2) $\mathbf{r}_\mathrm{T} = \mathbf{0}$, we show in Appendix III.8 that this is consistent with our continuum results: the presence of a shear-dominant Guest–Hutchinson mode implies that we can always choose a set of primitive lattice vectors such that $\mathbf{r}_\mathrm{T} \neq \mathbf{0}$.

12.2 Weyl zero modes

In the context of ball-and-spring lattices of the type studied in Kane and Lubensky [12] and Rocklin et al. [143], Weyl points are bulk floppy modes at isolated wavevectors associated with a topological winding number. Although a distinct class of Weyl points has been studied in finite-frequency topological acoustics [162–164], no previous continuum theories for topological mechanics have considered zero-frequency Weyl modes. Here we show that our augmented elasticity theory (11.4), (11.5) can exhibit Weyl points provided $n_w \geq 2$. For the simplest case $n_w = 2$, in two-dimensions, we derive an analytical expression for the existence and location of Weyl modes, which we use to construct the phase diagram in Fig. 12.3.

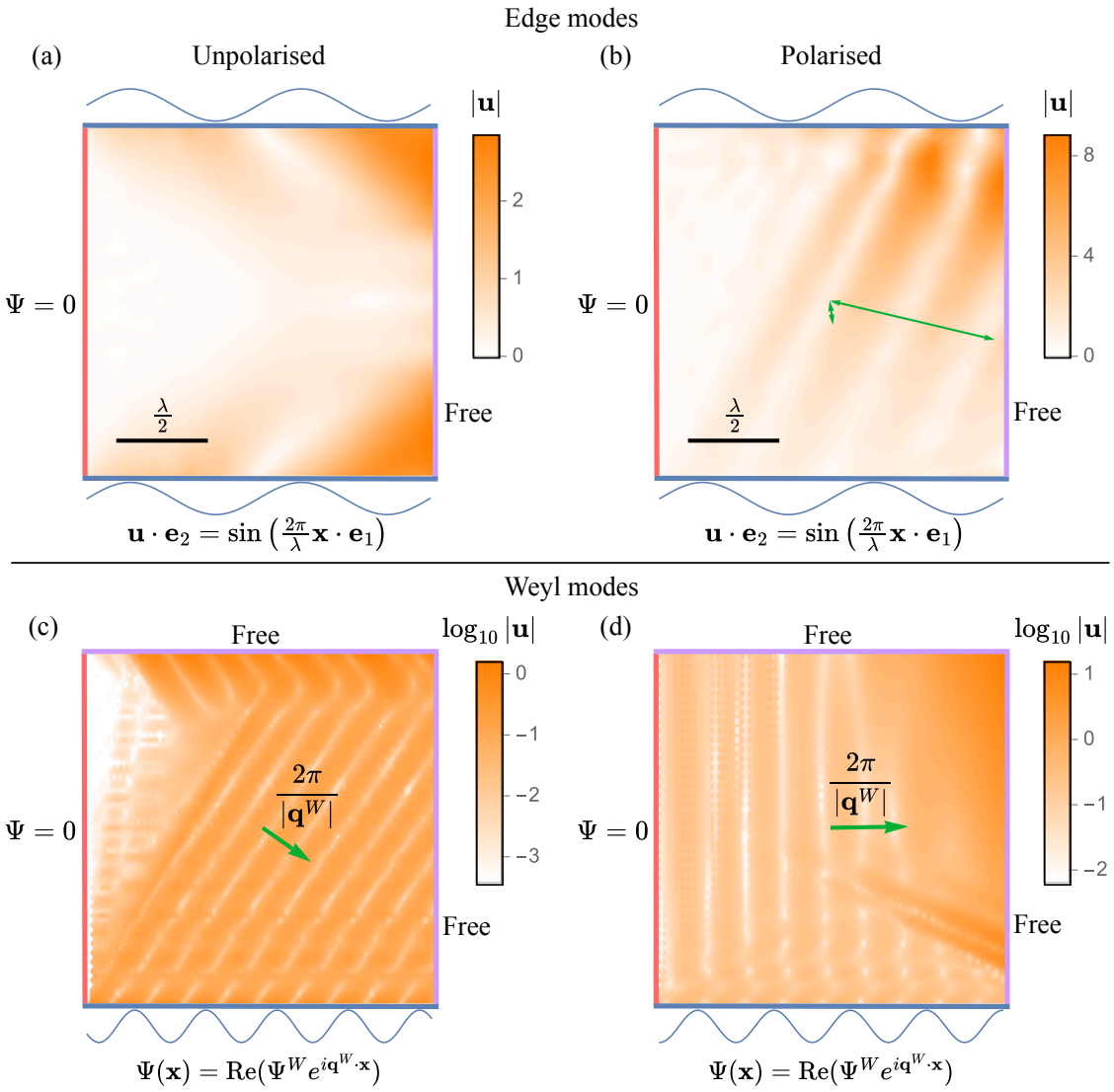


Fig. 12.2 Efficient computation for floppy modes in the continuum. We find floppy modes by solving (11.5) with stress measures $\Sigma = 0$, specifying either Dirichlet boundary conditions on the generalised displacements Ψ, or free boundary conditions if no displacements are specified (details in Appendix III.9). (a–b) Edge modes are shown using the magnitude of the displacement field $|\mathbf{u}(\mathbf{x},t)|$, for the generalised elastic moduli obtained by homogenising distorted kagome lattices: (a) unpolarised and (b) polarised. The length scale (shown by scale bars) in both cases is set by the prescribed displacements $\mathbf{u} \cdot \mathbf{e}_2 = \sin\left(\frac{2\pi}{\lambda}\mathbf{x} \cdot \mathbf{e}_1\right)$ on the two horizontal boundaries. The two green double-headed arrows in (b) are parallel to the soft directions (12.9) with length equal to the approximate wavelength predicted by the q_1 term in (12.13). (c–d) Bulk Weyl modes are shown using the logarithm of the displacement field magnitude $\log_{10}|\mathbf{u}(\mathbf{x},t)|$, for the generalised elastic moduli obtained by homogenising two different geometrical configurations of distorted double-kagome lattices. The green arrows are parallel to the Weyl mode wavevectors $\mathbf{q}^{\mathrm{W}}$ in (12.18), with length equal to the wavelength $2\pi/|\mathbf{q}^{\mathrm{W}}|$.

We begin by showing that $n_w \geq 2$ is necessary for the existence of Weyl points. These bulk floppy modes necessarily occur at isolated $\mathbf{q}$ values with real components that satisfy $\det \mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q}) = 0$. Therefore, the Weyl wavevectors $\mathbf{q}$ must satisfy two simultaneous constraints:

$$\mathrm{Re}\det \mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q}) = 0 \quad \text{and} \quad \mathrm{Im}\det \mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q}) = 0, \tag{12.15}$$

where $\mathrm{Re}\, z$ and $\mathrm{Im}\, z$ are the real and imaginary parts of the complex number z, respectively. For $n_w = 1$ and $d = 2$, these conditions applied to (12.1) give

$$P_2(\mathbf{q}) = 0 \quad \text{and} \quad P_3(\mathbf{q}) = 0$$

from the real and imaginary parts respectively. Since each $P_m(\mathbf{q})$ is a homogeneous polynomial in q_j with real coefficients, the solutions to each equation lie on lines passing through $\mathbf{q} = \mathbf{0}$. Weyl points only exist if these lines intersect at $\mathbf{q} \neq \mathbf{0}$, which is impossible in this case. Furthermore, $n_w = 0$ corresponds to solving $P_2(\mathbf{q}) = 0$, whose solutions are lines rather than isolated points in $\mathbf{q}$-space. By ruling out $n_w = 0$ and $n_w = 1$, we have demonstrated that $n_w \geq 2$ is necessary for Weyl points in two dimensions.

We now consider the simplest case $n_w = 2$ and derive conditions that the coordinates $\left(q_1^{\mathrm{W}}, q_2^{\mathrm{W}}\right)$ of these bulk Weyl modes must satisfy. Setting the real and imaginary parts to zero, as in (12.15), gives

$$-\frac{1}{2!}P_2(\mathbf{q}) + \frac{1}{4!}P_4(\mathbf{q}) = 0, \tag{12.16a}$$

$$P_3(\mathbf{q}) = 0. \tag{12.16b}$$

The coordinates of the Weyl points are located on lines in $\mathbf{q}$-space satisfying (12.16b), represented as $q_2 = \kappa q_1$, except when the line is given by $q_1 = 0$. Under the assumption $q_2 = \kappa q_1$, κ satisfies

$$P_3(\mathbf{e}_1 + \kappa \mathbf{e}_2) = \sum_{j=0}^{3} A_{j,3-j} \kappa^{3-j} = 0. \tag{12.17}$$

If $A_{0,3} \neq 0$, (12.17) is cubic and we obtain up to three real solutions for κ. We then substitute these κ values into $q_2 = \kappa q_1$ to solve for q_1^{W} in (12.16a). Therefore, if a bulk mode with an isolated wavevector exists in the continuum when $A_{0,3} \neq 0$, its $\mathbf{q}^{\mathrm{W}}$-coordinates $\left(q_1^{\mathrm{W}}, q_2^{\mathrm{W}}\right)$ must satisfy

$$q_1^{\mathrm{W}} = \pm\sqrt{12\frac{P_2(\mathbf{e}_1 + \kappa \mathbf{e}_2)}{P_4(\mathbf{e}_1 + \kappa \mathbf{e}_2)}}, \quad q_2^{\mathrm{W}} = \kappa q_1^{\mathrm{W}}. \tag{12.18}$$

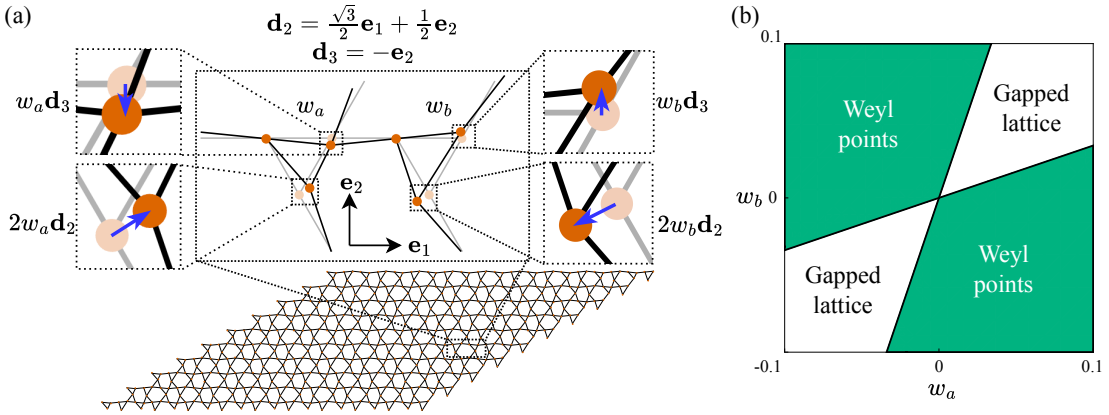


Fig. 12.3 Constructing a lattice with continuum Weyl points. (a) The distorted double-kagome lattice unit cell is based on two kagome unit cells distorted according to the parametrisation shown. The unit cell geometry is specified by the parameters w_a and w_b, which determine the magnitudes of a geometrical perturbation, shown using blue arrows, for each of the two sub-cells. The case $w_a = w_b$ reduces to the distorted kagome lattice with three sites per primitive unit cell. (b) Phase diagram for the presence of Weyl points in the lattice in part (a), calculated by requiring the expression under the square root in (12.18) to be positive. The green regions correspond to values of w_a, w_b for which Weyl points with $n_{\text{Weyl}} = \pm 1$ can be observed in the continuum.

If instead $A_{0,3} = 0$, then there are still three solutions to (12.16b), but one of the solutions corresponds to $q_1 = 0$ and leads to coordinates

$$q_1^{\text{W}} = 0, \quad q_2^{\text{W}} = \pm\sqrt{12\frac{P_2(\mathbf{e}_2)}{P_4(\mathbf{e}_2)}} = \pm\sqrt{12\frac{A_{0,2}}{A_{0,4}}}, \tag{12.19}$$

that satisfy both (12.16a) and (12.16b). In addition, there are two solutions of the form (12.18), where κ satisfies the quadratic equation $A_{1,2}\kappa^2 + A_{2,1}\kappa + A_{3,0} = 0$, obtained by setting $A_{0,3} = 0$ in (12.17). Any bulk floppy mode satisfies either (12.18) or (12.19), and both coordinates q_1^{W} and q_2^{W} must be real.

To check if a bulk floppy mode is a Weyl point, we need to show that the winding number (11.10) is non-zero for a path enclosing its wavevector $\mathbf{q}^{\text{W}}$, i.e.

$$n_{\text{Weyl}} = \frac{1}{2\pi\text{i}}\int_{\tau_0}^{\tau_1} \frac{\frac{\text{d}}{\text{d}\tau}\det\mathbf{P}_K^{\text{T}}\mathbf{L}(\mathbf{q}(\tau))}{\det\mathbf{P}_K^{\text{T}}\mathbf{L}(\mathbf{q}(\tau))}\,\text{d}\tau \neq 0, \tag{12.20}$$

where $\mathbf{q}(\tau)$ is a parametrisation of a loop in real $\mathbf{q}$-space that encloses the point $\left(q_1^{\mathrm{W}}, q_2^{\mathrm{W}}\right)$.

Fig. 10.1(c) illustrates the physical phenomena associated with a Weyl point in the bulk spectrum. The blue floppy mode is localised on the bottom edge of the material sample for $q_1 < q_1^{\mathrm{W}}$. When $q_1 = q_1^{\mathrm{W}}$, the decay length of this blue mode diverges, corresponding to a bulk mode, in contrast to the phenomenology of gapped lattices, where no bulk floppy modes exist. For $q_1 > q_1^{\mathrm{W}}$, the blue mode is an edge mode, but is localised on the top edge. Unlike a topologically polarised material, a material with Weyl points has edge floppy modes that switch sides as the wavevector is varied. Significantly, unlike the Weyl winding number in Rocklin et al. [143], the continuum definition of n_{Weyl} in (12.20) is formulated using the generalised elastic moduli and does not rely on the compatibility matrix of a discrete lattice.

To construct a lattice with Weyl points in the continuum, we start with a kagome lattice supercell consisting of two kagome unit cells, as shown in Fig. 12.3(a). These two kagome unit cells have a total of two local floppy modes at $\mathbf{q} = \mathbf{0}$, i.e. one mode for each kagome unit cell. We then apply a perturbation to each of the cells, parametrised by w_a and w_b, as shown in Fig. 12.3(a). The resulting *distorted double-kagome* (DDK) lattice has $n_w = 2$, and we show that the perturbed lattice harbours Weyl modes. In Fig. 12.3(b), we present a phase diagram of the DDK lattice for the parameters w_a and w_b. The green regions correspond to lattice geometries that exhibit Weyl points in the continuum. We obtain this phase diagram by identifying parameters w_a, w_b for which bulk floppy modes occur at isolated wavevectors with real, non-zero components given by (12.18), and that correspond to winding numbers $n_{\mathrm{Weyl}} = \pm 1$ computed using (12.20).

We illustrate Weyl phenomenology in the continuum by applying our homogenisation procedure to DDK lattices corresponding to w_a, w_b values in the Weyl regions of the phase diagram in Fig. 12.3(b). Setting the stress measures $\Sigma = 0$ to obtain floppy modes, we solve the partial differential equations (11.5) with generalised elastic moduli given by our homogenisation procedure. The results are plotted in Fig. 12.2(c) and (d), with details of the boundary conditions given in Appendix III.9. The bulk floppy modes associated with the Weyl points are visible as wavefronts that span the bulk of the sample, with direction and wavelength matching the Weyl wavevectors computed using (12.18). We have therefore shown two advantages to our continuum approach to Weyl modes: the analytically obtained Weyl phase diagram in Fig. 12.3, and efficient numerical solutions for Weyl phenomenology in a finite material sample shown in Fig. 12.2(c) and (d).

Chapter 13

Conclusions: topological floppy modes in the continuum

Motivated by the wide range of settings in which topological floppy modes occur, we presented the general form of an augmented elasticity theory in (11.4), (11.5) to capture both topological polarisation and Weyl points. Our formulation of topological mechanics in the continuum is independent of microscopic detail, because we define the continuum Maxwell criterion (11.9) and topological invariants (11.10) using only the generalised elastic moduli of our theory. We show that topological phenomena in the continuum can be classified according to the number of additional fields n_w required to capture them: topological edge modes require $n_w \geq 1$ and Weyl points require $n_w \geq 2$, as summarised in Fig. 13.1. Significantly, previous continuum theories [15–17, 156] have not captured Weyl points in the continuum.

Although our formulation does not rely on an underlying lattice, we show how our theory arises naturally from taking the continuum limit of a generic ball-and-spring lattice, using a systematic homogenisation procedure. This procedure shows that the additional fields in our theory can be interpreted as local soft modes of an underlying lattice. Therefore, our homogenisation procedure provides design principles for topological metamaterials by identifying the minimal number of local soft modes necessary for a desired behaviour. Our formulation also provides computationally efficient tools for studying topological floppy modes by relying on partial differential equations rather than simulations using the full lattice structure.

In practice, 3D-printed metamaterials are not Maxwell lattices due to effects not captured in ideal ball-and-spring models, such as bending stiffness at hinges [133, 140, 141] and nonlinearities [165, 166]. Nevertheless, asymmetric edge rigidity characteristic of topological polarisation can occur when the energy cost for hinge bending is much lower than for bond stretching [133, 140, 141]. Further work could involve generalising our approach to

incorporate these additional energy costs, as well as non-rectilinear constraints [167] and active effects [10].

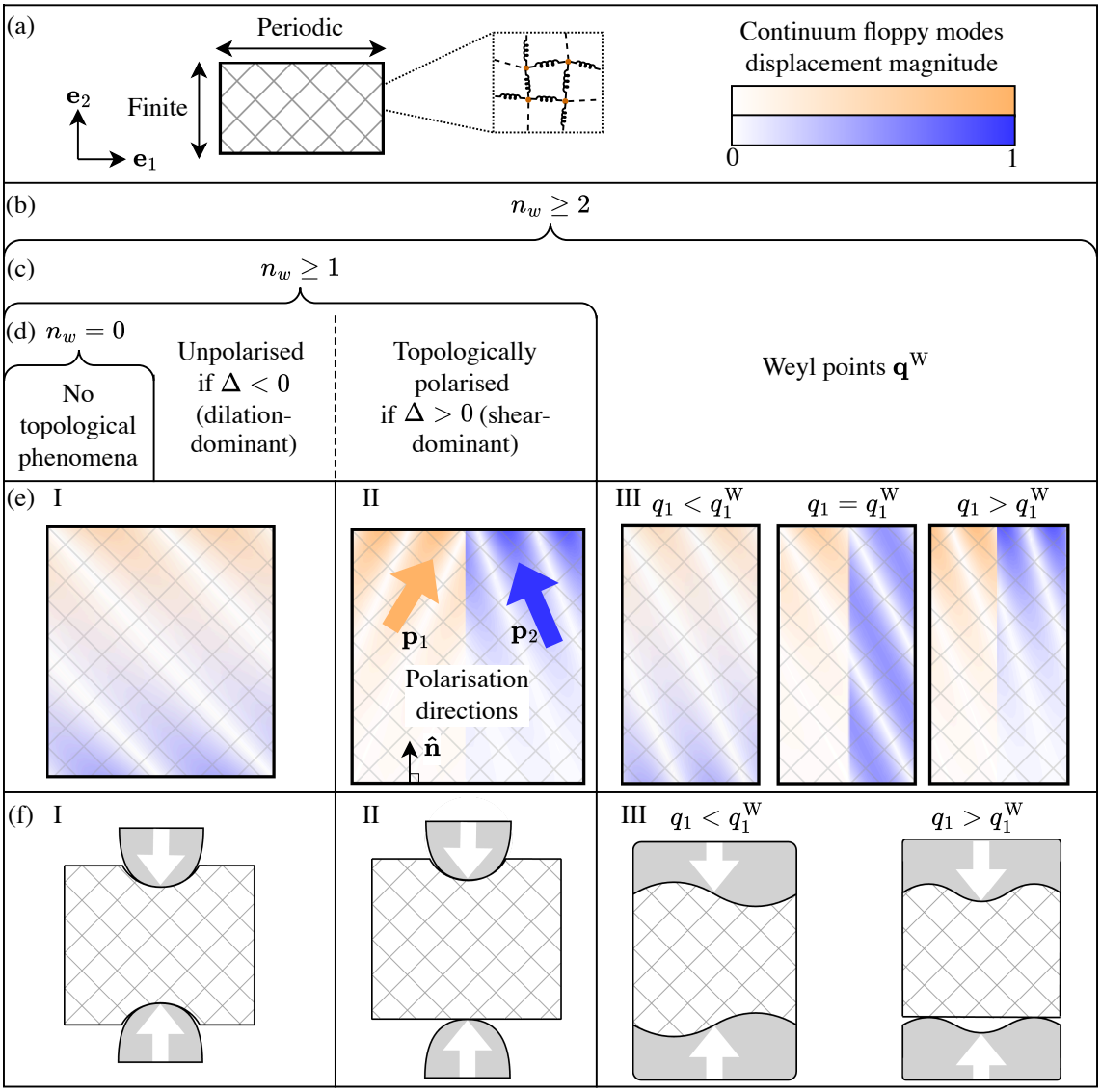


Fig. 13.1 Classification of topological phenomena in two dimensions. (a) We consider a continuum system finite in the $\mathbf{e}_2$-direction and periodic along the $\mathbf{e}_1$-direction. The displacement magnitudes of the two continuum edge floppy modes are indicated using colour gradients. (b–d) Topological phenomena are classified according to the number of additional fields n_w required to capture them. (e) Representative displacement magnitudes associated with the continuum floppy modes are shown in columns I–III, where modes localised on the same edge are displayed side-by-side for clarity. Column II illustrates how both floppy modes are localised on the same edge if the normal $\hat{\mathbf{n}}$ to the strip edges makes an acute angle with both polarisation directions $\mathbf{p}_1$ and $\mathbf{p}_2$. In column III, as q_1 passes through Weyl point coordinate q_1^{W}, the blue floppy mode switches from being localised at the bottom edge to the top edge, becoming a bulk mode at the transition point. (f) Columns I–III illustrate the influence of topological phenomena on edge rigidity using indentation tests. Edges with no localised floppy modes are more rigid, and the presence of a Weyl point in column III results in a wavevector-dependent edge rigidity. The deformations shown have been exaggerated for clarity.

Appendix III

III.1 Determinant of the compatibility matrix for a geometrically perturbed lattice

Here we derive (11.19), proving that (a) the terms lowest order in δ present in $\det \mathbf{C}_p(\mathbf{q};\varepsilon)$ are of order $d+n_w$ in δ and that (b) these terms are of order at least d in $\bar{q}_r$. We begin with (11.17) and the statements about the null spaces of $\mathbf{C}_0$ and $\mathbf{C}_0+\varepsilon\mathbf{C}_w$ that immediately follow (11.17). Let $\bar{\mathbf{C}}(\mathbf{q}) = \mathbf{C}(\mathbf{q}) - \mathbf{C}_0$ and $\bar{\mathbf{C}}_w(\mathbf{q}) = \mathbf{C}_w(\mathbf{q}) - \mathbf{C}_w$. Since the elements of $\mathbf{C}(\mathbf{q})$ and $\mathbf{C}_w(\mathbf{q})$ depend on $\mathbf{q}$ via terms proportional to $e^{\mathbf{i}\mathbf{n}\cdot\mathbf{q}}$ (cf. (11.16)), all the non-zero elements of $\bar{\mathbf{C}}(\mathbf{q})$ and $\bar{\mathbf{C}}_w(\mathbf{q})$ depend on $\mathbf{q}$ via terms proportional to $e^{\mathbf{i}\mathbf{n}\cdot\mathbf{q}}-1$. Let the kth rows of $\mathbf{C}_0$, $\bar{\mathbf{C}}(\mathbf{q})$, $\mathbf{C}_w$ and $\bar{\mathbf{C}}_w(\mathbf{q})$ be denoted c_k, $\bar{c}_k(\mathbf{q})$, w_k and $\bar{w}_k(\mathbf{q})$ respectively. These vectors are not written in bold font to distinguish them from d-dimensional vectors such as $\mathbf{q}$. Because $e^{\mathbf{i}\mathbf{n}\cdot\mathbf{q}}-1$ is $O(|\mathbf{n}\cdot\mathbf{q}|)$, and $|\mathbf{n}\cdot\mathbf{q}| = O(\delta)$, each of the vectors with an overbar $\bar{\cdot}$ is $O(\delta)$.

We write the determinant of an $m\times m$ matrix $\mathbf{A}$ as $\det\mathbf{A} = \mathcal{M}(a_1,\ldots,a_m)$, where $a_1,\ldots,a_m$ are the rows of $\mathbf{A}$ and $\mathcal{M}$ is an alternating, multilinear map from m-tuples of vectors to the complex numbers $\mathbb{C}$, that maps the rows of the $m\times m$ identity matrix to 1. This representation of the determinant is detailed in Rosén [168]. The determinant of the perturbed compatibility matrix is then

$$\det\mathbf{C}_p(\mathbf{q};\varepsilon) = \mathcal{M}(c_1(\mathbf{q};\varepsilon),\ldots,c_{n_\mathrm{b}}(\mathbf{q};\varepsilon)), \tag{III.1}$$

where the rows of $\mathbf{C}_p(\mathbf{q};\varepsilon)$ are

$$c_k(\mathbf{q};\varepsilon) = c_k + \varepsilon w_k + \bar{c}_k(\mathbf{q}) + \varepsilon\bar{w}_k(\mathbf{q}) + O(\varepsilon^2)$$

for $k=1,\ldots,n_\mathrm{b}$. Because $\mathcal{M}$ is multilinear, we can expand the right hand side of (III.1) so that it equals a sum involving terms such as $\mathcal{M}(c_1,\varepsilon w_2,\ldots,\bar{c}_{n_\mathrm{b}}(\mathbf{q}))$. Because the row space of $\mathbf{C}_0$ has dimension $n_\mathrm{b}-(d+n_w)$, there exists a choice of $n_\mathrm{b}-(d+n_w)$ rows c_k from $\mathbf{C}_0$

that form a basis for its row space [42]. Since all the other c_k are linear combinations of those basis vectors, each non-zero term in the expansion must have at least $d+n_w$ terms involving ε or $\mathrm{e}^{\mathbf{i}\mathbf{n}\cdot\mathbf{q}}-1$, so the terms of lowest order in δ are $O(\delta^{d+n_w})$. This is result (a).

To establish result (b), we note that because the row space of $\mathbf{C}_0+\varepsilon\mathbf{C}_w$ has dimension $n_\mathrm{b}-d$, there exists a choice of $n_\mathrm{b}-d$ rows $c_k+\varepsilon w_k$ that form a basis for this row space. Because all the other $c_k+\varepsilon w_k$ are linear combinations of those basis vectors, each non-zero term in the expansion of the right hand side of (III.1) when keeping terms of the form $c_k+\varepsilon w_k$ together must have at least d terms involving $\mathrm{e}^{\mathbf{i}\mathbf{n}\cdot\mathbf{q}}-1$, so the terms of lowest order in $\bar{q}_r$ have order d. This is result (b).

These results enable us to write down (11.19) as the general form taken by $\det\mathbf{C}_p(\mathbf{q};\varepsilon)$. The terms lowest order in δ are obtained as products of polynomials of degree m in $\bar{q}_r$ and powers of ε, where the order of each term in δ equals $d+n_w$. Therefore, $\bar{P}_m(\bar{q}_1,\ldots,\bar{q}_d)$ multiplies ε^{d+n_w-m}. The sum over m ranges from d to $d+n_w$ because each term has order in $\bar{q}_r$ of at least d (accounting for the lower limit of the sum) and ε is never raised to a negative power (accounting for the upper limit).

III.2 Details of the homogenisation procedure

III.2.1 Projecting onto subspaces

We begin by projecting the equations of motion for the $n_\mathrm{s}d$ degrees of freedom per lattice unit cell (11.24) onto each of the three subspaces defined by the eigendecomposition of $\mathbf{C}_0^\mathrm{T}\mathbf{K}\mathbf{C}_0$ in Section 11.2.3:

$$\begin{aligned}
a^2\mathbf{P}_u(\tilde{\mathbf{M}}\partial_t^2\Phi-F) &= \mathbf{P}_u\mathbf{C}_\mathbf{x}^\mathrm{T}\mathbf{K}(\mathbf{C}_0+\varepsilon\mathbf{C}_w+\mathbf{C}_\mathbf{x})\Phi+\mathbf{P}_u\mathbf{C}_\mathbf{x}^\mathrm{T}\mathbf{K}(\varepsilon\mathbf{C}_{\mathbf{x},w}+O(\varepsilon^2))\Phi \qquad \text{(III.2a)}\\
&\quad+\mathbf{P}_u\left(\varepsilon\mathbf{C}_{\mathbf{x},w}+O(\varepsilon^2)\right)^\mathrm{T}\mathbf{K}\left(\mathbf{C}_0+\varepsilon\mathbf{C}_w+\mathbf{C}_\mathbf{x}+\varepsilon\mathbf{C}_{\mathbf{x},w}+O(\varepsilon^2)\right)\Phi\\
a^2\mathbf{P}_w(\tilde{\mathbf{M}}\partial_t^2\Phi-F) &= \mathbf{P}_w\mathbf{C}_\mathbf{x}^\mathrm{T}\mathbf{K}(\mathbf{C}_0+\varepsilon\mathbf{C}_w+\mathbf{C}_\mathbf{x})\Phi+\mathbf{P}_w\mathbf{C}_\mathbf{x}^\mathrm{T}\mathbf{K}\left(\varepsilon\mathbf{C}_{\mathbf{x},w}+O(\varepsilon^2)\right)\Phi \qquad \text{(III.2b)}\\
&\quad+\mathbf{P}_w\left(\varepsilon\mathbf{C}_{\mathbf{x},w}+O(\varepsilon^2)\right)^\mathrm{T}\mathbf{K}\left(\mathbf{C}_0+\varepsilon\mathbf{C}_w+\mathbf{C}_\mathbf{x}+\varepsilon\mathbf{C}_{\mathbf{x},w}+O(\varepsilon^2)\right)\Phi\\
&\quad-\mathbf{P}_w\varepsilon\mathbf{C}_w^\mathrm{T}\mathbf{K}(\mathbf{C}_0+\varepsilon\mathbf{C}_w+\mathbf{C}_\mathbf{x})\Phi-\mathbf{P}_w\varepsilon\mathbf{C}_w^\mathrm{T}\mathbf{K}\left(\varepsilon\mathbf{C}_{\mathbf{x},w}+O(\varepsilon^2)\right)\Phi\\
a^2\mathbf{P}_v(\tilde{\mathbf{M}}\partial_t^2\Phi-F) &= -\mathbf{P}_v(\mathbf{C}_0+\varepsilon\mathbf{C}_w)^\mathrm{T}\mathbf{K}(\mathbf{C}_0+\varepsilon\mathbf{C}_w+\mathbf{C}_\mathbf{x})\Phi \qquad \text{(III.2c)}\\
&\quad-\mathbf{P}_v(\mathbf{C}_0+\varepsilon\mathbf{C}_w)^\mathrm{T}\mathbf{K}(\varepsilon\mathbf{C}_{\mathbf{x},w}+O(\varepsilon^2))\Phi\\
&\quad+\mathbf{P}_v\mathbf{C}_\mathbf{x}^\mathrm{T}\mathbf{K}(\mathbf{C}_0+\varepsilon\mathbf{C}_w+\mathbf{C}_\mathbf{x})\Phi+\mathbf{P}_v\mathbf{C}_\mathbf{x}^\mathrm{T}\mathbf{K}(\varepsilon\mathbf{C}_{\mathbf{x},w}+O(\varepsilon^2))\Phi\\
&\quad+\mathbf{P}_v(\varepsilon\mathbf{C}_{\mathbf{x},w}+O(\varepsilon^2))^\mathrm{T}\mathbf{K}(\mathbf{C}_0+\varepsilon\mathbf{C}_w+\mathbf{C}_\mathbf{x}+\varepsilon\mathbf{C}_{\mathbf{x},w}+O(\varepsilon^2))\Phi,
\end{aligned}$$

where we drop the explicit dependence on $(\mathbf{x},t)$ for brevity, and the right hand side of (11.24) has been expanded. Using $\varepsilon = O(\delta)$ and $a = L\delta$ in (III.2c), we obtain (11.25) by neglecting terms $O(\delta^2)$ and above. To see this, we first neglect terms that are clearly $O(\delta^2)$ and above and are left with

$$\mathbf{P}_v(\mathbf{C}_0+\varepsilon\mathbf{C}_w)^{\mathrm{T}}\mathbf{K}\mathbf{C}\Phi_v+\mathbf{P}_v(\mathbf{C}_0+\varepsilon\mathbf{C}_w)^{\mathrm{T}}\mathbf{K}(\varepsilon\mathbf{C}_w+\mathbf{C}_{\mathbf{x}})\Phi \approx \mathbf{P}_v\mathbf{C}_{\mathbf{x}}^{\mathrm{T}}\mathbf{K}(\mathbf{C}_0+\varepsilon\mathbf{C}_w+\mathbf{C}_{\mathbf{x}})\Phi \tag{III.3}$$

after expanding some brackets. Next, we establish that $\Phi_v = O(\delta)$ by noting that the terms $\mathbf{P}_v(\mathbf{C}_0+\varepsilon\mathbf{C}_w)^{\mathrm{T}}\mathbf{K}(\varepsilon\mathbf{C}_w+\mathbf{C}_{\mathbf{x}})\Phi$ and $\mathbf{P}_v\mathbf{C}_{\mathbf{x}}^{\mathrm{T}}\mathbf{K}(\mathbf{C}_0+\varepsilon\mathbf{C}_w+\mathbf{C}_{\mathbf{x}})\Phi$ are at least $O(\delta)$, so the first term on the left hand side of the above equation is also $O(\delta)$. Therefore, $\mathbf{P}_v\mathbf{C}_0^{\mathrm{T}}\mathbf{K}\mathbf{C}_0\Phi_v = O(\delta)$, and since $\mathbf{P}_v\mathbf{C}_0^{\mathrm{T}}\mathbf{K}\mathbf{C}_0$ is invertible on the range space of $\mathbf{P}_v$ by construction, we conclude that $\Phi_v = O(\delta)$. Returning to (III.3), we can discard more terms now revealed to be $O(\delta^2)$ to obtain (11.25).

We then substitute (11.25) into (III.2a) and (III.2b) and neglect $O(\delta^3)$ terms, because terms of lowest order are $O(\delta^2)$. From this, we obtain, after algebraic manipulation, (11.28) in the main text. We note that the null space of $\tilde{\mathbf{K}}$ is the orthogonal complement of the null space of $\mathbf{C}_0^{\mathrm{T}}$, so the action of $\tilde{\mathbf{K}}$ involves projecting onto unit-cell periodic ($\mathbf{q}=\mathbf{0}$) states of self-stress of the geometrically unperturbed lattice.

III.2.2 Generalised elastic moduli and other terms in the continuum equations of motion

Here we express the generalised elastic moduli, inertial constants and body force/torque density in terms of lattice quantities. By substituting these expressions into the continuum equations (11.4), (11.5), we show that these equations are equivalent to (11.28), which is the continuum limit of the lattice equations of motion.

To express the generalised elastic moduli concisely, we first define some auxiliary linear maps. Let H be a linear map from the space of second-order tensors $\mathbf{A}$ to the space of bond extensions $\Xi = [e_1 \dots e_{n_\mathrm{b}}]^{\mathrm{T}}$, and $\{G_m\}_{m=1}^{n_w}$ be linear maps from vectors $\mathbf{c}$ to bond extensions Ξ. These linear maps H and G_m ($m=1,\dots,n_w$) are given by

$$H[\mathbf{A}] = \sum_{r=1}^{d}(-\mathrm{i})(\partial_{\bar{q}_r}\mathbf{C}(\mathbf{q})|_{\mathbf{q}=\mathbf{0}})D[\mathbf{A}\cdot a^{-1}\mathbf{a}_r], \quad G_m[\mathbf{c}] = \sum_{r=1}^{d}(-\mathrm{i})(\partial_{\bar{q}_r}\mathbf{C}(\mathbf{q})|_{\mathbf{q}=\mathbf{0}})w_m(a^{-1}\mathbf{a}_r\cdot\mathbf{c}), \tag{III.4}$$

where D is the linear map from vectors $\mathbf{v}$ to site displacements satisfying $D[\mathbf{e}_j] = \sqrt{n_\mathrm{s}}U_j$ for $j=1,\dots,d$, recalling that $\{\mathbf{e}_j\}_{j=1}^{d}$ is an orthonormal basis, and $\{U_j\}_{j=1}^{d}$ are the unit cell site displacements Φ for uniform translations along $\mathbf{e}_j$. Thus, we can regard D as mapping

displacement vectors $\mathbf{u}$ to the unit cell site displacements Φ obtained by translating all sites by $\mathbf{u}$. These linear maps are used to express the elastic tensors, which satisfy

$$\mathsf{C} : \mathbf{A} = H^{\mathrm{T}}[\tilde{\mathbf{K}} H[\mathbf{A}]] \tag{III.5a}$$

$$\mathfrak{B}_m \cdot \mathbf{c} = H^{\mathrm{T}}\left[\tilde{\mathbf{K}} G_m[\mathbf{c}]\right] \tag{III.5b}$$

$$\mathbf{N}_m = H^{\mathrm{T}}\left[a^{-1}\varepsilon \tilde{\mathbf{K}} \mathbf{C}_w W_m\right] \tag{III.5c}$$

$$\mathbf{M}_{km} \cdot \mathbf{c} = G_k^{\mathrm{T}}[\tilde{\mathbf{K}} G_m[\mathbf{c}]] \tag{III.5d}$$

$$\mathbf{h}_{km} = G_k^{\mathrm{T}}[\tilde{\mathbf{K}} a^{-1}\varepsilon \mathbf{C}_w W_m] \tag{III.5e}$$

$$J_{km} = (a^{-1}\varepsilon)^2 W_k^{\mathrm{T}} \mathbf{C}_w^{\mathrm{T}} \tilde{\mathbf{K}} \mathbf{C}_w W_m. \tag{III.5f}$$

for all second-order tensors $\mathbf{A}$ and vectors $\mathbf{c}$, and where $\tilde{\mathbf{K}}$ is given in (11.29). Since $|\mathbf{a}_r| = O(a)$ and $\varepsilon = O(\delta) = O(a)L^{-1}$, the expressions $a^{-1}\mathbf{a}_r$ and $a^{-1}\varepsilon$ are independent of a in the continuum limit.

The inertial terms proportional to the second time derivatives of the fields $\mathbf{u}(\mathbf{x},t)$ and $\{\varphi_m(\mathbf{x},t)\}_{m=1}^{n_w}$ contain the following constants for $m,k = 1,\ldots,n_w$:

$$\rho = \sum_{h=1}^{n_{\mathrm{s}}} M_h, \qquad \mathbf{p}_k = D^{\mathrm{T}}[\mathbf{P}_u \tilde{\mathbf{M}} W_k], \qquad \mu_{mk} = W_k^{\mathrm{T}} \tilde{\mathbf{M}} W_m, \tag{III.6}$$

where M_h is the mass density associated with the hth site. We note that $\mu_{km} = \mu_{mk}$. For the special case of all sites in the unit cell having equal masses, $\tilde{\mathbf{M}} = \mu\mathbf{I}$ and $\mathbf{p}_k = \mathbf{0}$ for all $k = 1,\ldots,n_w$, and $\mu_{km} = 0$ for $k \neq m$.

The body force/torque densities derived from the external forces $\mathbf{f}_h(\mathbf{x},t)$ applied to lattice sites are given by

$$\mathbf{f}(\mathbf{x},t) = \sum_{h=1}^{n_{\mathrm{s}}} \mathbf{f}_h(\mathbf{x},t), \qquad \tau_k(\mathbf{x},t) \;\; = W_k^{\mathrm{T}} F(\mathbf{x},t) \tag{III.7}$$

for $k = 1,\ldots,n_w$, which concludes the mapping between (11.4), (11.5) and the continuum limit (11.28).

III.2.3 The condition $\nu = 1$ is necessary to capture topological polarisation

We establish that $\varepsilon = O(\delta)$ is necessary to capture topological polarisation by showing that when $\nu \neq 1$ and $\nu \geq 0$ in $\varepsilon = O(\delta^\nu)$, the continuum equations of motion obtained from

retaining only terms lowest order in δ present in (11.24) are spatial inversion symmetric. We consider the following cases, which exhaust the possible values of ν.

1. $\nu = 0$

   We repeat the procedure used to obtain the continuum equations of motion in Appendix III.2.1, but with $\mathbf{C}_0$ replaced by $\mathbf{C}_0 + \varepsilon\mathbf{C}_w + O(\varepsilon^2)$ and Φ_v replaced by $\Phi_w + \Phi_v$. These steps are equivalent to setting $n_w = 0$, i.e. considering a lattice with no local soft modes. Thus $\Phi_w + \Phi_v$ is eliminated from the equations of motion, resulting in standard linear elasticity. This case is therefore unable to capture topological polarisation.

2. $\nu > 0$

   To account for possible scaling of $|\Phi_w|$ with δ, we write $\Phi_w = O(\delta^{\sigma-\nu})$ where $\sigma \geq \nu$, so that $\varepsilon\Phi_w = O(\delta^\sigma)$.

   (a) $0 < \sigma < 1$

   Retaining terms lowest order in δ present in (III.2c) gives

   $$\Phi_v = -\mathbf{X}\mathbf{C}_0^{\mathrm{T}}\mathbf{K}\mathbf{C}_w\varepsilon\Phi_w,$$

   where $\mathbf{X}$ is the pseudoinverse introduced in (11.26). Substituting this expression into (III.2a) and (III.2b), and retaining terms lowest order in δ yields

   $$\mathbf{P}_u\mathbf{C}_{\mathbf{x}}^{\mathrm{T}}\tilde{\mathbf{K}}\varepsilon\mathbf{C}_w\Phi_w = 0$$
   $$\mathbf{P}_w\varepsilon\mathbf{C}_w^{\mathrm{T}}\tilde{\mathbf{K}}\varepsilon\mathbf{C}_w\Phi_w = 0$$

   as the effective continuum equations of motion. These are not only spatial inversion symmetric but also do not include any dependence on derivatives of Φ_u, so they are independent of strain.

   (b) $\sigma > 1$

      i. $\sigma = \nu$

      Retaining terms lowest order in δ in (III.2c) yields

      $$\Phi_v = -\mathbf{X}\mathbf{C}_0^{\mathrm{T}}\mathbf{K}\mathbf{C}_{\mathbf{x}}(\Phi_u + \Phi_w),$$

which is substituted into (III.2a) and (III.2b) to obtain the continuum equations of motion (after discarding terms not of lowest order in δ)

$$a^2\mathbf{P}_u(\tilde{\mathbf{M}}\partial_t^2\Phi - F) = \mathbf{P}_u\mathbf{C}_\mathbf{x}^\mathrm{T}\tilde{\mathbf{K}}\mathbf{C}_\mathbf{x}(\Phi_u + \Phi_w)$$
$$a^2\mathbf{P}_w(\tilde{\mathbf{M}}\partial_t^2\Phi - F) = \mathbf{P}_w\mathbf{C}_\mathbf{x}^\mathrm{T}\tilde{\mathbf{K}}\mathbf{C}_\mathbf{x}(\Phi_u + \Phi_w).$$

These equations are spatial inversion symmetric.

ii. $\sigma > \nu$

Retaining terms lowest order in δ in (III.2c) yields

$$\Phi_v = -\mathbf{X}\mathbf{C}_0^\mathrm{T}\mathbf{K}\mathbf{C}_\mathbf{x}\Phi_u,$$

which is substituted into (III.2a) and (III.2b) to obtain the continuum equations of motion (after discarding terms not of lowest order in δ)

$$a^2\mathbf{P}_u(\tilde{\mathbf{M}}\partial_t^2\Phi - F) = \mathbf{P}_u\mathbf{C}_\mathbf{x}^\mathrm{T}\tilde{\mathbf{K}}\mathbf{C}_\mathbf{x}\Phi_u$$
$$a^2\mathbf{P}_w(\tilde{\mathbf{M}}\partial_t^2\Phi - F) = \mathbf{P}_w(\mathbf{C}_\mathbf{x} - \varepsilon\mathbf{C}_w)^\mathrm{T}\tilde{\mathbf{K}}\mathbf{C}_\mathbf{x}\Phi_u.$$

If $\nu = 1$, the equations can break spatial inversion symmetry, and are special cases of (11.28) obtained by discarding terms involving Φ_w because they are higher order in δ than Φ_u. To proceed, we rule out separately $\nu < 1$ and $\nu > 1$. If $\nu < 1$, then neglecting terms higher order in δ present in the second equation above gives

$$\mathbf{P}_w\varepsilon\mathbf{C}_w^\mathrm{T}\tilde{\mathbf{K}}\mathbf{C}_\mathbf{x}\Phi_u = 0,$$

so the continuum equations of motion are spatial inversion symmetric. If $\nu > 1$, then the same equation becomes

$$a^2\mathbf{P}_w(\tilde{\mathbf{M}}\partial_t^2\Phi - F) = \mathbf{P}_w\mathbf{C}_\mathbf{x}^\mathrm{T}\tilde{\mathbf{K}}\mathbf{C}_\mathbf{x}\Phi_u$$

after neglecting terms higher order in δ, so the continuum equations of motion are again spatial inversion symmetric.

(c) $\sigma = 1$

i. $0 < \nu < 1$

Retaining terms lowest order in δ present in (III.2c) gives (11.27), which is substituted into (III.2a) and (III.2b) to obtain the continuum equations of motion (after

discarding terms not of lowest order in δ)

$$a^2\mathbf{P}_u(\tilde{\mathbf{M}}\partial_t^2\Phi - F) = \mathbf{P}_u\mathbf{C}_{\mathbf{x}}^{\mathrm{T}}\tilde{\mathbf{K}}(\mathbf{C}_{\mathbf{x}}\Phi_u + \varepsilon\mathbf{C}_w\Phi_w)$$
$$\varepsilon\mathbf{P}_w\mathbf{C}_w^{\mathrm{T}}\tilde{\mathbf{K}}\mathbf{C}_w\Phi_w = -\mathbf{P}_w\mathbf{C}_w^{\mathrm{T}}\tilde{\mathbf{K}}\mathbf{C}_{\mathbf{x}}\Phi_u.$$

Let $\mathbf{P}_1$ be the orthogonal projection onto the orthogonal complement of the null space of $\tilde{\mathbf{K}}\mathbf{C}_w$, so that $\tilde{\mathbf{K}}\mathbf{C}_w\Phi_w = \tilde{\mathbf{K}}\mathbf{C}_w\mathbf{P}_1\Phi_w$. The restriction of $\mathbf{P}_w\mathbf{C}_w^{\mathrm{T}}\tilde{\mathbf{K}}\mathbf{C}_w\mathbf{P}_w$ to the range space of $\mathbf{P}_1$ is invertible. We recall that the range space of $\mathbf{C}_w^{\mathrm{T}}\tilde{\mathbf{K}}$ is equal to the orthogonal complement of the null space of $\tilde{\mathbf{K}}\mathbf{C}_w$ [42], so the restriction of $\mathbf{P}_w\mathbf{C}_w^{\mathrm{T}}\tilde{\mathbf{K}}\mathbf{C}_w\mathbf{P}_w$ to the range space of $\mathbf{P}_1$ is an operator on that subspace. Multiplying the second equation by the inverse of this restriction of $\mathbf{P}_w\mathbf{C}_w^{\mathrm{T}}\tilde{\mathbf{K}}\mathbf{C}_w\mathbf{P}_w$ gives

$$\mathbf{P}_1\Phi_w = -\mathbf{Y}\mathbf{P}_w\mathbf{C}_w^{\mathrm{T}}\tilde{\mathbf{K}}\mathbf{C}_{\mathbf{x}}\Phi_u,$$

where $\mathbf{Y}$ is the aforementioned inverse. This equation is substituted into the $\mathbf{P}_u$ equation of motion to eliminate Φ_w, leading to

$$a^2\mathbf{P}_u(\tilde{\mathbf{M}}\partial_t^2\Phi - F) = \mathbf{P}_u\mathbf{C}_{\mathbf{x}}^{\mathrm{T}}(\tilde{\mathbf{K}} - \tilde{\mathbf{K}}\mathbf{C}_w\mathbf{Y}\mathbf{C}_w^{\mathrm{T}}\tilde{\mathbf{K}})\mathbf{C}_{\mathbf{x}}\Phi_u.$$

Thus, we have eliminated Φ_w using similar reasoning to that used in eliminating Φ_v. The resulting continuum equations of motion are those of standard linear elasticity with modified effective spring constants. These equations of motion are spatial inversion symmetric.

ii. $\nu = 1$
The continuum equations of motion are those given in (11.28) and so are able to break spatial inversion symmetry.

We have shown that the continuum equations of motion obtained by retaining lowest order terms in δ are spatial inversion symmetric whenever $\nu \neq 1$. Therefore, $\nu = 1$ is necessary for the equations to capture topological polarisation.

III.2.4 Recovering the compatibility matrix determinant from generalised elastic moduli

We show that the generalised elastic moduli of the homogenised theory contain enough information to recover $\det\mathbf{C}_p(\mathbf{q};\varepsilon)$ to lowest order in δ. We use this to prove that $\det\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q})$ in Section 11.1.1 is proportional to the terms lowest order in δ present in $\det\mathbf{C}_p(\mathbf{q};\varepsilon)$ when

the continuum theory arises from a periodic lattice satisfying the Maxwell criterion, provided the generalised elastic moduli also satisfy the continuum Maxwell criterion.

First, we define

$$\mathbf{A}(\check{\mathbf{q}},\mathbf{q};\varepsilon) = \mathbf{C}_p(\check{\mathbf{q}};\varepsilon)^{\mathrm{T}}\mathbf{K}\mathbf{C}_p(\mathbf{q};\varepsilon) \tag{III.8}$$

and we let $\mathbf{S}_u$, $\mathbf{S}_w$ and $\mathbf{S}_v$ be matrices whose columns are given by $\{\sqrt{n_\mathrm{s}}U_n\}_{n=1}^{d}$, $\{W_m\}_{m=1}^{n_w}$, and $\{V_k\}_{k=1}^{n_\mathrm{s}d-(d+n_w)}$ respectively. These three sets of vectors are the eigenvectors of $\mathbf{C}_0^{\mathrm{T}}\mathbf{K}\mathbf{C}_0$ introduced in Section 11.2.3. Let $\mathbf{S} = [\mathbf{S}_u\,\mathbf{S}_w\,\mathbf{S}_v]$ be the matrix containing columns from $\mathbf{S}_u$, $\mathbf{S}_w$ and $\mathbf{S}_v$. We see that

$$\det\mathbf{C}_p(\check{\mathbf{q}};\varepsilon)\,\det\mathbf{C}_p(\mathbf{q};\varepsilon) \propto \det\mathbf{A}(\check{\mathbf{q}},\mathbf{q};\varepsilon) \propto \det\mathbf{S}^{\mathrm{T}}\mathbf{A}(\check{\mathbf{q}},\mathbf{q};\varepsilon)\mathbf{S}, \tag{III.9}$$

since $\det\mathbf{K} > 0$ and $\det\mathbf{S} > 0$ are constants independent of $\mathbf{q}$. We express $\det\mathbf{S}^{\mathrm{T}}\mathbf{A}(\check{\mathbf{q}},\mathbf{q};\varepsilon)\mathbf{S}$ in terms of the generalised elastic moduli. For notational convenience, we write $\mathbf{S} = [\mathbf{S}_{u,w}\,\mathbf{S}_v]$, where $\mathbf{S}_{u,w}$ combines the columns of $\mathbf{S}_x$ and $\mathbf{S}_w$. Then,

$$\begin{aligned}\det\mathbf{S}^{\mathrm{T}}\mathbf{A}(\check{\mathbf{q}},\mathbf{q};\varepsilon)\mathbf{S} &= \det\begin{bmatrix}\mathbf{S}_{u,w}^{\mathrm{T}}\mathbf{A}(\check{\mathbf{q}},\mathbf{q};\varepsilon)\mathbf{S}_{u,w} & \mathbf{S}_{u,w}^{\mathrm{T}}\mathbf{A}(\check{\mathbf{q}},\mathbf{q};\varepsilon)\mathbf{S}_v\\ \mathbf{S}_v^{\mathrm{T}}\mathbf{A}(\check{\mathbf{q}},\mathbf{q};\varepsilon)\mathbf{S}_{u,w} & \mathbf{S}_v^{\mathrm{T}}\mathbf{A}(\check{\mathbf{q}},\mathbf{q};\varepsilon)\mathbf{S}_v\end{bmatrix}\\ &= \det\mathbf{S}_v^{\mathrm{T}}\mathbf{A}(\check{\mathbf{q}},\mathbf{q};\varepsilon)\mathbf{S}_v\,\det\tilde{\mathbf{A}}(\check{\mathbf{q}},\mathbf{q};\varepsilon),\end{aligned} \tag{III.10}$$

for sufficiently small components $q_r, \check{q}_r$ of $\mathbf{q},\check{\mathbf{q}}$, and perturbation ε, where we write $\mathbf{A} = \mathbf{A}(\check{\mathbf{q}},\mathbf{q};\varepsilon)$ to define

$$\tilde{\mathbf{A}}(\check{\mathbf{q}},\mathbf{q};\varepsilon) = \mathbf{S}_{u,w}^{\mathrm{T}}(\mathbf{A} - \mathbf{A}\mathbf{S}_v(\mathbf{S}_v^{\mathrm{T}}\mathbf{A}\mathbf{S}_v)^{-1}\mathbf{S}_v^{\mathrm{T}}\mathbf{A})\mathbf{S}_{u,w},$$

and we have used the matrix identity

$$\det\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{bmatrix} = \det\mathbf{A}_{22}\,\det(\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}) \tag{III.11}$$

for invertible $\mathbf{A}_{22}$. The requirement that $\mathbf{S}_v^{\mathrm{T}}\mathbf{A}(\check{\mathbf{q}},\mathbf{q};\varepsilon)\mathbf{S}_v$ be invertible is satisfied for sufficiently small $\check{q}_r$, q_r, ε because the columns of $\mathbf{S}_v$ are eigenvectors of $\mathbf{A}(\mathbf{0},\mathbf{0};0) = \mathbf{C}_0^{\mathrm{T}}\mathbf{K}\mathbf{C}_0$ with strictly positive eigenvalues.

Using the definition of $\mathbf{A}(\check{\mathbf{q}},\mathbf{q};\varepsilon)$, algebraic manipulations give

$$\begin{aligned}\tilde{\mathbf{A}}(\check{\mathbf{q}},\mathbf{q};\varepsilon) &= \mathbf{S}_{u,w}^{\mathrm{T}}\mathbf{C}_p(\check{\mathbf{q}};\varepsilon)^{\mathrm{T}}(\tilde{\mathbf{K}} + O(\delta))\mathbf{C}_p(\mathbf{q};\varepsilon)\mathbf{S}_{u,w}\\ &= \begin{bmatrix}\mathbf{S}_x^{\mathrm{T}}\mathbf{R}(\check{\mathbf{q}},\mathbf{q};\varepsilon)\mathbf{S}_x & \mathbf{S}_x^{\mathrm{T}}\mathbf{R}(\check{\mathbf{q}},\mathbf{q};\varepsilon)\mathbf{S}_w\\ \mathbf{S}_w^{\mathrm{T}}\mathbf{R}(\check{\mathbf{q}},\mathbf{q};\varepsilon)\mathbf{S}_x & \mathbf{S}_w^{\mathrm{T}}\mathbf{R}(\check{\mathbf{q}},\mathbf{q};\varepsilon)\mathbf{S}_w\end{bmatrix},\end{aligned}$$

where $\mathbf{R}(\check{\mathbf{q}},\mathbf{q};\varepsilon) = \mathbf{C}_p(\check{\mathbf{q}};\varepsilon)^{\mathrm{T}}(\tilde{\mathbf{K}}+O(\delta))\mathbf{C}_p(\mathbf{q};\varepsilon)$. We relate $\tilde{\mathbf{A}}(\check{\mathbf{q}},\mathbf{q};\varepsilon)$ to the generalised elastic moduli using (III.4), giving

$$(\mathbf{C}_p(\mathbf{q};\varepsilon)\mathbf{S}_x)_{jk} = a(H[\mathbf{e}_k\otimes(\mathrm{i}\mathbf{q})])_j + O(\delta^2) \tag{III.12a}$$

$$(\mathbf{C}_p(\mathbf{q};\varepsilon)\mathbf{S}_w)_{jk} = a(G_k[(\mathrm{i}\mathbf{q})])_j + \varepsilon\mathbf{C}_w W_k + O(\delta^2). \tag{III.12b}$$

Since the values of $\check{\mathbf{q}}$, $\mathbf{q}$ are complex, the inner products $\cdot$ and : must be complexified so that they are antilinear in their first argument: $(c\mathbf{a})\cdot\mathbf{b} = c^*\mathbf{a}\cdot\mathbf{b}$ and $(c\mathbf{A}):\mathbf{B} = c^*\mathbf{A}:\mathbf{B}$ for $c\in\mathbb{C}$, where * denotes complex conjugation. For consistency with the complexified inner products, the tensor product is complexified such that $\mathbf{a}\otimes(c\mathbf{b}) = c^*\mathbf{a}\otimes\mathbf{b}$. The operators H and G_m introduced in (III.4) must also be complexified to be compatible with the complexified inner products. This modification involves replacing the transposes $^{\mathrm{T}}$ in (III.5) by conjugate transposes † (Hermitian transposes).

We introduce the generalised elastic moduli $\hat{\mathbf{K}}$ by computing the matrix (with respect to some fixed orthonormal basis $\{\mathbf{e}_j\}_{j=1}^d$ of d-dimensional space)

$$\mathbf{L}(\check{\mathbf{q}})^\dagger\hat{\mathbf{K}}\mathbf{L}(\mathbf{q}) = \begin{bmatrix}\mathbf{F}(\check{\mathbf{q}},\mathbf{q}) & \mathbf{G}(\mathbf{q},\check{\mathbf{q}})^\dagger\\ \mathbf{G}(\check{\mathbf{q}},\mathbf{q}) & \mathbf{H}(\check{\mathbf{q}},\mathbf{q})\end{bmatrix}, \tag{III.13}$$

where $\mathbf{L}(\mathbf{q})$ is defined in Section 11.1.1 as the map from generalised displacements Ψ to generalised strains Λ, and the blocks are given by the matrices

$$(\mathbf{F}(\check{\mathbf{q}},\mathbf{q}))_{jk} = (\mathbf{e}_j\otimes\check{\mathbf{q}}):\mathsf{C}:(\mathbf{e}_k\otimes\mathbf{q}) \tag{III.14a}$$

$$(\mathbf{G}(\check{\mathbf{q}},\mathbf{q}))_{jk} = (-\mathfrak{B}_j\cdot\check{\mathbf{q}}+\mathrm{i}\mathbf{N}_j):(\mathbf{e}_k\otimes\mathbf{q}) \tag{III.14b}$$

$$(\mathbf{H}(\check{\mathbf{q}},\mathbf{q}))_{jk} = \check{\mathbf{q}}\cdot\mathbf{M}_{jk}\cdot\mathbf{q}+\mathrm{i}(\mathbf{h}_{kj}\cdot\mathbf{q}-\check{\mathbf{q}}\cdot\mathbf{h}_{jk})+J_{jk}. \tag{III.14c}$$

Since the elements of $\mathbf{L}(\mathbf{q})$ have polynomial dependence on $\mathrm{i}\mathbf{q}$, we see that $\mathbf{L}(\mathbf{q})^{\mathrm{T}} = \mathbf{L}(-\mathbf{q}^*)^\dagger$, where the complex conjugate of a complex vector $\mathbf{q} = \mathbf{q}^R+\mathrm{i}\mathbf{q}^I$ expressed in terms of real vectors $\mathbf{q}^R,\mathbf{q}^I$ is given by $\mathbf{q}^* = \mathbf{q}^R-\mathrm{i}\mathbf{q}^I$. Similarly, since the elements of $\mathbf{C}_p(\mathbf{q};\varepsilon)$ have polynomial dependence on $\mathrm{e}^{\mathrm{i}\mathbf{n}\cdot\mathbf{q}}$ for some vectors $\mathbf{n}$, $\mathbf{C}_p(\mathbf{q};\varepsilon)^{\mathrm{T}} = \mathbf{C}_p(-\mathbf{q}^*;\varepsilon)^\dagger$.

We can now compute the following blocks in $\tilde{\mathbf{A}}(\check{\mathbf{q}},\mathbf{q};\varepsilon)$:

$$\begin{aligned}(\mathbf{S}_x^{\mathrm{T}}\mathbf{C}_p(\check{\mathbf{q}};\varepsilon)^{\mathrm{T}}\tilde{\mathbf{K}}\mathbf{C}_p(\mathbf{q};\varepsilon)\mathbf{S}_x)_{jk} &= (\mathbf{S}_x^{\mathrm{T}}\mathbf{C}_p(-\check{\mathbf{q}}^*;\varepsilon)^\dagger\tilde{\mathbf{K}}\mathbf{C}_p(\mathbf{q};\varepsilon)\mathbf{S}_x)_{jk}\\ &= a^2(\mathbf{e}_j\otimes(-\mathrm{i}\check{\mathbf{q}}^*)):\mathsf{C}:(\mathbf{e}_k\otimes(\mathrm{i}\mathbf{q}))+O(\delta^3)\\ &= -a^2(\mathbf{e}_j\otimes\check{\mathbf{q}}^*):\mathsf{C}:(\mathbf{e}_k\otimes\mathbf{q})+O(\delta^3)\\ &= a^2(\mathbf{F}(-\check{\mathbf{q}}^*,\mathbf{q}))_{jk}+O(\delta^3)\end{aligned} \tag{III.15a}$$

$$
\begin{aligned}
(\mathbf{S}_w^{\mathrm{T}}\mathbf{C}_p(\check{\mathbf{q}};\varepsilon)^{\mathrm{T}}\tilde{\mathbf{K}}\mathbf{C}_p(\mathbf{q};\varepsilon)\mathbf{S}_x)_{jk} &= (\mathbf{S}_w^{\mathrm{T}}\mathbf{C}_p(-\check{\mathbf{q}}^*;\varepsilon)^{\dagger}\tilde{\mathbf{K}}\mathbf{C}_p(\mathbf{q};\varepsilon)\mathbf{S}_x)_{jk} \\
&= a^2(\mathfrak{B}_j\cdot(-\mathrm{i}\check{\mathbf{q}}^*)+\mathbf{N}_j):(\mathbf{e}_k\otimes(\mathrm{i}\mathbf{q}))+O(\delta^3) \\
&= a^2(-\mathfrak{B}_j\cdot(-\check{\mathbf{q}}^*)+\mathrm{i}\mathbf{N}_j):(\mathbf{e}_k\otimes\mathbf{q})+O(\delta^3) \\
&= a^2(\mathbf{G}(-\check{\mathbf{q}}^*,\mathbf{q}))_{jk}+O(\delta^3)
\end{aligned}
\tag{III.15b}
$$

$$
\begin{aligned}
(\mathbf{S}_x^{\mathrm{T}}\mathbf{C}_p(\check{\mathbf{q}};\varepsilon)^{\mathrm{T}}\tilde{\mathbf{K}}\mathbf{C}_p(\mathbf{q};\varepsilon)\mathbf{S}_w)_{jk} &= (\mathbf{S}_x^{\mathrm{T}}\mathbf{C}_p(-\check{\mathbf{q}}^*;\varepsilon)^{\dagger}\tilde{\mathbf{K}}\mathbf{C}_p(\mathbf{q};\varepsilon)\mathbf{S}_w)_{jk} \\
&= a^2(\mathbf{e}_j\otimes(-\mathrm{i}\check{\mathbf{q}}^*)):(\mathfrak{B}_k\cdot(\mathrm{i}\mathbf{q})+\mathbf{N}_k)+O(\delta^3) \\
&= a^2(\mathbf{e}_j\otimes(-\check{\mathbf{q}}^*)):(-\mathfrak{B}_k\cdot\mathbf{q}+\mathrm{i}\mathbf{N}_k)+O(\delta^3) \\
&= a^2((-\mathfrak{B}_k\cdot\mathbf{q}+\mathrm{i}\mathbf{N}_k):(\mathbf{e}_j\otimes(-\check{\mathbf{q}}^*)))^*+O(\delta^3) \\
&= a^2(\mathbf{G}(\mathbf{q},-\check{\mathbf{q}}^*)^{\dagger})_{jk}+O(\delta^3)
\end{aligned}
\tag{III.15c}
$$

$$
\begin{aligned}
(\mathbf{S}_w^{\mathrm{T}}\mathbf{C}_p(\check{\mathbf{q}};\varepsilon)^{\mathrm{T}}\tilde{\mathbf{K}}\mathbf{C}_p(\mathbf{q};\varepsilon)\mathbf{S}_w)_{jk} &= (\mathbf{S}_w^{\mathrm{T}}\mathbf{C}_p(-\check{\mathbf{q}}^*;\varepsilon)^{\dagger}\tilde{\mathbf{K}}\mathbf{C}_p(\mathbf{q};\varepsilon)\mathbf{S}_w)_{jk} \\
&= a^2(-\mathrm{i}\check{\mathbf{q}}^*)\cdot\mathbf{M}_{jk}\cdot(\mathrm{i}\mathbf{q}) \\
&\qquad +(-\mathrm{i}\check{\mathbf{q}}^*)\cdot\mathbf{h}_{jk}+\mathbf{h}_{kj}\cdot(\mathrm{i}\mathbf{q})+J_{jk})+O(\delta^3) \\
&= a^2((-\check{\mathbf{q}}^*)\cdot\mathbf{M}_{jk}\cdot\mathbf{q} \\
&\qquad +\mathrm{i}(\mathbf{h}_{kj}\cdot\mathbf{q}-(-\check{\mathbf{q}}^*)\cdot\mathbf{h}_{jk})-J_{jk})+O(\delta^3) \\
&= a^2(\mathbf{H}(-\check{\mathbf{q}}^*,\mathbf{q}))_{jk}+O(\delta^3),
\end{aligned}
\tag{III.15d}
$$

where $\mathbf{F}(\check{\mathbf{q}},\mathbf{q})$, $\mathbf{G}(\check{\mathbf{q}},\mathbf{q})$ and $\mathbf{H}(\check{\mathbf{q}},\mathbf{q})$ are introduced in (III.14).

Therefore, we can use the blocks in $\tilde{\mathbf{A}}(\check{\mathbf{q}},\mathbf{q};\varepsilon)$ computed in (III.15), together with (III.13), to obtain

$$
\begin{aligned}
\tilde{\mathbf{A}}(\check{\mathbf{q}},\mathbf{q};\varepsilon) &= a^2\begin{bmatrix}\mathbf{F}(-\check{\mathbf{q}}^*,\mathbf{q}) & \mathbf{G}(\mathbf{q},-\check{\mathbf{q}}^*)^{\dagger}\\ \mathbf{G}(-\check{\mathbf{q}}^*,\mathbf{q}) & \mathbf{H}(-\check{\mathbf{q}}^*,\mathbf{q})\end{bmatrix}+O(\delta^3) \\
&= a^2\mathbf{L}(-\check{\mathbf{q}}^*)^{\dagger}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q})+O(\delta^3) \\
&= a^2\mathbf{L}(\check{\mathbf{q}})^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q})+O(\delta^3).
\end{aligned}
$$

Combining this result with (III.9), (III.10), and using

$$
\det\mathbf{S}_v^{\mathrm{T}}\mathbf{A}(\check{\mathbf{q}},\mathbf{q};\varepsilon)\mathbf{S}_v = \det\mathbf{S}_v^{\mathrm{T}}\mathbf{A}(\mathbf{0},\mathbf{0};0)\mathbf{S}_v+O(\delta),
$$

we obtain

$$
\det\mathbf{C}_p(\check{\mathbf{q}};\varepsilon)\det\mathbf{C}_p(\mathbf{q};\varepsilon)\propto a^{2(d+n_w)}\det\mathbf{L}(\check{\mathbf{q}})^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q})+O\left(\delta^{2(d+n_w)+1}\right) \tag{III.16}
$$

where the constant of proportionality is positive and we recall that $a=\delta L$.

From the definition of $\mathbf{P}_K$ in Section 11.1.1,

$$\begin{aligned}\det\mathbf{L}(\mathbf{q})^{\mathrm{T}}\hat{\mathbf{K}}\mathbf{L}(\mathbf{q}) &= \det\mathbf{L}(\mathbf{q})^{\mathrm{T}}\mathbf{P}_K\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}^{\frac{1}{2}}\mathbf{P}_K\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}^{\frac{1}{2}}\mathbf{P}_K\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}) = (\det\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}^{\frac{1}{2}}\mathbf{P}_K\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}))^2 \\ &= (\det\mathbf{P}_K^{\mathrm{T}}\hat{\mathbf{K}}^{\frac{1}{2}}\mathbf{P}_K)^2(\det\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}))^2\end{aligned} \tag{III.17}$$

when the Maxwell criterion (11.9) is satisfied, and where $\hat{\mathbf{K}}^{\frac{1}{2}}$ is the unique positive semi-definite square root of the positive semi-definite matrix $\hat{\mathbf{K}}$. Therefore, setting $\check{\mathbf{q}} = \mathbf{q}$ in (III.16) and using (III.17) gives

$$(\det\mathbf{C}_p(\mathbf{q};\varepsilon))^2 \propto (-a^2)^{d+n_w}(\det\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}))^2 + O\left(\delta^{2(d+n_w)+1}\right). \tag{III.18}$$

We have shown that the terms lowest order in δ in $\det\mathbf{C}_p(\mathbf{q};\varepsilon)$ are proportional to $\det\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q})$. Our result here is also consistent with our derivation in Appendix III.1 showing that the terms of lowest order in δ present in $\det\mathbf{C}_p(\mathbf{q};\varepsilon)$ have order $d+n_w$. Since $\det\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q})$ depends only on the generalised elastic moduli, we have shown that the generalised elastic moduli contain enough information to recover $\det\mathbf{C}_p(\mathbf{q};\varepsilon)$ to lowest order in δ, up to a constant of proportionality. This constant is irrelevant to the study of zero modes.

III.3 Counting continuum edge modes with a topological invariant

Here we derive the bulk–boundary correspondence in (12.3). Let $g_\alpha(\tau) = \det\mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}_\alpha^k(\tau))$. Since $g_\alpha(\tau)$ is a polynomial of degree $d+n_w$ in τ with α-dependent coefficients, we can write the roots of $g_\alpha(\tau)=0$ as $\tau = f_1(\alpha),\ldots,f_{d+n_w}(\alpha)$, where f_j is a continuous function. Using the theory of Puiseux series [161], the set of functions f_j is partitioned into cycles, and f_j can be expressed as a fractional power series

$$f_j(\alpha) = \sum_{m=0}^{\infty}\gamma_{k,m}\alpha^{\frac{m}{p_j}}, \tag{III.19}$$

where p_j is the number of elements in the cycle to which f_j belongs. Since $1 \le p_j \le d+n_w$, we see that

$$\lim_{\alpha\to 0^+}\frac{f_j(\alpha)}{\alpha^c} = 0 \tag{III.20}$$

whenever $f_j(0)=0$ and $0<c<\frac{1}{d+n_w}$. This result, which we use below, is the reason for the upper bound on c in (12.3). If all the functions f_j satisfying $f_j(0)=0$ are analytic

(i.e. $p_j = 1$), then $0 < c < 1$ is sufficient for (III.20) to hold for all functions f_j satisfying $f_j(0) = 0$.

Since $g_\alpha(\tau) = Z\prod_{k=1}^{d+n_w}(\tau - f_j(\alpha))$ for some constant Z, the integrand in (12.3) is

$$\frac{\frac{\mathrm{d}}{\mathrm{d}\tau}\det \mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}_\alpha^k(\tau))}{\det \mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}_\alpha^k(\tau))} = \sum_{j=1}^{d+n_w}\frac{1}{\tau - f_j(\alpha)}. \tag{III.21}$$

Considering $\alpha \in \mathbb{R}$, each root $f_j(\alpha)$ contributes to the integral (12.3) a term given by

$$\begin{aligned}\int_{-\alpha^c}^{\alpha^c}\frac{1}{\tau - f_j(\alpha)}\,\mathrm{d}\tau &= \log(\alpha^c - f_j(\alpha)) - \log(-\alpha^c - f_j(\alpha)) \\ &= \log\left|\frac{\alpha^c - f_j(\alpha)}{-\alpha^c - f_j(\alpha)}\right| + \mathrm{i}(\arg(\alpha^c - f_j(\alpha)) - \arg(-\alpha^c - f_j(\alpha))),\end{aligned} \tag{III.22}$$

where $\arg z$ is the argument of $z \in \mathbb{C}$. To evaluate this expression, we choose a branch of the natural logarithm. Since branches of the logarithm differ by integer multiples of $2\pi\mathrm{i}$, the integral is independent of our branch choice because it is a difference between logarithmic terms. We choose the principal branch of the logarithm, with branch cut along the negative real axis, because the straight line in the complex plane joining $-\alpha^c - f_j(\alpha)$ and $\alpha^c - f_j(\alpha)$ is parallel to the real axis. This line lies either in the positive or negative imaginary half-plane, since $\operatorname{Im} f_j(\alpha) \neq 0$ for $\alpha > 0$, by the assumption that there are no bulk modes. Therefore, this line does not cross the branch cut. With this branch choice, we have $\arg z \in (-\pi, \pi]$, and when $z \in \mathbb{C} \setminus (-\infty, 0]$,

$$\arg z = |\operatorname{sgn}(\operatorname{Re}(z))| \arctan\left(\frac{\operatorname{Im}(z)}{\operatorname{Re}(z)}\right) + H(-\operatorname{Re}(z))\operatorname{sgn}(\operatorname{Im}(z))\,\pi, \tag{III.23}$$

where arctan denotes the inverse tangent function,

$$\operatorname{sgn}(x) = \begin{cases} -1 & x < 0 \\ 0 & x = 0 \\ 1 & x > 0, \end{cases} \quad \text{and } H(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{2} & x = 0 \\ 1 & x > 0 \end{cases}.$$

By considering separately $f_j(0) = 0$ and $f_j(0) \neq 0$, we see that

$$\lim_{\alpha \to 0^+} \log\left|\frac{\alpha^c - f_j(\alpha)}{-\alpha^c - f_j(\alpha)}\right| = \log 1 = 0, \tag{III.24}$$

in both cases, where we have used (III.20) for the case $f_j(0) = 0$. Therefore, the real part of the integral vanishes in the limit $\alpha \to 0^+$.

To evaluate the imaginary part of the integral, we use (III.23) to obtain

$$\begin{aligned}\arg(\pm\alpha^c - f_j(\alpha)) &= |\mathrm{sgn}(\mathrm{Re}(\pm\alpha^c - f_j(\alpha)))| \arctan\left(\frac{\mathrm{Im}(\pm\alpha^c - f_j(\alpha))}{\mathrm{Re}(\pm\alpha^c - f_j(\alpha))}\right) \\ &\quad + H(-\mathrm{Re}(\pm\alpha^c - f_j(\alpha)))\,\mathrm{sgn}(\mathrm{Im}(\pm\alpha^c - f_j(\alpha)))\,\pi, \\ &= -\arctan\left(\frac{\mathrm{Im}(f_j(\alpha))}{\pm\alpha^c - \mathrm{Re}(f_j(\alpha))}\right) \\ &\quad - H(\mp\alpha^c + \mathrm{Re}(f_j(\alpha)))\,\mathrm{sgn}(\mathrm{Im}(f_j(\alpha)))\,\pi,\end{aligned} \tag{III.25}$$

for values of $\alpha > 0$ sufficiently close to zero, since $\lim_{\alpha\to 0^+} |\mathrm{sgn}(\mathrm{Re}(\pm\alpha^c - f_j(\alpha)))| = 1$.

When j corresponds to $f_j(0) \neq 0$, we consider two cases: $\mathrm{Re}(f_j(0)) \neq 0$ and $\mathrm{Re}(f_j(0)) = 0$. When $\mathrm{Re}(f_j(0)) \neq 0$,

$$\begin{aligned}&\lim_{\alpha\to 0^+} \arg(\pm\alpha^c - f_j(\alpha)) \\ &= \lim_{\alpha\to 0^+} \left[\arctan\left(\frac{\mathrm{Im}(f_j(\alpha))}{\mathrm{Re}(f_j(\alpha))}\right) - H(\mathrm{Re}(f_j(\alpha)))\,\mathrm{sgn}(\mathrm{Im}(f_j(\alpha)))\,\pi\right],\end{aligned} \tag{III.26}$$

so the limit is independent of the sign in $\pm\alpha^c$, leading to

$$\lim_{\alpha\to 0^+} (\arg(\alpha^c - f_j(\alpha)) - \arg(-\alpha^c - f_j(\alpha))) = 0. \tag{III.27}$$

When $\mathrm{Re}(f_j(0)) = 0$, (III.20) holds with $f_j(\alpha)$ replaced by $\mathrm{Re}(f_j(\alpha))$, and by assumption $\mathrm{Im}(f_j(0)) \neq 0$, so we obtain

$$\begin{aligned}&\lim_{\alpha\to 0^+} \arg(\pm\alpha^c - f_j(\alpha)) \\ &= \lim_{\alpha\to 0^+} \left[\arctan\left(\frac{\mathrm{Im}(f_j(\alpha))}{\mp\alpha^c}\right) - H(\mp 1)\,\mathrm{sgn}(\mathrm{Im}(f_j(\alpha)))\,\pi\right], \\ &= \left(\lim_{\alpha\to 0^+} \mathrm{sgn}(\mathrm{Im}(f_j(\alpha)))\right)\left(\mp\frac{1}{2} - H(\mp 1)\right)\pi.\end{aligned} \tag{III.28}$$

We again find that this case leads to (III.27) because the limit is independent of the sign in $\pm\alpha^c$. In summary, when j corresponds to $f_j(0) \neq 0$, the limit of the imaginary part of the integral in (III.22) vanishes. Therefore, in view of (III.21), roots $f_j(\alpha)$ such that $f_j(0) \neq 0$ do not contribute to the topological invariant (12.3). This proves that the topological invariant does not count edge modes that are irrelevant to the continuum limit, i.e. edge modes that do not satisfy (12.2).

Finally, when $f_j(0) = 0$, we obtain from (III.25)

$$\lim_{\alpha\to 0^+} \arg(\pm\alpha^c - f_j(\alpha)) = \lim_{\alpha\to 0^+} \left(-H(\mp 1)\,\mathrm{sgn}(\mathrm{Im}(f_j(\alpha)))\,\pi\right), \tag{III.29}$$

which leads to

$$\lim_{\alpha\to 0^+} (\arg(\alpha^c - f_j(\alpha)) - \arg(-\alpha^c - f_j(\alpha))) = \left(\lim_{\alpha\to 0^+} \mathrm{sgn}(\mathrm{Im}(f_j(\alpha)))\right)\pi. \tag{III.30}$$

Bringing all our results of this section together gives

$$\begin{aligned}
\lim_{\alpha\to 0^+} \frac{1}{\pi \mathrm{i}} \int_{-\alpha^c}^{\alpha^c} \frac{\frac{\mathrm{d}}{\mathrm{d}\tau} \det \mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q}_\alpha^k(\tau))}{\det \mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q}_\alpha^k(\tau))}\,\mathrm{d}\tau &= \lim_{\alpha\to 0^+} \frac{1}{\pi \mathrm{i}} \sum_{j=1}^{d+n_w} \int_{-\alpha^c}^{\alpha^c} \frac{1}{\tau - f_j(\alpha)}\,\mathrm{d}\tau \\
&= \frac{1}{\pi \mathrm{i}} \sum_{\substack{j=1: \\ f_j(0)=0}}^{d+n_w} \mathrm{i} \left(\lim_{\alpha\to 0^+} \mathrm{sgn}(\mathrm{Im}(f_j(\alpha)))\right)\pi \\
&= \sum_{\substack{j=1: \\ f_j(0)=0}}^{d+n_w} \left(\lim_{\alpha\to 0^+} \mathrm{sgn}(\mathrm{Im}(f_j(\alpha)))\right) = N_{+,k} - N_{-,k},
\end{aligned}$$

which proves (12.3) because $N_{+,k}$ is the number of roots with $\lim_{\alpha\to 0^+} \mathrm{sgn}(\mathrm{Im}(f_j(\alpha)) = 1$ and $N_{-,k}$ is the number of roots with $\lim_{\alpha\to 0^+} \mathrm{sgn}(\mathrm{Im}(f_j(\alpha)) = -1$, where $N_{\pm,k}$ only count roots satisfying $f_j(0) = 0$.

III.4 Guest–Hutchinson modes, Poisson ratios and Δ

Guest–Hutchinson modes of two-dimensional Maxwell lattices as studied in Lubensky et al. [13], Mao and Lubensky [14], and Rocklin et al. [159] are defined using standard linear elasticity, as strains in the null space of the elasticity tensor. To define Guest–Hutchinson modes in our generalised elasticity theory, we first obtain an effective standard linear-elastic continuum from our augmented theory. We then prove that $\Delta > 0$ is equivalent to the Guest–Hutchinson mode being shear-dominant and that $\Delta < 0$ is equivalent to the Guest–Hutchinson mode being dilation-dominant. We also show that in two-dimensional elasticity with one Guest–Hutchinson mode, the Poisson ratio with respect to the principal strain directions of the Guest–Hutchinson mode is positive when the mode is shear-dominant, and negative when the mode is dilation-dominant.

We consider spatially uniform generalised strains Λ, which implies that $\nabla\varphi_m = \mathbf{0}$ because the strain measures φ_m are constants. For a fixed linearised strain $\nabla^s\mathbf{u}$, we solve for the

values of $\varphi_1, \ldots, \varphi_{n_w}$ that minimise the elastic energy density. Physically, we are imposing a spatially uniform strain $\nabla^s \mathbf{u}$ and allowing the the other kinematic fields to relax to a spatially uniform equilibrium. Using (11.2) and the elastic moduli in (11.5), the elastic energy density in this spatially uniform state is given by

$$W = \frac{1}{2}\left(\nabla^s \mathbf{u} : \mathsf{C} : \nabla^s \mathbf{u} + 2A^{\mathrm{T}}\Lambda_\varphi + \Lambda_\varphi^{\mathrm{T}}\mathbf{J}\Lambda_\varphi\right), \tag{III.31}$$

where $A \in \mathbb{R}^{n_w}$ has components $A_m = \mathbf{N}_m : \nabla^s \mathbf{u}$, the vector $\Lambda_\varphi = [\varphi_1 \ldots \varphi_{n_w}]^{\mathrm{T}}$ contains the additional fields, and $\mathbf{J}$ is the symmetric $n_w \times n_w$ matrix with (k,m)-component equal to the generalised elastic modulus J_{km}. Since the matrix of generalised elastic moduli $\hat{\mathbf{K}}$ is positive semi-definite, $\mathbf{J}$ is also positive semi-definite, implying that W is a convex function of Λ_φ. Thus, stationary points of W correspond to global minima. Differentiating W with respect to Λ_φ yields the stationary point condition

$$\mathbf{J}\Lambda_\varphi = -A. \tag{III.32}$$

The minimum elastic energy density W for given $\nabla^s \mathbf{u}$ is therefore

$$W = \frac{1}{2}\left(\nabla^s \mathbf{u} : \mathsf{C} : \nabla^s \mathbf{u} - A^{\mathrm{T}}\mathbf{J}^{-1}A\right) = \frac{1}{2}\nabla^s \mathbf{u} : \left(\mathsf{C} - \sum_{k,m=1}^{n_w} (\mathbf{J}^{-1})_{km}\mathbf{N}_k \otimes \mathbf{N}_m\right) : \nabla^s \mathbf{u}, \tag{III.33}$$

where $\mathbf{J}^{-1}$ is the inverse of $\mathbf{J}$ if it exists, and its Moore–Penrose pseudoinverse if not. When $\mathbf{J}$ is not invertible, (III.32) does not have unique solutions Λ_φ, but since any stationary point of W is a *global* minimum, the minimum W is independent of the choice of Λ_φ satisfying (III.32) and is given by (III.33).

We can therefore define an effective elasticity tensor

$$\mathsf{C}_{\mathrm{eff}} = \mathsf{C} - \sum_{k,m=1}^{n_w} (\mathbf{J}^{-1})_{km}\mathbf{N}_k \otimes \mathbf{N}_m. \tag{III.34}$$

Thus, when we impose a spatially uniform strain $\nabla^s \mathbf{u}$ and allow the system to relax to a spatially uniform equilibrium state, the system has elastic energy density equal to that of a standard linear continuum with elasticity tensor $\mathsf{C}_{\mathrm{eff}}$. The Guest–Hutchinson modes in our generalised elasticity theory are defined to be strains in the null space of $\mathsf{C}_{\mathrm{eff}}$. Physically, these Guest–Hutchinson modes are uniform zero-energy deformations when the additional fields $\varphi_1, \ldots \varphi_{n_w}$ are free to relax to spatially uniform equilibrium values.

Since Δ is defined in terms of the coefficients of the $\mathbf{q}$-expansion of $\det \mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q})$, we need to relate $\mathsf{C}_{\mathrm{eff}}$ to this expansion in order to link Δ to Guest–Hutchinson modes in our theory.

We take the determinant of the matrix in (III.13) and use the matrix identity (III.11) to obtain

$$\det \mathbf{L}(\check{\mathbf{q}})^\dagger \hat{\mathbf{K}} \mathbf{L}(\mathbf{q}) = \det \mathbf{H}(\check{\mathbf{q}}, \mathbf{q}) \det(\mathbf{F}(\check{\mathbf{q}}, \mathbf{q}) - \mathbf{G}(-\mathbf{q}^*, \check{\mathbf{q}})^\dagger \mathbf{H}(\check{\mathbf{q}}, \mathbf{q})^{-1} \mathbf{G}(\check{\mathbf{q}}, \mathbf{q})) \tag{III.35}$$

when $\mathbf{H}(\check{\mathbf{q}}, \mathbf{q})$ is invertible. If $\mathbf{J}$ is invertible, then (III.14c) implies that $\mathbf{H}(\check{\mathbf{q}}, \mathbf{q})$ is invertible for sufficiently small components $q_j, \check{q}_j$. Setting $\check{\mathbf{q}} = \mathbf{q}$, and using (III.14), the terms lowest order in q_j present in $\mathbf{F}(\mathbf{q}, \mathbf{q}) - \mathbf{G}(-\mathbf{q}^*, \mathbf{q})^\dagger \mathbf{H}(\mathbf{q}, \mathbf{q})^{-1} \mathbf{G}(\mathbf{q}, \mathbf{q})$ are proportional to the $d \times d$ matrix $\mathbf{Z}(\mathbf{q})$ whose elements are given by

$$(\mathbf{Z}(\mathbf{q}))_{jk} = (\mathbf{e}_j \otimes \mathbf{q}) : \mathsf{C}_{\text{eff}} : (\mathbf{e}_k \otimes \mathbf{q}). \tag{III.36}$$

Therefore, the terms lowest order in q_j present in $\det \mathbf{L}(\mathbf{q})^\dagger \hat{\mathbf{K}} \mathbf{L}(\mathbf{q})$ are proportional to $\det \mathbf{Z}(\mathbf{q})$. Since $|\det \mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q})|^2 \propto \det \mathbf{L}(\mathbf{q})^\dagger \hat{\mathbf{K}} \mathbf{L}(\mathbf{q})$ and the terms lowest order in q_j present in $|\det \mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q})|^2$ are proportional to $|P_d(\mathbf{q})|^2$ in (12.1), we see that

$$|P_d(\mathbf{q})|^2 \propto \det \mathbf{Z}(\mathbf{q}). \tag{III.37}$$

The results we have derived so far in this section apply in d spatial dimensions. Because Δ is defined for $d = 2$, we consider only this case here. Since $\mathbf{Z}(\mathbf{q})$ is a 2×2 matrix with elements quadratic in q_j, it is a homogeneous polynomial of degree 4 in q_j. We conclude that $\det \mathbf{Z}(\mathbf{q})$ is proportional to $|P_2(\mathbf{q})|^2$ in (12.1). From this, we show that when $\Delta \neq 0$, there is exactly one Guest–Hutchinson mode. To begin, $\Delta \neq 0$ implies that $P_2(\mathbf{q})$ is not identically zero. Since C_{eff} has both major and minor symmetry, it is a symmetric linear operator on the space of symmetric second-order tensors, which is three-dimensional when $d = 2$. Clearly, $\mathsf{C}_{\text{eff}} \neq 0$ because $P_2(\mathbf{q})$ is not identically zero. Therefore, the number of non-zero eigenvalues of C_{eff} is at least 1 and at most 3. Suppose that there are two Guest–Hutchinson modes. Then by the rank–nullity theorem, C_{eff} has $3 - 2 = 1$ non-zero eigenvalues and takes the form $\mathsf{C}_{\text{eff}} = \lambda \mathbf{E} \otimes \mathbf{E}$ for $\lambda > 0$ and some symmetric second-order tensor $\mathbf{E}$. Then, using $(\mathbf{a} \otimes \mathbf{b}) : \mathbf{E} = \mathbf{a} \cdot (\mathbf{E} \cdot \mathbf{b})$,

$$\begin{aligned}(\mathbf{Z}(\mathbf{q}))_{jk} &= (\mathbf{e}_j \otimes \mathbf{q}) : \lambda \mathbf{E} \otimes \mathbf{E} : (\mathbf{e}_k \otimes \mathbf{q}) = \lambda (\mathbf{e}_j \cdot (\mathbf{E} \cdot \mathbf{q}))((\mathbf{e}_k \cdot (\mathbf{E} \cdot \mathbf{q}))^* \\ &= \lambda (\mathbf{e}_j \cdot (\mathbf{E} \cdot \mathbf{q}))((\mathbf{E} \cdot \mathbf{q}) \cdot \mathbf{e}_k) = \mathbf{e}_j \cdot (\lambda (\mathbf{E} \cdot \mathbf{q}) \otimes (\mathbf{E} \cdot \mathbf{q})) \cdot \mathbf{e}_k.\end{aligned}$$

So, $\det \mathbf{Z}(\mathbf{q}) = \det(\lambda (\mathbf{E} \cdot \mathbf{q}) \otimes (\mathbf{E} \cdot \mathbf{q})) = 0$ identically, because $(\mathbf{E} \cdot \mathbf{q}) \otimes (\mathbf{E} \cdot \mathbf{q})$ has a non-trivial null space equal to the set of vectors orthogonal to $\mathbf{E} \cdot \mathbf{q}$. This implies that $P_2(\mathbf{q})$ is identically zero, which is a contradiction. Therefore, $\Delta \neq 0$ implies that there is at most one Guest–Hutchinson mode.

Finally, we show that $\Delta > 0$ and $\Delta < 0$ are equivalent to shear-dominant and dilation-dominant Guest–Hutchinson modes, respectively. We recall that a Guest–Hutchinson mode $\mathbf{E}_{\mathrm{GH}}$ in two dimensions is shear-dominant when $\det \mathbf{E}_{\mathrm{GH}} < 0$ and dilation-dominant when $\det \mathbf{E}_{\mathrm{GH}} > 0$, corresponding to principal strains of opposite signs and of the same sign, respectively [14, 159]. Since all the expressions we work with are homogeneous polynomials in q_1, q_2, we can set $q_1 = 1$ so that solving for q_2 is equivalent to solving for the ratio q_2/q_1, which determines the direction of $\mathbf{q}$ in 2D. Thus, we set $\mathbf{q} = \mathbf{e}_1 + q_2\mathbf{e}_2$, where $q_2 \in \mathbb{C}$. So q_2 satisfies $P_2(\mathbf{q}) = 0$ if and only if $\det \mathbf{Z}(\mathbf{q}) = 0$. When this last equality holds, there exists a vector $\mathbf{d}$ whose components d_k satisfy

$$(\mathbf{Z}(\mathbf{q}))_{jk} d_k = 0,$$

where we can scale $\mathbf{d}$ so that $d_1 = 1$. Since $\mathsf{C}_{\mathrm{eff}}$ is symmetric, it can be diagonalised to obtain

$$\mathsf{C}_{\mathrm{eff}} = \lambda_1 \mathbf{E}_1 \otimes \mathbf{E}_1 + \lambda_2 \mathbf{E}_2 \otimes \mathbf{E}_2, \tag{III.38}$$

where $\lambda_1, \lambda_2 > 0$ are the two non-zero eigenvalues of $\mathsf{C}_{\mathrm{eff}}$ and $\mathbf{E}_1, \mathbf{E}_2$ are eigenstrains orthogonal to the Guest–Hutchinson mode. We find that

$$\begin{aligned} 0 = d_j^*(\mathbf{Z}(\mathbf{q}))_{jk} d_k = d_j^*(\mathbf{e}_j \otimes \mathbf{q}) : \mathsf{C}_{\mathrm{eff}} : (\mathbf{e}_k \otimes \mathbf{q})\, d_k = (\mathbf{d} \otimes \mathbf{q}) : \mathsf{C}_{\mathrm{eff}} : (\mathbf{d} \otimes \mathbf{q}) \\ = \sum_{j=1}^{2} \lambda_j ((\mathbf{d} \otimes \mathbf{q}) : \mathbf{E}_j)(\mathbf{E}_j : (\mathbf{d} \otimes \mathbf{q})) = \lambda_1 |\mathbf{E}_1 : (\mathbf{d} \otimes \mathbf{q})|^2 + \lambda_2 |\mathbf{E}_2 : (\mathbf{d} \otimes \mathbf{q})|^2, \end{aligned} \tag{III.39}$$

from which we conclude that $\mathbf{E}_j : (\mathbf{d} \otimes \mathbf{q}) = 0$ for $j = 1, 2$. Thus, $\mathsf{C}_{\mathrm{eff}} : \mathrm{sym}(\mathbf{d} \otimes \mathbf{q}) = \mathbf{0}$. We show now that $\mathbf{d}$ satisfies $P_2(\mathbf{d}^*) = 0$. Since $\mathbf{E}_j$ are symmetric second-order tensors, $\mathbf{E}_j : (\mathbf{d} \otimes \mathbf{q}) = \mathbf{E}_j : (\mathbf{q}^* \otimes \mathbf{d}^*)$. Therefore, $\mathsf{C}_{\mathrm{eff}} : (\mathbf{q}^* \otimes \mathbf{d}^*) = \mathbf{0}$, implying that

$$(\mathbf{e}_j \otimes \mathbf{d}^*) : \mathsf{C}_{\mathrm{eff}} : (\mathbf{e}_k \otimes \mathbf{d}^*) q_k^* = 0$$

for $j = 1, 2$. We see that $\det \mathbf{Z}(\mathbf{d}^*) = 0$ from (III.36), so $P_2(\mathbf{d}^*) = 0$ by (III.37).

When $\Delta > 0$, the values q_2^A, q_2^B satisfying $P_2(\mathbf{e}_1 + q_2\mathbf{e}_2) = 0$ are real and distinct, so $\mathbf{q}$ and $\mathbf{d}$ in (III.39) are vectors with real components. Therefore, $\mathrm{sym}(\mathbf{d} \otimes \mathbf{q})$ is the Guest–Hutchinson mode. We show that if $\mathbf{q} = \mathbf{e}_1 + q_2^A \mathbf{e}_2$, then $\mathbf{d} = \mathbf{e}_1 + q_2^B \mathbf{e}_2$. Suppose instead that $\mathbf{d} = \mathbf{q}$. Then, employing the reasoning in the previous paragraph applied to $\mathbf{e}_1 + q_2^B \mathbf{e}_2$, we find a Guest–Hutchinson mode corresponding to the other solution q_2^B, given by $\mathrm{sym}((\mathbf{e}_1 + q_2^B \mathbf{e}_2) \otimes \check{\mathbf{q}})$, where $\check{\mathbf{q}}$ satisfies $P_2(\check{\mathbf{q}}^*) = 0$. Therefore, we have two independent Guest–Hutchinson modes: $\mathbf{q} \otimes \mathbf{q}$, and $\mathrm{sym}((\mathbf{e}_1 + q_2^B \mathbf{e}_2) \otimes \check{\mathbf{q}})$, since $\mathbf{q} = \mathbf{e}_1 + q_2^A \mathbf{e}_2$ and $\mathbf{e}_1 + q_2^B \mathbf{e}_2$ are linearly independent.

We have arrived at a contradiction with our previously established result that there is at most one Guest–Hutchinson mode when $P_2(\mathbf{q})$ is not identically zero. Therefore, when $\Delta > 0$, the Guest–Hutchinson mode is

$$\mathbf{E}_{\mathrm{GH}} = \mathrm{sym}((\mathbf{e}_1 + q_2^A \mathbf{e}_2) \otimes (\mathbf{e}_1 + q_2^B \mathbf{e}_2)), \tag{III.40}$$

with

$$\begin{aligned} \det \mathbf{E}_{\mathrm{GH}} &= \det \begin{bmatrix} 1 & \frac{1}{2}(q_2^B + q_2^A) \\ \frac{1}{2}(q_2^A + q_2^B) & q_2^A q_2^B \end{bmatrix} = q_2^A q_2^B - \frac{1}{4}\left((q_2^A)^2 + 2q_2^A q_2^B + (q_2^A)^2\right) \\ &= -\frac{1}{4}\left(q_2^A - q_2^B\right)^2 < 0. \end{aligned} \tag{III.41}$$

Thus, $\Delta > 0$ implies that the Guest–Hutchinson mode is shear-dominant.

When $\Delta < 0$, the values q_2^A, q_2^B satisfying $P_2(\mathbf{e}_1 + q_2 \mathbf{e}_2) = 0$ are a complex conjugate pair $q_2^{A,B} = \zeta \pm \mathrm{i}\xi$ for $\zeta, \xi \in \mathbb{R}$. We let $\mathbf{q} = \mathbf{e}_1 + (\zeta + \mathrm{i}\xi)\mathbf{e}_2$, so that the other solution is $\mathbf{q}^*$. In the reasoning leading to (III.39), we showed that the Guest–Hutchinson mode $\mathbf{E}_{\mathrm{GH}}$ takes the form $\mathrm{sym}(\mathbf{q} \otimes \mathbf{d})$ for some vector $\mathbf{d}$ satisfying $P_2(\mathbf{d}^*) = 0$. The symmetrisation operation when the vectors $\mathbf{a}$ and $\mathbf{b}$ are complex is given by

$$\mathrm{sym}(\mathbf{a} \otimes \mathbf{b}) = \frac{1}{2}(\mathbf{a} \otimes \mathbf{b} + \mathbf{b}^* \otimes \mathbf{a}^*).$$

From $P_2(\mathbf{d}^*) = 0$, we see that $\mathbf{d} = \mathbf{q}$ or $\mathbf{d} = \mathbf{q}^*$. Therefore, the candidates for the Guest–Hutchinson mode $\mathbf{E}_{\mathrm{GH}}$ are $\mathrm{sym}(\mathbf{q} \otimes \mathbf{q}) = \mathrm{sym}(\mathbf{q}^* \otimes \mathbf{q}^*)$, $\mathrm{sym}(\mathbf{q} \otimes \mathbf{q}^*)$ and $\mathrm{sym}(\mathbf{q}^* \otimes \mathbf{q})$. If $\mathsf{C}_{\mathrm{eff}} : \mathbf{E} = \mathbf{0}$ and $\mathbf{E} = \mathbf{E}_R + \mathrm{i}\mathbf{E}_I$, where $\mathbf{E}_R, \mathbf{E}_I$ are real second-order tensors, then $\mathsf{C}_{\mathrm{eff}} : \mathbf{E}_R = \mathbf{0}$ and $\mathsf{C}_{\mathrm{eff}} : \mathbf{E}_I = \mathbf{0}$. Therefore, $\mathbf{E}_R$ and $\mathbf{E}_I$ are both proportional to the Guest–Hutchinson mode $\mathbf{E}_{\mathrm{GH}}$. Suppose $\mathsf{C}_{\mathrm{eff}} : \mathrm{sym}(\mathbf{q} \otimes \mathbf{q}^*)$. Expanding $\mathrm{sym}(\mathbf{q} \otimes \mathbf{q}^*)$ gives

$$\begin{aligned} \mathrm{sym}(\mathbf{q} \otimes \mathbf{q}^*) &= \frac{1}{2}(\mathbf{q} \otimes \mathbf{q}^* + (\mathbf{q}^*)^* \otimes \mathbf{q}^*) = \mathbf{q} \otimes \mathbf{q}^* \\ &= (\mathbf{e}_1 + \zeta \mathbf{e}_2) \otimes (\mathbf{e}_1 + \zeta \mathbf{e}_2) - \xi^2 \mathbf{e}_2 \otimes \mathbf{e}_2 + \mathrm{i}\xi(\mathbf{e}_2 \otimes (\mathbf{e}_1 + \zeta \mathbf{e}_2) + (\mathbf{e}_1 + \zeta \mathbf{e}_2) \otimes \mathbf{e}_2). \end{aligned} \tag{III.42}$$

The real and imaginary parts of $\mathrm{sym}(\mathbf{q} \otimes \mathbf{q}^*)$ are linearly independent second-order tensors, because the real part has a non-zero $\mathbf{e}_1 \otimes \mathbf{e}_1$ component whereas the imaginary part does not. Therefore, $\mathsf{C}_{\mathrm{eff}} : \mathrm{sym}(\mathbf{q} \otimes \mathbf{q}^*) \neq \mathbf{0}$, otherwise there would be two independent Guest–Hutchinson modes given by the real and imaginary parts of $\mathrm{sym}(\mathbf{q} \otimes \mathbf{q}^*)$. Similarly, $\mathsf{C}_{\mathrm{eff}} : \mathrm{sym}(\mathbf{q}^* \otimes \mathbf{q}) \neq \mathbf{0}$, which we can see by replacing ξ with $-\xi$ in (III.42). The last candidate

for $\mathbf{E}$ is

$$\operatorname{sym}(\mathbf{q}\otimes\mathbf{q}) = \frac{1}{2}(\mathbf{q}\otimes\mathbf{q}+\mathbf{q}^*\otimes\mathbf{q}^*) = (\mathbf{e}_1+\zeta\mathbf{e}_2)\otimes(\mathbf{e}_1+\zeta\mathbf{e}_2)+\xi^2\mathbf{e}_2\otimes\mathbf{e}_2, \tag{III.43}$$

where the imaginary parts cancel to give a purely real second-order tensor. Since there exists $\mathbf{q}$ satisfying $\det\mathbf{Z}(\mathbf{q})=0$, and (III.39) implies that for this $\mathbf{q}$ there exists a vector $\mathbf{d}$ such that $\mathsf{C}_{\text{eff}} : \operatorname{sym}(\mathbf{d}\otimes\mathbf{q}) = \mathbf{0}$, a Guest–Hutchinson mode $\mathbf{E}_{\text{GH}}$ must exist. Because we have ruled out the other candidates, $\mathbf{E}_{\text{GH}} = \operatorname{sym}(\mathbf{q}\otimes\mathbf{q})$, with

$$\det\mathbf{E}_{\text{GH}} = \det\begin{bmatrix} 1 & \zeta \\ \zeta & \zeta^2+\xi^2 \end{bmatrix} = \xi^2 > 0. \tag{III.44}$$

Thus, we have shown that $\Delta<0$ implies that the Guest–Hutchinson mode is dilation-dominant.

Bringing our results together, we see that the converses to the previous implications also hold: when $\Delta\neq 0$, the presence of a shear-dominant Guest–Hutchinson mode implies that $\Delta>0$ and the presence of a dilation-dominant Guest–Hutchinson mode implies that $\Delta<0$. Therefore, we have established that $\Delta>0$ is equivalent to the presence of a shear-dominant Guest–Hutchinson mode and $\Delta<0$ is equivalent to the presence of a dilation-dominant Guest–Hutchinson mode.

Finally, we show that a shear-dominant Guest–Hutchinson mode is equivalent to a positive Poisson ratio, and that a dilation-dominant one is equivalent to a negative Poisson ratio. This result holds for two-dimensional elasticity with exactly one Guest–Hutchinson mode. Let the principal strains (eigenvalues) of the Guest–Hutchinson mode $\mathbf{E}_{\text{GH}}$ be s_1, s_2, with corresponding principal directions (eigenvectors) $\mathbf{v}_1, \mathbf{v}_2$. The Guest–Hutchinson mode is therefore

$$\mathbf{E}_{\text{GH}} = s_1\mathbf{v}_1\otimes\mathbf{v}_1 + s_2\mathbf{v}_2\otimes\mathbf{v}_2, \tag{III.45}$$

and the orthogonality of the eigenstrains of C_{eff} in (III.38) implies that $\mathbf{E}_j : \mathbf{E}_{\text{GH}} = 0$ for $j=1,2$, which leads to

$$\mathbf{E}_j : (\mathbf{v}_1\otimes\mathbf{v}_1) = -\frac{s_2}{s_1}\mathbf{E}_j : (\mathbf{v}_2\otimes\mathbf{v}_2) \tag{III.46}$$

for $j=1,2$, where we assumed that $\det\mathbf{E}_{\text{GH}} = s_1 s_2 \neq 0$. To define the relevant Poisson ratio, we consider imposing a normal strain ε_{11} along the $\mathbf{v}_1$-direction, with stress-free material surfaces normal to the $\mathbf{v}_2$-direction, i.e. the applied stress has zero $\mathbf{v}_2\otimes\mathbf{v}_2$-component. The Poisson ratio relative to these directions is $\nu_{21} = -\frac{\varepsilon_{22}}{\varepsilon_{11}}$, where ε_{22} is the transverse strain, perpendicular to the direction of imposed stretch. The strain is $\nabla^{\text{s}}\mathbf{u} = \varepsilon_{11}\mathbf{v}_1\otimes\mathbf{v}_1 + \varepsilon_{22}\mathbf{v}_2\otimes\mathbf{v}_2$,

and the stress-free boundary condition implies that $(\mathbf{v}_2 \otimes \mathbf{v}_2) : (\mathsf{C}_{\text{eff}} : \nabla^{\text{s}}\mathbf{u}) = 0$, which gives the Poisson ratio

$$\nu_{21} = \frac{\lambda_1(\mathbf{E}_1 : (\mathbf{v}_1 \otimes \mathbf{v}_1))(\mathbf{E}_1 : (\mathbf{v}_2 \otimes \mathbf{v}_2)) + \lambda_2(\mathbf{E}_1 : (\mathbf{v}_1 \otimes \mathbf{v}_1))(\mathbf{E}_2 : (\mathbf{v}_2 \otimes \mathbf{v}_2))}{\lambda_1(\mathbf{E}_1 : (\mathbf{v}_2 \otimes \mathbf{v}_2))^2 + \lambda_2(\mathbf{E}_2 : (\mathbf{v}_2 \otimes \mathbf{v}_2))^2} = -\frac{s_2}{s_1},$$

where the last equality is obtained using (III.46). Using $\det \mathbf{E}_{\text{GH}} = s_1 s_2$ completes the proof.

III.5 Analyticity of q_2 as a function of q_1 corresponding to continuum zero modes

Our goal in this section is to prove that when solutions to $\det \mathbf{P}_K^{\text{T}} \mathbf{L}(q_1 \mathbf{e}_1 + q_2 \mathbf{e}_2) = 0$ are expressed in the form $q_2 = F(q_1)$, where $\lim_{q_1 \to 0} F(q_1) = 0$, i.e. the zero mode satisfies (12.2), F is an analytic function on some neighbourhood of zero.

In this derivation, we assume $A_{0,2} \neq 0$, which is valid as long as the unit vector $\mathbf{e}_2$ corresponding to q_2 is not pointing along a soft direction (as defined in Section 12.1.2). In this case, as we showed in the main text, there are two continuum zero modes, corresponding to functions F_1, F_2 such that the wavevector components of each zero mode are given by $q_2 = F_k(q_1)$, $k = 1, 2$. We apply the theory of Puiseux series [161]: F_k can always be written as a Puiseux series (fractional power series)

$$F_k(q_1) = \sum_{m=1}^{\infty} \gamma_{k,m} q_1^{\frac{m}{m_k}} \tag{III.47}$$

for sufficiently small $|q_1|$, where m_k is some positive integer. To prove that F_k is analytic, we need to show that the lowest order non-zero term in this expansion is proportional to q_1^a where $a \geq 1$. In the theory of Puiseux series, the set of functions F_k with the same value of $F_k(0)$ is partitioned into cycles, and m_k is the number of elements in the cycle to which F_k belongs. Therefore, since there are only two functions F_k representing solutions to $\det \mathbf{P}_K^{\text{T}} \mathbf{L}(\mathbf{q}) = 0$ that satisfy $F_k(0) = 0$, we have $m_k = 1$ or $m_k = 2$. We show that F_k is analytic by contradiction. Suppose F_k is not analytic. Then, $m_k = 2$ because $m_k = 1$ implies that F_k is analytic. Therefore, F_k belongs to a cycle of two elements, implying that this cycle contains both F_1 and F_2. Since at least one of the F_k is not analytic, $\gamma_{k,1} \neq 0$ for some $k \in \{1, 2\}$. Otherwise, the lowest order non-zero term in the Puiseux expansion of F_k would be proportional to q_1^a where $a \geq 1$, implying that F_k is analytic. The coefficients $\gamma_{1,1}$ and $\gamma_{2,1}$ are related by $\gamma_{1,1} = -\gamma_{2,1}$ because F_1 and F_2 belong to the same cycle (in general, coefficients of Puiseux series belonging to the same cycle of p elements are related by the pth

roots of unity) [161]. Therefore, both F_1 and F_2 are not analytic and the lowest order non-zero terms in their Puiseux series are proportional to $q_1^{\frac{1}{2}}$. The product of roots of $\det \mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q})$, taken as a polynomial in q_2 with q_1-dependent coefficients, has lowest order non-zero term proportional to q_1, which is the product of the lowest order non-zero terms in F_1 and F_2. This is the case because the other functions corresponding to solutions of $\det \mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}) = 0$ have non-zero constant terms in their expansions. However, the lowest order term possible in the product of roots is proportional to $A_{2,0}q_1^2$, leading to a contradiction. We therefore conclude that both F_1 and F_2 are analytic whenever the unit vector $\mathbf{e}_2$ corresponding to q_2 is not pointing along a soft direction.

III.6 Orthogonal soft directions

Here we consider the case of orthogonal soft directions and show that the results previously derived for the non-orthogonal case also apply here. We demonstrate that the wavevector of the floppy mode associated with a soft direction $\mathbf{e}_1$ can be parametrised as $\mathbf{q} = q_1\mathbf{e}_1 + F(q_1)\mathbf{e}_2$, where F is an analytic function. When soft directions are orthogonal, choosing an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ such that $\mathbf{e}_1$ is parallel to a soft direction necessarily means that $\mathbf{e}_2$ is parallel to the other soft direction. Therefore, the coefficients $A_{j,k}$ of the polynomials $P_m(q_1\mathbf{e}_1 + q_2\mathbf{e}_2) = \sum_{j=1}^{m} A_{j,m-j} q_1^j q_2^{m-j}$ in (12.1) satisfy $A_{2,0} = A_{0,2} = 0$. In particular $A_{0,2} = 0$ means that the results proven in Appendix III.5 do not apply in this case. However, when $1 \leq n_w \leq 2$, we can still show that there is at least one solution $\mathbf{q}$, with $q_2 = F(q_1)$, to $\det \mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}) = 0$ such that F is analytic and satisfies $F(0) = 0$, $F'(0) = 0$. The last property concerning $F'(0)$ implies that this floppy mode has wavevector $\mathbf{q} = q_1\mathbf{e}_1 + \mathrm{i}\frac{1}{2}\mathrm{Im}(F''(0))q_1^2\mathbf{e}_2 + O(q_1^3)$. Therefore, this floppy mode can be associated with the soft direction parallel to $\mathbf{e}_1$, as explained following (12.11). This association is sufficient for the results proven in Appendix III.7 to hold, in particular (12.13). Thus, the characterisation of topological polarisation in Section 12.1.2 holds when the soft directions are orthogonal, if $1 \leq n_w \leq 2$. This is a sufficient, not necessary condition. Weaker conditions are possible but we give this condition because it is easiest to state and covers the cases in which two-dimensional elasticity can exhibit topological polarisation and Weyl modes.

We have assumed that there are no bulk floppy modes, so there are no solutions $\mathbf{q}$ with purely real components to $\det \mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q}) = 0$. Therefore, there exists at least one $3 \leq m' \leq 2 + n_w$ such that $A_{0,m'} \neq 0$, otherwise $q_1 = 0, q_2 \neq 0$ would be a bulk zero mode. The lower bound on m' comes from $A_{0,2} = 0$. Let m be the least such m'. Treating $\det \mathbf{P}_K^{\mathrm{T}}\mathbf{L}(\mathbf{q})$ as a polynomial in q_2 with q_1-dependent coefficients, we see that there are m roots in q_2 that tend to zero as $q_1 \to 0$. Similarly, there is at least one $3 \leq s' \leq 2 + n_w$ such that $A_{s',0} \neq 0$, otherwise

$q_1 \neq 0, q_2 = 0$ would be a bulk floppy mode. The lower bound on s' comes from $A_{2,0} = 0$. Let s be the least such s'. If $1 \leq n_w \leq 2$, then $m - 1 \leq 3$. Suppose (for contradiction) that none of the functions F_j corresponding to zero modes are analytic. Let $p > 1$ be the number of Puiseux cycles among the roots (in q_2) of $\det \mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q}) = 0$, where m_k is the number of elements in the kth cycle, $1 \leq k \leq p$. If none of the roots correspond to analytic functions, then the lowest order terms in q_1 present in the Puiseux expansions in the kth cycle have q_1 raised to $\frac{m_k - 1}{m_k}$ at most (cf. Appendix III.5). Therefore, the lowest order term in q_1 present in the product of roots has order of at most $\sum_{k=1}^{p} \frac{m_k - 1}{m_k} m_k = m - p$, since $\sum_{k=1}^{p} m_k = m$. However, $A_{s,0} \neq 0$, so the the lowest order term in the product of roots has order of $s \geq 3$ in q_1. But the hypothesis $1 \leq n_w \leq 2$ implies that this lowest order term has order $m - p < m - 1 \leq 3$, which is a contradiction. Therefore, at least one root in q_2 corresponds to an analytic function F. The value of $F'(0)$ is determined by the equation $P_2(1, F'(0)) = A_{1,1} F'(0) = 0$ (cf. the discussion after (12.7)), which gives the stated result $F'(0) = 0$. Thus, we have shown that when the soft directions are orthogonal and $\mathbf{e}_1$ is parallel to a soft direction, the floppy mode associated with that soft direction corresponds to an analytic function.

III.7 Characterising edge floppy modes in terms of polarisation directions

Here, we derive the expression for the floppy mode wavevector (12.13) in a two-dimensional system. Let $\{\mathbf{e}_1^S, \mathbf{e}_2^S\}$ be an orthonormal basis with $\mathbf{e}_1^S$ parallel to a soft direction, and $\{\mathbf{e}_1, \mathbf{e}_2\}$ be an orthonormal basis obtained by rotating $\{\mathbf{e}_1^S, \mathbf{e}_2^S\}$ anticlockwise through an angle θ. We showed in Section 12.1.2 that there is a floppy mode associated with the soft direction parallel to $\mathbf{e}_1^S$, with wavevector $\mathbf{q} = q_1^S \mathbf{e}_1 + F_S(q_1^S)\mathbf{e}_2$. The function $F_S(q_1^S)$ is analytic, as proven in Appendix III.5, and satisfies $F_S(0) = 0$ and $F_S'(0) = 0$. We express the wavevector of this floppy mode in the rotated basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ as $\mathbf{q} = q_1 \mathbf{e}_1 + F_\theta(q_1)\mathbf{e}_2$, where F_θ is a continuous function satisfying $\det \mathbf{P}_K^{\mathrm{T}} \mathbf{L}(q_1 \mathbf{e}_1 + F_\theta(q_1)\mathbf{e}_2) = 0$ for q_1 in some neighbourhood of zero. The components of the floppy mode wavevector $\mathbf{q}$ in these two bases are related by

$$\begin{bmatrix} q_1 \\ F_\theta(q_1) \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} q_1^S \\ F_S(q_1^S) \end{bmatrix}, \tag{III.48}$$

so by substituting the expression for q_1 from the first row into the second row, we obtain

$$F_\theta(q_1^S \cos\theta + F_S(q_1^S)\sin\theta) = -q_1^S \sin\theta + F_S(q_1^S)\cos\theta. \tag{III.49}$$

To show that F_θ is analytic, let $h_\theta(q_1^S) = q_1^S \cos\theta + F_S(q_1^S)\sin\theta$. Then the preceding equation shows that $F_\theta \circ h_\theta$ is an analytic function, since the right-hand side of (III.49) is analytic in q_1^S by the analyticity of F_S. Using $F_S'(0) = 0$, we see that $h_\theta'(0) = \cos\theta \neq 0$ for $\theta \neq \pm\pi/2$. By Theorem 5.7.17 in Asmar and Grafakos [154], h_θ is invertible on some neighbourhood of zero with an analytic inverse. Restricted to this neighbourhood, h_θ can be taken to be invertible with an analytic inverse for all q_1 under consideration. Then, $F_\theta = (F_\theta \circ h_\theta) \circ h_\theta^{-1}$ so F_θ is the composition of analytic functions and is therefore analytic on some neighbourhood of zero.

Differentiating (III.49) with respect to q_1^S and using the analyticity of F_S and F_θ for $\theta \neq \pm\pi/2$ gives

$$F_\theta'(q_1^S\cos\theta + F_S(q_1^S)\sin\theta)(\cos\theta + F_S'(q_1^S)\sin\theta) = -\sin\theta + F_S'(q_1^S)\cos\theta. \tag{III.50}$$

Setting $q_1^S = 0$ gives

$$F_\theta'(0) = -\tan\theta. \tag{III.51}$$

Differentiating (III.50) with respect to q_1 results in

$$\begin{aligned} F_\theta''(q_1^S\cos\theta + F_S(q_1^S)\sin\theta)(\cos\theta + F_S'(q_1^S)\sin\theta)^2 \\ + F_\theta'(q_1^S\cos\theta + F_S(q_1^S)\sin\theta)\sin\theta\, F_S''(q_1^S) = F_S''(q_1^S)\cos\theta, \end{aligned} \tag{III.52}$$

and setting $q_1^S = 0$ gives

$$F_\theta''(0) = \frac{1}{\cos^3\theta} F_S''(0). \tag{III.53}$$

Therefore, the floppy mode wavevector $\mathbf{q} = q_1^S\mathbf{e}_1^S + \frac{1}{2}F_S''(0)(q_1^S)^2\mathbf{e}_2^S + O((q_1^S)^3)$ associated with the soft direction parallel to $\mathbf{e}_1^S$ can be expressed in terms of an orthonormal basis rotated by θ as

$$\mathbf{q} = q_1(\mathbf{e}_1 - \tan\theta\,\mathbf{e}_2) + \frac{1}{2}\frac{F_S''(0)}{\cos^3\theta} q_1^2\mathbf{e}_2 + O(q_1^3). \tag{III.54}$$

We see that $\mathbf{e}_1^S = \cos\theta\,\mathbf{e}_1 - \sin\theta\,\mathbf{e}_2 = \cos\theta\,(\mathbf{e}_1 - \tan\theta\,\mathbf{e}_2)$, and $\cos\theta$ can be expressed in terms of the polarisation direction $\mathbf{p}$ associated with the soft direction parallel to $\mathbf{e}_1^S$ using $\cos\theta = \mathbf{e}_2 \cdot \mathbf{e}_2^S = -\mathrm{sgn}(\mathrm{Im}(F_S''(0)))\,\mathbf{e}_2 \cdot \mathbf{p}$ using (12.12). Therefore, (III.54) becomes (12.13).

III.8 Topological polarisation in discrete lattices

Here we connect our continuum definition of topological polarisation with the definition for discrete lattices. The polarisation lattice vector $\mathbf{r}_\mathrm{T}$ defined in previous work on discrete

lattices [12, 14] can be used to indicate whether floppy modes are symmetrically or asymmetrically localised on opposite edges of a strip, provided an appropriate gauge choice of unit cell is made. Rocklin et al. [159] use this choice to obtain $\mathbf{r}_{\mathrm{T}} = \mathbf{0}$ for a lattice with a shear-dominant Guest–Hutchinson mode, indicating that this lattice is apparently unpolarised. We show that this result is consistent with our result that a shear-dominant Guest–Hutchinson mode implies that the lattice is topologically polarised. We first illustrate how $\mathbf{r}_{\mathrm{T}} = \mathbf{0}$ is obtained, with reference to our polarisation directions $\mathbf{p}_1, \mathbf{p}_2$ introduced in Section 12.1.2. Given a lattice whose primitive lattice vectors $\{\mathbf{a}_1, \mathbf{a}_2\}$ are dual to reciprocal lattice vectors $\{\mathbf{b}_1, \mathbf{b}_2\}$ that satisfy $\mathrm{sgn}(\mathbf{b}_j \cdot \mathbf{p}_1) = -\mathrm{sgn}(\mathbf{b}_j \cdot \mathbf{p}_2)$ for $j = 1, 2$, then $n_1 = n_2 = 0$ in the definition of the polarisation lattice vector $\mathbf{r}_{\mathrm{T}} = -(n_1\mathbf{a}_1 + n_2\mathbf{a}_2)$ [12, 14, 159], for an appropriate unit cell choice. In other words, if each reciprocal lattice vector $\mathbf{b}_j$ makes an acute angle with one polarisation direction and an obtuse angle with the other, then $\mathbf{r}_{\mathrm{T}} = \mathbf{0}$ for an appropriate gauge choice of unit cell. In this case, the lattice is apparently unpolarised, because whenever edges normal to $\mathbf{b}_j$ are considered, floppy modes are symmetrically localised on opposite edges. However, $\mathbf{r}_{\mathrm{T}} = \mathbf{0}$ does not mean that the lattice is unpolarised, because we can choose a different set of primitive lattice vectors so that $\mathbf{r}_{\mathrm{T}} \neq \mathbf{0}$.

Given a set of primitive reciprocal lattice vectors $\{\mathbf{b}_1, \mathbf{b}_2\}$ and a lattice with a shear-dominant Guest–Hutchinson mode, we construct a new set $\{\tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2\}$ of reciprocal lattice vectors such that at least one vector makes an acute angle with both polarisation directions. Whenever the Guest–Hutchinson mode of the lattice is shear-dominant, we can choose a direction $\hat{\mathbf{n}}$ that makes an acute angle with both polarisation directions $\mathbf{p}_1, \mathbf{p}_2$. We take an integer linear combination of the reciprocal lattice vectors $m_1\mathbf{b}_1 + m_2\mathbf{b}_2$, $m_1, m_2 \in \mathbb{Z}$, that is sufficiently aligned with $\hat{\mathbf{n}}$ to make an acute angle with both polarisation directions. Let $\tilde{b}_{1,j} = m_j/\gcd(m_1, m_2)$ for $j = 1, 2$, where $\gcd(m_1, m_2)$ is the greatest common divisor of (m_1, m_2). Therefore, $\gcd(\tilde{b}_{1,1}, \tilde{b}_{1,2}) = 1$. We set

$$\tilde{\mathbf{b}}_1 = \tilde{b}_{1,1}\mathbf{b}_1 + \tilde{b}_{1,2}\mathbf{b}_2. \tag{III.55}$$

By construction, $\tilde{\mathbf{b}}_1 \cdot \mathbf{p}_m > 0$ for $m = 1, 2$. We must now choose $\tilde{\mathbf{b}}_2$ such that $\{\tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2\}$ generates the same reciprocal lattice as $\{\mathbf{b}_1, \mathbf{b}_2\}$. By a result in the theory of lattices [169], $\{\tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2\}$ and $\{\mathbf{b}_1, \mathbf{b}_2\}$ generate the same lattice if and only if the (integer) coefficients $\tilde{b}_{2,1}, \tilde{b}_{2,2}$ in $\tilde{\mathbf{b}}_2 = \tilde{b}_{2,1}\mathbf{b}_1 + \tilde{b}_{2,2}\mathbf{b}_2$ satisfy

$$\det \begin{bmatrix} \tilde{b}_{1,1} & \tilde{b}_{1,2} \\ \tilde{b}_{2,1} & \tilde{b}_{2,2} \end{bmatrix} = \tilde{b}_{1,1}\tilde{b}_{2,2} - \tilde{b}_{1,2}\tilde{b}_{2,1} = \pm 1. \tag{III.56}$$

Using a result from abstract algebra (the remarks following Definition 6.8 in Fraleigh [170]) that guarantees the existence of integers r,s that satisfy $\gcd(\tilde{b}_{1,1},\tilde{b}_{1,2}) = \tilde{b}_{1,1}r + \tilde{b}_{1,2}s$, we conclude that there exist $\tilde{b}_{2,1},\tilde{b}_{2,2}$ that satisfy $\tilde{b}_{1,1}\tilde{b}_{2,2} - \tilde{b}_{1,2}\tilde{b}_{2,1} = 1$. Therefore, we have shown the existence of a new set of reciprocal lattice vectors $\{\tilde{\mathbf{b}}_1, \tilde{\mathbf{b}}_2\}$ such that $\tilde{\mathbf{b}}_1$ makes an acute angle with both polarisation directions. As a result, the polarisation lattice vector $\tilde{\mathbf{R}}_\mathrm{T}$ computed with respect to the primitive lattice vectors $\{\tilde{\mathbf{a}}_1, \tilde{\mathbf{a}}_2\}$ dual to the new set of reciprocal lattice vectors is non-zero because the winding number $\tilde{n}_1 \neq 0$ accounts for the asymmetric floppy mode localisation between the edges orthogonal to $\tilde{\mathbf{b}}_1$. Thus, in the context of discrete lattices whose edge floppy modes coincide with the continuum edge modes of (12.2), the presence of a shear-dominant Guest–Hutchinson mode implies that we can always choose a set of primitive lattice vectors such that $\mathbf{r}_\mathrm{T} \neq \mathbf{0}$, consistent with our continuum results.

III.9 Solving partial differential equations for topological floppy modes

To obtain numerical solutions for floppy modes from the the system of partial differential equations (11.5), we first compute a basis for the range space of $\hat{\mathbf{K}}$ from its eigenvectors that correspond to non-zero eigenvalues. We use the corresponding projection matrix $\mathbf{P}_K^\mathrm{T}$ and the compact representation (11.6) to obtain $\mathbf{P}_K^\mathrm{T}\hat{\mathbf{K}}\boldsymbol{\Lambda} = 0$, which is a system of $n_K = \operatorname{rank}\hat{\mathbf{K}}$ partial differential equations in $d + n_w$ dependent variables $u_1, \ldots, u_d, \varphi_1, \ldots, \varphi_{n_w}$. We see that the continuum Maxwell criterion (11.9) guarantees that the number of equations n_K equals the number of dependent variables $d + n_w$. We add a small diffusive term for stability, so the system of equations solved numerically is

$$\mathbf{P}_K^\mathrm{T}\hat{\mathbf{K}}\boldsymbol{\Lambda} + c\nabla^2\Psi = 0, \tag{III.57}$$

where $\Psi = [\mathbf{u}\ \varphi_1 \ldots \varphi_{n_w}]^\mathrm{T}$ are the dependent variables, ∇^2 is the Laplacian, and c is a constant with numerical value much smaller than the elements of $\mathbf{P}_K^\mathrm{T}\hat{\mathbf{K}}$. In our numerical computations, the elements of $\mathbf{P}_K^\mathrm{T}\hat{\mathbf{K}}$ have numerical values in the range 10^{-2} to 1, and we use $0 < c \leq 10^{-10}$. We use Mathematica to solve (III.57) and specify the Dirichlet boundary conditions below. We use free boundary conditions wherever Dirichlet conditions are not specified.

III.9.1 Edge modes

We solve (III.57) with $n_w = 1$ on the square-shaped domain $\Omega = \{(x_1, x_2) \in \mathbb{R}^2 : |x_1|, |x_2| \leq 0.5\}$, and set $\lambda = 0.5$ to impose Dirichlet boundary conditions

$$\Psi(\mathbf{x}) = \mathbf{0} \text{ on } x_1 = -0.5, \qquad \mathbf{u}(\mathbf{x}) \cdot \mathbf{e}_2 = \sin(4\pi x_1) \text{ on } |x_2| = 0.5.$$

III.9.2 Weyl modes

We solve (III.57) with $n_w = 2$ on the square-shaped domain $\Omega = \{(x_1, x_2) \in \mathbb{R}^2 : |x_1|, |x_2| \leq 1\}$, with Dirichlet boundary conditions

$$\Psi(\mathbf{x}) = \mathbf{0} \text{ on } x_1 = -1, \qquad \Psi(\mathbf{x}) = \mathrm{Re}(\Psi^{\mathrm{W}} e^{\mathrm{i}\mathbf{q}^{\mathrm{W}} \cdot \mathbf{x}}) \text{ on } x_2 = -1,$$

where $\mathbf{q}^{\mathrm{W}}$ is the Weyl wavevector whose coordinates are given by (12.18), and the Weyl mode shape Ψ^{W} satisfies $\mathbf{P}_K^{\mathrm{T}} \mathbf{L}(\mathbf{q}^{\mathrm{W}}) \Psi^{\mathrm{W}} = 0$.

References

[1] I. Tan et al. New two-parameter constitutive models for rubber-like materials: Revisiting the relationship between single chain stretch and continuum deformation. *European Journal of Mechanics - A/Solids* 108 (2024), p. 105398.

[2] I. Tan and A. Souslov. Classifying topological floppy modes in the continuum. *New Journal of Physics* 27 (2025), p. 055002.

[3] J. R. Rice. *Mechanics of solids*. *Encyclopaedia Britannica*. 2017.

[4] C. A. Truesdell. Cauchy and the modern mechanics of continua. *Revue d'histoire des sciences* 45.1 (1992), pp. 5–24.

[5] C. Truesdell and W. Noll. *The Non-Linear Field Theories of Mechanics, 3rd edition*. Springer, 2004.

[6] M. E. Gurtin. *An introduction to continuum mechanics*. Academic Press, 1982.

[7] H. Dal et al. On the Performance of Isotropic Hyperelastic Constitutive Models for Rubber-Like Materials: A State of the Art Review. *Applied Mechanics Reviews* 73.2 (2021).

[8] M. Destrade et al. Third- and fourth-order elasticities of biological soft tissues. *The Journal of the Acoustical Society of America* 127.4 (2010), pp. 2103–2106.

[9] M. E. Gurtin et al. *The mechanics and thermodynamics of continua*. Cambridge University Press, 2010.

[10] C. Scheibner et al. Odd elasticity. *Nature Physics* 16.4 (2020), pp. 475–480.

[11] A. Yavari and A. Goriely. *Nonlinear Cauchy Elasticity*. arXiv preprint. 2025.

[12] C. L. Kane and T. C. Lubensky. Topological boundary modes in isostatic lattices. *Nature Physics* 10 (2014), pp. 39–45.

[13] T. C. Lubensky et al. Phonons and elasticity in critically coordinated lattices. *Reports on Progress in Physics* 78.7 (2015), p. 073901.

[14] X. Mao and T. C. Lubensky. Maxwell Lattices and Topological Mechanics. *Annual Review of Condensed Matter Physics* 9 (2018), pp. 413–433.

[15] K. Sun and X. Mao. Continuum Theory for Topological Edge Soft Modes. *Physical Review Letters* 124.20 (2020).

[16] A. Saremi and Z. Rocklin. Topological Elasticity of Flexible Structures. *Physical Review X* 10.1 (2020).

[17] H. Nassar et al. Microtwist elasticity: A continuum approach to zero modes and topological polarization in Kagome lattices. *Journal of the Mechanics and Physics of Solids* 144 (2020), p. 104107.

[18] A. Gent. Rubber elasticity: basic concepts and behavior. *Science and technology of rubber*. Ed. by J. E. Mark et al. Elsevier, 2005.

[19] P. Steinmann et al. Hyperelastic models for rubber-like materials: consistent tangent operators and suitability for Treloar's data. *Archive of Applied Mechanics* 82.9 (2012), pp. 1183–1217.

[20] O. Yasa et al. An Overview of Soft Robotics. *Annual Review of Control, Robotics, and Autonomous Systems* 6 (2023), pp. 1–29.

[21] O. Gonzalez and A. M. Stuart. *A first course in continuum mechanics*. Cambridge University Press, 2008.

[22] L. R. G. Treloar. *The Physics of Rubber Elasticity*. Oxford University Press, 2005.

[23] M. Destrade et al. Methodical fitting for mathematical models of rubber-like materials. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 473.2198 (2017), p. 20160811.

[24] A. Anssari-Benam et al. On the central role of the invariant I2 in nonlinear elasticity. *International Journal of Engineering Science* 163 (2021), p. 103486.

[25] A. Anssari-Benam and C. O. Horgan. A three-parameter structurally motivated robust constitutive model for isotropic incompressible unfilled and filled rubber-like materials. *European Journal of Mechanics - A/Solids* 95 (2022), p. 104605.

[26] R. W. Ogden et al. Fitting hyperelastic models to experimental data. *Computational Mechanics* 34.6 (2004), pp. 484–502.

[27] A. Anssari-Benam and A. Bucchi. A generalised neo-Hookean strain energy function for application to the finite deformation of elastomers. *International Journal of Non-Linear Mechanics.* 128 (2021), p. 103626.

[28] R. S. Rivlin. Large elastic deformations of isotropic materials IV. Further developments of the general theory. *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 241 (1948), pp. 379–397.

[29] R. W. Ogden. On isotropic tensors and elastic moduli. *Mathematical Proceedings of the Cambridge Philosophical Society* 75.3 (1974), pp. 427–436.

[30] A. N. Gent. A New Constitutive Relation for Rubber. *Rubber Chemistry and Technology* 69.1 (1996), pp. 59–61.

[31] C. O. Horgan and G. Saccomandi. A Molecular-Statistical Basis for the Gent Constitutive Model of Rubber Elasticity. *Journal of Elasticity* 68.1/3 (2002), pp. 167–176.

[32] E. M. Arruda and M. C. Boyce. A three-dimensional constitutive model for the large stretch behavior of rubber elastic materials. *Journal of the Mechanics and Physics of Solids* 41.2 (1993), pp. 389–412.

[33] V. J. Amores et al. On the network orientational affinity assumption in polymers and the micro–macro connection through the chain stretch. *Journal of the Mechanics and Physics of Solids* 148 (2021), p. 104279.

[34] P. D. Wu and E. van der Giessen. On improved 3-D non-Gaussian network models for rubber elasticity. *Mechanics Research Communications* 19.5 (1992), pp. 427–433.

[35] C. Miehe et al. A micro-macro approach to rubber-like materials–Part I: the non-affine micro-sphere model of rubber elasticity. *Journal of the Mechanics and Physics of Solids* 52.11 (2004), pp. 2617–2660.

[36] H. Dal et al. An extended eight-chain model for hyperelastic and finite viscoelastic response of rubberlike materials: Theory, experiments and numerical aspects. *Journal of the Mechanics and Physics of Solids* 145 (2020), p. 104159.

[37] S. Toki et al. New Insights into Structural Development in Natural Rubber during Uniaxial Deformation by In Situ Synchrotron X-ray Diffraction. *Macromolecules* 35.17 (2002), pp. 6578–6584.

[38] S. Toki et al. Molecular orientation and structural development in vulcanized polyisoprene rubbers during uniaxial deformation by in situ synchrotron X-ray diffraction. *Polymer*. In Honour of Ian Ward's 75th Birthday 44.19 (2003), pp. 6003–6011.

[39] M. Ott et al. Microscopic Study of Chain Deformation and Orientation in Uniaxially Strained Polymer Networks: NMR Results versus Different Network Models. *Macromolecules* 47.21 (2014), pp. 7597–7611.

[40] J. H. Irving and J. G. Kirkwood. The Statistical Mechanical Theory of Transport Processes. IV. The Equations of Hydrodynamics. *Journal of Chemical Physics* 18.6 (1950), pp. 817–829.

[41] W. Noll. Die Herleitung der Grundgleichungen der Thermomechanik der Kontinua aus der statistischen Mechanik. *Journal of Rational Mechanics and Analysis* 4.5 (1955), pp. 627–646.

[42] S. Axler. *Linear Algebra Done Right*. Undergraduate Texts in Mathematics. Cham: Springer International Publishing, 2024.

[43] M. Carroll. Must Elastic Materials be Hyperelastic? *Mathematics and Mechanics of Solids* 14.4 (2009), pp. 369–376.

[44] R. W. Ogden. Large deformation isotropic elasticity–on the correlation of theory and experiment for incompressible rubberlike solids. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 326.1567 (1972), pp. 565–584.

[45] D. E. Carlson and D. A. Tortorelli. On hyperelasticity with internal constraints. *Journal of Elasticity* 42.1 (1996), pp. 91–98.

[46] P. D. Wu and E. van der Giessen. On improved network models for rubber elasticity and their applications to orientation hardening in glassy polymers. *Journal of the Mechanics and Physics of Solids* 41.3 (1993), pp. 427–456.

[47] M. Doi and S. F. Edwards. *The Theory of Polymer Dynamics*. Clarendon Press, 1988.

[48] F. Manca et al. Elasticity of flexible and semiflexible polymers with extensible bonds in the Gibbs and Helmholtz ensembles. *Journal of Chemical Physics* 136.15 (2012), p. 154906.

[49] M. Rubinstein and R. H. Colby. *Polymer Physics*. Oxford University Press, 2003.

[50] M. C. Wang and E. Guth. Statistical theory of networks of non-Gaussian flexible chains. *Journal of Chemical Physics* 20.7 (1952), pp. 1144–1157.

[51] P. J. Flory and J. Rehner Jr. Statistical Mechanics of Cross-Linked Polymer Networks I. Rubberlike Elasticity. *Journal of Chemical Physics* 11.11 (1943), pp. 512–520.

[52] P. J. Flory. Statistical thermodynamics of random networks. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 351.1666 (1976), pp. 351–380.

[53] L. R. G. Treloar. Stress-strain data for vulcanised rubber under various types of deformation. *Transactions of the Faraday Society* 40 (1944), p. 59.

[54] M. F. Beatty. An Average-Stretch Full-Network Model for Rubber Elasticity. *Journal of Elasticity* 70.1-3 (2003), pp. 65–86.

[55] L. Zhan et al. A new micro–macro transition for hyperelastic materials. *Journal of the Mechanics and Physics of Solids* 171 (2023), p. 105156.

[56] S. Kawabata et al. Experimental survey of the strain energy density function of isoprene rubber vulcanizate. *Macromolecules* 14.1 (1981), pp. 154–162.

[57] V. J. Amores et al. A model for rubber-like materials with three parameters obtained from a tensile test. *European Journal of Mechanics - A/Solids* 100 (2023), p. 104931.

[58] J. U. Sommer et al. Theoretical investigation of the segment-segment correlation in topological constrained networks. *Colloid and Polymer Science* 268.2 (1990), pp. 148–154.

[59] B. Mergell and R. Everaers. Tube Models for Rubber–Elastic Systems. *Macromolecules* 34.16 (2001), pp. 5675–5686.

[60] N. C. Admal and E. B. Tadmor. A Unified Interpretation of Stress in Molecular Systems. *Journal of Elasticity* 100.1-2 (2010), pp. 63–143.

[61] J. Diani and P. Gilormini. Combining the logarithmic strain and the full-network model for a better understanding of the hyperelastic behavior of rubber-like materials. *Journal of the Mechanics and Physics of Solids* 53.11 (2005), pp. 2579–2596.

[62] R. Akbari et al. Reverse physically motivated frameworks for investigation of strain energy function in rubber-like elasticity. *International Journal of Mechanical Sciences* 221 (2022), p. 107110.

[63] S. Abbott. *Understanding Analysis*. Springer, 2015.

[64] P. Glaister. 75.41 A "Flat" Function with Some Interesting Properties and an Application. *The Mathematical Gazette* 75.474 (1991), pp. 438–440.

[65] C. O. Horgan and J. G. Murphy. Poynting Effects in Soft Elastic Materials: A Review of Recent Results. *Journal of Elasticity* 157.2 (2025).

[66] L. Meunier et al. Mechanical experimental characterisation and numerical modelling of an unfilled silicone rubber. *Polym. Test.* 27.6 (2008), pp. 765–777.

[67] D. C. Moreira and L. C. S. Nunes. Comparison of simple and pure shear for an incompressible isotropic hyperelastic material under large deformation. *Polymer Testing* 32.2 (2013), pp. 240–248.

[68] P. A. Janmey et al. Negative normal stress in semiflexible biopolymer gels. *Nature Materials* 6 (2007), pp. 48–51.

[69] M. Kröger. Simple, admissible, and accurate approximants of the inverse Langevin and Brillouin functions, relevant for strong polymer deformations and flows. *Journal of Non-Newtonian Fluid Mechanics* 223 (2015), pp. 77–87.

[70] P. Bažant and B. H. Oh. Efficient Numerical Integration on the Surface of a Sphere. *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik* 66.1 (1986), pp. 37–49.

[71] V. I. Lebedev. Values of the nodes and weights of ninth to seventeenth order gauss-markov quadrature formulae invariant under the octahedron group with inversion. *USSR Computational Mathematics and Mathematical Physics* 15.1 (1975), pp. 44–51.

[72] V. I. Lebedev. Quadratures on a sphere. *USSR Computational Mathematics and Mathematical Physics* 16.2 (1976), pp. 10–24.

[73] V. I. Lebedev. Spherical quadrature formulas exact to orders 25–29. *Siberian Mathematical Journal* 18.1 (1977), pp. 99–107.

[74] C. Beentjes. *Quadrature on a spherical surface.* Tech. rep. Oxford University, 2015.

[75] A. Anssari-Benam et al. Assessment of a new isotropic hyperelastic constitutive model for a range of rubberlike materials and deformations. *Rubber Chemistry and Technology* 95.2 (2021), pp. 200–217.

[76] G. Munoz et al. Predicting Mechanical Constitutive Laws of Elastomers with Mesoscale Simulations. *Macromolecules* 55.5 (2022), pp. 1487–1494.

[77] M. Laczkovich and V. T. Sós. *Real Analysis: Series, Functions of Several Variables, and Applications*. Springer, 2017.

[78] E. Pucci and G. Saccomandi. A Note on the Gent Model for Rubber-Like Materials. *Rubber Chemistry and Technology* 75.5 (2002), pp. 839–852.

[79] S. Ramaswamy. The Mechanics and Statistics of Active Matter. *Annual Review of Condensed Matter Physics* 1.1 (2010), pp. 323–345.

[80] M. C. Marchetti et al. Hydrodynamics of soft active matter. *Reviews of Modern Physics* 85.3 (2013), pp. 1143–1189.

[81] K. Kruse et al. Generic theory of active polar gels: a paradigm for cytoskeletal dynamics. *The European Physical Journal E* 16.1 (2005), pp. 5–16.

[82] S. Ramaswamy et al. Active nematics on a substrate: Giant number fluctuations and long-time tails. *Europhysics Letters* 62.2 (2003), pp. 196–202.

[83] J. Toner and Y. Tu. Flocks, herds, and schools: A quantitative theory of flocking. *Physical Review E* 58.4 (1998), pp. 4828–4858.

[84] R. W. Nash et al. Run-and-Tumble Particles with Hydrodynamics: Sedimentation, Trapping, and Upstream Swimming. *Physical Review Letters* 104.25 (2010), p. 258101.

[85] H. H. Wensink et al. Meso-scale turbulence in living fluids. *Proceedings of the National Academy of Sciences* 109.36 (2012), pp. 14308–14313.

[86] E. Lauga. Bacterial Hydrodynamics. *Annual Review of Fluid Mechanics* 48.1 (2016), pp. 105–130.

[87] J. L. Moran and J. D. Posner. Phoretic Self-Propulsion. *Annual Review of Fluid Mechanics* 49 (2017), pp. 511–540.

[88] G. Tucci et al. Nonreciprocal collective dynamics in a mixture of phoretic Janus colloids. *New Journal of Physics* 26.7 (2024), p. 073006.

[89] D. Das and E. Lauga. Active Particles Powered by Quincke Rotation in a Bulk Fluid. *Physical Review Letters* 122.19 (2019), p. 194503.

[90] R. Singh and R Adhikari. Generalized Stokes laws for active colloids and their applications. *Journal of Physics Communications* 2.2 (2018), p. 025025.

[91] G. Turk et al. Stokes traction on an active particle. *Physical Review E* 106.1 (2022), p. 014601.

[92] R. Dreyfus et al. Microscopic artificial swimmers. *Nature* 437.7060 (2005), pp. 862–865.

[93] E. Han et al. Low-Reynolds-number, biflagellated Quincke swimmers with multiple forms of motion. *Proceedings of the National Academy of Sciences* 118.29 (2021), e2022000118.

[94] A. P. Petroff et al. Fast-Moving Bacteria Self-Organize into Active Two-Dimensional Crystals of Rotating Cells. *Physical Review Letters* 114.15 (2015), p. 158102.

[95] T. H. Tan et al. Odd dynamics of living chiral crystals. *Nature* 607.7918 (2022), pp. 287–293.

[96] J. Agudo-Canalejo and R. Golestanian. Active Phase Separation in Mixtures of Chemically Interacting Particles. *Physical Review Letters* 123.1 (2019), p. 018101.

[97] R. Di Leonardo et al. Bacterial ratchet motors. *Proceedings of the National Academy of Sciences* 107.21 (2010), pp. 9541–9545.

[98] P. Baconnier et al. Selective and collective actuation in active solids. *Nature Physics* 18.10 (2022), pp. 1234–1239.

[99] C. Hernández-López et al. Model of Active Solids: Rigid Body Motion and Shape-Changing Mechanisms. *Physical Review Letters* 132.23 (2024).

[100] J. Veenstra et al. Adaptive locomotion of active solids. *Nature* 639 (2025), pp. 935–941.

[101] S. Shankar and L. Mahadevan. Active hydraulics and odd elasticity of muscle fibres. *Nature Physics* 20.9 (2024), pp. 1501–1508.

[102] A. Alexanderian et al. Irreducibility of a Symmetry Group Implies Isotropy. *Journal of Elasticity* 102 (2011), pp. 151–174.

[103] L. Braverman et al. Topological Defects in Solids with Odd Elasticity. *Physical Review Letters* 127.26 (2021).

[104] K. A. Pericak-Spector et al. The Representation Theorem for Linear, Isotropic Tensor Functions in Even Dimensions. *Journal of Elasticity* 57.2 (1999), pp. 157–164.

[105] C. C. Wang. A new representation theorem for isotropic functions: An answer to Professor G. F. Smith's criticism of my papers on representations for isotropic functions. *Archive for Rational Mechanics and Analysis* 36 (1970), pp. 198–223.

[106] Z. H. Guo. An alternative proof of the representation theorem for isotropic, linear asymmetric stress-strain relations. *Quarterly of Applied Mathematics* 41.1 (1983), pp. 119–123.

[107] C. S. Jog. A Concise Proof of the Representation Theorem for Fourth-Order Isotropic Tensors. *Journal of Elasticity* 85.2 (2006), pp. 119–124.

[108] J. P. Jaric. On the Representation of Symmetric Isotropic 4-Tensors. *Journal of Elasticity* 51.1 (1998), pp. 73–79.

[109] F. D. Murnaghan. Finite Deformations of an Elastic Solid. *American Journal of Mathematics* 59.2 (1937), p. 235.

[110] L. D. Landau et al. *Theory of elasticity: volume 7*. Elsevier, 2012.

[111] G. Saccomandi and L. Vergori. Some Remarks on the Weakly Nonlinear Theory of Isotropic Elasticity. *Journal of Elasticity* 147.1-2 (2021), pp. 33–58.

[112] Y. Du et al. Connecting weakly nonlinear elasticity theories of isotropic hyperelastic materials. *Mathematics and Mechanics of Solids* 29.9 (2024), pp. 1900–1914.

[113] A. K. Subramaniyan and C. T. Sun. Continuum interpretation of virial stress in molecular simulations. *International Journal of Solids and Structures* 45.14-15 (2008), pp. 4340–4346.

[114] C. Heussinger and E. Frey. Floppy Modes and Nonaffine Deformations in Random Fiber Networks. *Physical Review Letters* 97.10 (2006).

[115] C. P. Broedersz et al. Criticality and isostaticity in fibre networks. *Nature Physics* 7.12 (2011), pp. 983–988.

[116] D. L. Humphries et al. The mechanics of phantom Mikado networks. *Journal of Physics Communications* 2.5 (2018), p. 055015.

[117] R. A. Brown et al. Tensional homeostasis in dermal fibroblasts: Mechanical responses to mechanical loading in three-dimensional substrates. *Journal of Cellular Physiology* 175.3 (1998), pp. 323–332.

[118] F. Alisafaei et al. Tension anisotropy drives fibroblast phenotypic transition by self-reinforcing cell–extracellular matrix mechanical feedback. *Nature Materials* 24 (2025), pp. 955–965.

[119] C. A. R. Jones et al. Micromechanics of cellularized biopolymer networks. *Proceedings of the National Academy of Sciences* 112.37 (2015).

[120] M. S. Hall et al. Fibrous nonlinear elasticity enables positive mechanical feedback between cells and ECMs. *Proceedings of the National Academy of Sciences* 113.49 (2016), pp. 14043–14048.

[121] T. Mizutani et al. Cellular stiffness response to external deformation: Tensional homeostasis in a single fibroblast. *Cell Motility* 59.4 (2004), pp. 242–248.

[122] M. Pósfai et al. Impact of physicality on network structure. *Nature Physics* 20 (2024), pp. 142–149.

[123] M. Destrade et al. Scalar evolution equations for shear waves in incompressible solids: a simple derivation of the Z, ZK, KZK and KP equations. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 467.2131 (2010), pp. 1823–1834.

[124] J. R. Munkres. *Analysis on Manifolds*. Westview Press, 1991.

[125] K. Bertoldi et al. Flexible mechanical metamaterials. *Nature Reviews Materials* 2.11 (2017).

[126] E. Barchiesi et al. Mechanical metamaterials: a state of the art. *Mathematics and Mechanics of Solids* 24.1 (2019), pp. 212–234.

[127] R. Lakes. Foam Structures with a Negative Poisson's Ratio. *Science* 235.4792 (1987), pp. 1038–1040.

[128] K. Bertoldi et al. Negative Poisson's Ratio Behavior Induced by an Elastic Instability. *Advanced Materials* 22.3 (2010), pp. 361–366.

[129] X. N. Liu et al. An elastic metamaterial with simultaneously negative mass density and bulk modulus. *Applied Physics Letters* 98.25 (2011).

[130] Z. G. Nicolaou and A. E. Motter. Mechanical metamaterials with negative compressibility transitions. *Nature Materials* 11.7 (2012), pp. 608–613.

[131] B. Florijn et al. Programmable Mechanical Metamaterials. *Physical Review Letters* 113.17 (2014).

[132] C. Coulais et al. Static non-reciprocity in mechanical metamaterials. *Nature* 542.7642 (2017), pp. 461–464.

[133] A. Bergne et al. Scalable 3D printing for topological mechanical metamaterials. *Extreme Mechanics Letters* 57 (2022), p. 101911.

[134] C. Widstrand et al. Stress focusing and damage protection in topological Maxwell metamaterials. *International Journal of Solids and Structures* 274 (2023), p. 112268.

[135] D. J. Thouless et al. Quantized Hall Conductance in a Two-Dimensional Periodic Potential. *Physical Review Letters* 49.6 (1982), pp. 405–408.

[136] M. Z. Hasan and C. L. Kane. Colloquium: Topological insulators. *Reviews of Modern Physics* 82.4 (2010), pp. 3045–3067.

[137] E. Prodan and C. Prodan. Topological Phonon Modes and Their Role in Dynamic Instability of Microtubules. *Physical Review Letters* 103.24 (2009).

[138] R. Süsstrunk and S. D. Huber. Observation of phononic helical edge states in a mechanical topological insulator. *Science* 349.6243 (2015), pp. 47–50.

[139] L. M. Nash et al. Topological mechanics of gyroscopic metamaterials. *Proceedings of the National Academy of Sciences* 112.47 (2015), pp. 14495–14500.

[140] O. R. Bilal et al. Intrinsically Polar Elastic Metamaterials. *Advanced Materials* 29.26 (2017).

[141] Z. Tang et al. Fully Polarized Topological Isostatic Metamaterials in Three Dimensions. *Physical Review Letters* 133.10 (2024).

[142] L. Zhang and X. Mao. Fracturing of topological Maxwell lattices. *New Journal of Physics* 20.6 (2018), p. 063034.

[143] D. Z. Rocklin et al. Mechanical Weyl Modes in Topological Maxwell Lattices. *Physical Review Letters* 116.13 (2016).

[144] O. Stenull et al. Topological Phonons and Weyl Lines in Three Dimensions. *Physical Review Letters* 117.6 (2016).

[145] D. Zhou et al. Topological Boundary Floppy Modes in Quasicrystals. *Physical Review X* 9.2 (2019).

[146] D. Zhou et al. Topological Edge Floppy Modes in Disordered Fiber Networks. *Physical Review Letters* 120.6 (2018).

[147] B. G.-g. Chen et al. Topological Mechanics of Origami and Kirigami. *Physical Review Letters* 116.13 (2016).

[148] H. Liu et al. Topological floppy modes in models of epithelial tissues. *Soft Matter* 17.38 (2021), pp. 8624–8641.

[149] A. S. Meeussen et al. Geared Topological Metamaterials with Tunable Mechanical Stability. *Physical Review X* 6.4 (2016).

[150] N. W. Ashcroft and N. D. Mermin. *Solid state physics*. Holt, Rinehart & Winston, 1976.

[151] J. Lee. *Introduction to topological manifolds*. Springer Science & Business Media, 2010.

[152] J. C. Maxwell. L. on the calculation of the equilibrium and stiffness of frames. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* (1864).

[153] C. R. Calladine. Buckminster Fuller's "Tensegrity" structures and Clerk Maxwell's rules for the construction of stiff frames. *International Journal of Solids and Structures* 14.2 (1978), pp. 161–172.

[154] N. H. Asmar and L. Grafakos. *Complex analysis with applications*. Springer, 2018.

[155] G. Baardink. "Topological Zero-Modes and States of Self-Stress in 3D Lattices". Master's thesis. Leiden University, 2017.

[156] R. Xia et al. Microtwist homogenization of three-dimensional Pyrochlore lattices on zero modes and mechanical polarization. *Journal of the Mechanics and Physics of Solids* 155 (2021), p. 104564.

[157] M. Guzman et al. Model-free characterization of topological edge and corner states in mechanical networks. *Proceedings of the National Academy of Sciences* 121.4 (2024).

[158] S. Forest and K. Sab. Finite-deformation second-order micromorphic theory and its relations to strain and stress gradient models. *Mathematics and Mechanics of Solids* 25.7 (2020), pp. 1429–1449.

[159] D. Z. Rocklin et al. Transformable topological mechanical metamaterials. *Nature Communications* 8.1 (2017).

[160] S. D Guest and J. W Hutchinson. On the determinacy of repetitive structures. *Journal of the Mechanics and Physics of Solids* 51.3 (2003), pp. 383–391.

[161] K. Knopp. *Theory of functions, Parts I and II*. Dover Publications, Inc, 1996.

[162] X. Shi et al. Elastic Weyl Points and Surface Arc States in Three-Dimensional Structures. *Physical Review Applied* 12.2 (2019).

[163] S. S. Ganti et al. Weyl points and topological surface states in a three-dimensional sandwich-type elastic lattice. *New Journal of Physics* 22.8 (2020), p. 083001.

[164] M. Xiao et al. Synthetic gauge flux and Weyl points in acoustic systems. *Nature Physics* 11.11 (2015), pp. 920–924.

[165] B. G.-g. Chen et al. Nonlinear conduction via solitons in a topological mechanical insulator. *Proceedings of the National Academy of Sciences* 111.36 (2014), pp. 13004–13009.

[166] N. Upadhyaya et al. Nuts and bolts of supersymmetry. *Physical Review Research* 2.4 (2020).

[167] N. Lera et al. Topological mechanical metamaterial with nonrectilinear constraints. *Physical Review B* 98.1 (2018).

[168] A. Rosén. *Geometric Multivector Analysis: From Grassmann to Dirac*. Birkhäuser Advanced Texts Basler Lehrbücher. Cham: Springer International Publishing, 2019.

[169] D. Micciancio and S. Goldwasser. *Complexity of Lattice Problems*. Boston, MA: Springer US, 2002.

[170] J. B. Fraleigh. *A First Course In Abstract Algebra, Seventh Edition*. Pearson Education, 2003.