

A New Framework for the Interpretation of Modulated
Martensites in Shape Memory Alloys

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Vincent Jusuf

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Ryan Scott Elliott, Adviser

June 2013

© Vincent Jusuf 2013

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank my adviser, Professor Ryan S. Elliott. I deeply appreciate your insight, guidance and patience. It has truly been my honor to be able to work with you over the past years.

I thank my dissertation committee, Professor Ellad Tadmor, Professor Richard James, and Professor Stefano Gonella, for their insight and advice.

I thank my family for their love, encouragement, and support without which I would not be the person I am today.

I thank my beloved Catherine whose love and support have been invaluable to me and continue to bring me joy. You are my sunshine.

During these years at the University of Minnesota, I have been fortunate to work alongside talented colleagues whom I now call friends. I especially would like to thank my office mates Amartya Banerjee, Nikhil Admal, and Vivekanand Dabade for sharing with me an immeasurable amount of knowledge. You have made long hours at the office more enjoyable. I would also like to acknowledge Venkata Guthikonda Vijay Srivastava, Shankar Krishnan and Amit Singh for their influence in shaping my research.

Last but not least, I thank my puppies Cosmo and Chili for giving me unconditional love and untold hours of joy and amusement.

To my loving family, Catherine, Cosmo and Chili.

Abstract

Shape memory alloys (SMAs) are a class of materials with unusual properties that have been attributed to the material undergoing a *martensitic phase transformation* (MPT). An MPT consists of the material's crystal structure evolving in a coordinated fashion from a high symmetry austenite phase to a low symmetry martensite phase. Often in SMAs, the austenite is a B2 cubic configuration that transforms into a Modulated Martensite (MM) phase. MMs are long-period stacking order structures consisting of $(110)_{\text{cubic}}$ basal planes. First-principles computational results have shown that the minimum energy phase for these materials is not a MM, but a short-period structure called the ground state martensite. It is commonly argued that energy contributions associated with kinematic compatibility constraints at the austenite-martensite interface explain the experimental observation of metastable MMs, as opposed to the expected ground state martensite phase. To date, a general approach for predicting the properties of the MM structure that will be observed for a particular material has not been available.

In this work, we develop a new framework for the interpretation of MMs as natural features of the material's energy landscape (expressed as a function of the lattice parameters and individual atomic positions within a perfect infinite crystal). From this energy-based framework, a new understanding of MMs as a mixture of two short-period base martensite phases is developed. Using only a small set of input data associated with the two base martensites, this Modulated Martensite Mixture Model (M^4) is capable of accurately predicting the energy, lattice constants, and structural details of an arbitrary modulated martensite phase. This is demonstrated by comparing the M^4 predictions to computational results from a particular empirical atomistic model.

Contents

	Page
Contents	iv
List of figures	viii
List of tables	xiii
Chapter 1 Introduction to shape memory alloys	1
1.0.1 Shape memory effect and pseudo-elasticity	2
1.0.2 Shape memory alloy applications	4
1.0.2.1 Military applications	4
1.0.2.2 Medical applications	4
1.0.2.3 Engineering applications	7
1.0.2.4 Safety applications	8
1.0.3 Current limitations of shape memory alloys	8
1.0.3.1 Predicting transition temperatures	8
1.0.3.2 Reversibility and fatigue	9
1.0.3.3 Environmental response	10
1.1 Mechanisms behind the shape memory effect and pseudo-elasticity	10

1.1.1	Lattice-scale description	11
1.1.1.1	Crystal structure	11
1.1.1.2	Cauchy-Born kinematics: Quasi-uniform deformation . . .	14
1.1.1.3	Energy density for perfect crystals	14
1.1.1.4	Cauchy-Born energy density	16
1.1.1.5	Homogenized continuum energy density	17
1.1.1.6	Stability of crystalline solids	19
1.1.1.7	Multi-welled SMA energy landscape	23
1.1.1.8	A more detailed description of a SMA energy landscape .	25
1.1.2	Micro-scale description	30
1.1.2.1	Twinning	30
1.1.2.2	Energy minimizing deformations for uniform displacement boundary conditions	32
1.1.2.3	Austenite-martensite interface	35
1.1.3	Macro-scale description	38
1.2	Modulated martensites	39
1.2.1	Adaptive martensite viewpoint of modulated martensites	46
1.2.2	Modulated martensites as periodic stacking of close-packed planes .	48
1.3	Probing the energy landscape of a typical SMA	49
1.4	Outline	50
Chapter 2	Branch-following and bifurcation method	52
2.1	Branch-following	53
2.1.1	Critical points	54
2.1.2	Symmetry of an equilibrium path	55

2.2	Lyapunov-Schmidt-Koiter decomposition and asymptotic bifurcation analysis	59
2.2.1	LSK decomposition	59
2.2.2	Asymptotic bifurcation analysis	61
Chapter 3	Branch-following and bifurcation study of a typical shape memory alloy	65
3.1	Problem description	65
3.1.1	Effective interaction potential model	66
3.1.2	Review: Cauchy-Born kinematics, free energy density, equilibrium equations	67
3.1.2.1	Cauchy-Born kinematics	67
3.1.2.2	Cascading Cauchy-Born	68
3.1.2.3	Free energy density and equilibrium equations	69
3.1.3	Morse EIPs with temperature dependent potential parameters	71
3.2	BFB result of a AuCd model	71
3.2.1	Some conclusions from <i>Guthikonda and Elliott (2011)</i>	74
3.3	Modulated martensites in a generic model of SMAs	77
3.3.1	Discussion	83
3.4	The effect of a kinematic compatibility constraint on modulated martensites	84
3.5	Summary and conclusion of the BFB study of AuCd for period-extending bifurcations	87
Chapter 4	The modulated martensite mixture model	88
4.1	The M^4 in detail	89
4.1.1	Two step decomposition of the deformation of a planar unit cell	92

4.1.2	The M^4 mixture rules	97
4.1.3	Recipe to construct a modulated martensite according to the M^4 . . .	98
4.1.4	Two interpretations of the <i>Balandraud and Zanzotto</i> (2007) model . . .	103
4.2	The modulated martensites k_5 , k_7 , and k_9 as interpreted by the M^4 and <i>Balandraud and Zanzotto</i> (2007)	104
4.3	Relative errors of the M^4 predictions	108
4.3.1	Discussion on the M^4 , B&Z ₁ , and B&Z ₂ prediction relative errors . . .	110
4.4	Using the M^4 to search for the most kinematically compatible modulated martensite	111
4.4.1	Constrained simulation result	115
4.5	The M^4 : Summary and conclusions	117
Chapter 5 Summary and conclusions		119
5.1	Some open questions	122
Bibliography		123
Appendix A Symmetry		131
A.1	Symmetry	131
A.1.1	Group Representation Theory	132
A.1.1.1	Invariant Subspaces and Irreducible Representations	132
A.1.2	Some Consequences of Symmetry	134
A.1.2.1	Branch-following in the presence of symmetry	134
A.1.2.2	Implications of symmetry on asymptotic bifurcation anal- ysis at a symmetry-breaking bifurcation point	136
Appendix B A Branch-following and bifurcation study of a one dimensional phase transforming material model		141

B.1	A one dimensional perfect crystal description	142
B.1.1	Internal potential energy, Cauchy-Born kinematics, and equilibrium equations in a one dimensional biatomic crystal model by <i>Dobson et al. (2008)</i>	144
B.2	Simulation results using a 24 atom representative unit cell	146
B.2.1	Period-extending bifurcations from \mathbf{u}_{24}^{2m}	147
B.2.1.1	General symmetry-breaking bifurcation points	147
B.2.1.2	Special symmetry-breaking bifurcation points	149
B.3	Simulation results using a 10 and 14 atom representative unit cell	155
B.3.1	Bifurcations from a D_n -symmetric path	155
B.3.2	Period-extending bifurcations from \mathbf{u}_{10}^{2m} and \mathbf{u}_{14}^{2m}	159
B.4	Summary	161
Appendix C An example of the determination of the forms of the coefficients in the expansion of the bifurcation equation		162
C.1	Form of $E_{,\xi\xi\lambda}$	163
C.2	Form of $E_{,\xi\xi\xi\xi}$	164
C.3	A remark regarding symmetric bifurcation points	164
Appendix D The application of a kinematic compatibility constraint		166
Appendix E The M^4 algorithm		168
E.1	B2, B19, and $L1_0$ input data	168

List of Figures

1.1	Experiment on NiTi thin wire showing the properties of shape memory and pseudo-elasticity	3
1.2	SMA NiTi used as a mechanical coupler	5
1.3	A schematic of orthodontic braces	5
1.4	The Mitek Anchor	6
1.5	A schematic of a cardio-vascular stent	6
1.6	Shape memory alloys in eyeglass frames	7
1.7	Graph of transition temperature of NiTi vs composition	9
1.8	A schematic showing transformation temperature hysteresis in a typical SMA	10
1.9	An example of a 4-Lattice.	13
1.10	A chart showing the performance of various descriptions of atomic interactions	15
1.11	The room-temperature phonon dispersion curves of α -Fe	23
1.12	A schematic of an energy landscape for a SMA.	24
1.13	A schematic of the energy well showing the austenite phase as a local minimum state.	27
1.14	A schematic of the energy well at the transition temperature.	28
1.15	A two-dimensional MPT.	29

1.16	A schematic of the energy wells at the transition temperature accounting for frame-indifference.	29
1.17	A deformation involving two variants of martensite separated by an interface.	31
1.18	A schematic representation of a twin.	32
1.19	A deformation with alternating gradients.	34
1.20	A deformation with alternating gradients and an interpolation layer.	34
1.21	A typical austenite-martensite interface.	36
1.22	Schematic representation depicting solution to the austenite-martensite interface equation (from (<i>Bhattacharya (2003)</i>)).	37
1.23	A schematic description of the shape memory effect.	38
1.24	An example of modulated martensite found in Cu-Al-Ni alloys.	41
1.25	A figure of the austenite phase in Cu-Al alloys <i>Nishiyama and Kajiwara (1963)</i>	42
1.26	An example of modulated martensites in Cu-Al alloys <i>Nishiyama and Kajiwara (1963)</i>	43
1.27	An example of a structure exhibiting two modulations of modulated martensites.	43
1.28	A plot of hysteresis vs λ_2 for various materials	44
1.29	DFT results for Ni_2MnGa	45
1.30	Modulated martensites as periodic stacking of close-packed planes	49
2.1	Starting from $\mathbf{c}(\alpha_0)$ a predictor step is taken onto $(\mathbf{u}^*, \lambda^*)$	53
2.2	Starting from $(\mathbf{u}^*, \lambda^*)$ the corrector process iteratively solves for $\mathbf{c}(\alpha_1)$	54
2.3	A square symmetric structure.	55
2.4	Symmetry of the square.	56
2.5	An example of a turning point found along the equilibrium path $\mathbf{c}(\alpha)$	58

2.6	An example of a bifurcation point found along the equilibrium path $c_1(\alpha)$. . .	58
2.7	A schematic description of an asymmetric bifurcation path	62
2.8	A schematic description of a symmetric bifurcation path	63
3.1	Results from the BFB study of AuCd using the model of <i>Guthikonda and Elliott</i> (2011)	73
3.2	The unit cells of the various structures found in the BFB study of AuCd using the model of <i>Guthikonda and Elliott</i> (2011)	74
3.3	A plot of energy versus normalized temperature from the BFB result of AuCd in <i>Guthikonda and Elliott</i> (2011)	75
3.4	The dispersion relation showing the onset of the first instability	76
3.5	A schematic of aB2 crystal subject to period-doubling disturbances	76
3.6	The dispersion relation showing the approach to the long wavelength instability	77
3.7	A schematic of a B2 crystal at the period-preserving bifurcation point . . .	77
3.8	Dispersion relation of the B2 configuration as a function of temperature . .	78
3.9	Examples of simulated cells (shown in gray) necessary to capture period-extending deformations	79
3.10	BFB study of AuCd showing period-extending bifurcation points	81
3.11	Energy of the stable structures found in the BFB study of AuCd	82
3.12	Kinematic compatibility of modulated structures as compared to the ground state martensite	85
3.13	The effect of a kinematic compatibility constraint on the energy of various structures found in the BFB study of AuCd	86
4.1	The dispersion relation of the B2 configuration as a function of temperature	90
4.2	The two base martensites	90
4.3	A portion of a crystal with B2 symmetry	92

4.4	A projection of two variants B19 crystal.	93
4.5	The four plane types in the M^4 after an in-plane atomic motion	93
4.6	The plane types B^+ and B^- in the M^4 after the application of uniform deformations F_{B^+} and F_{B^-}	94
4.7	A projection of two variants of $L1_0$ crystals.	95
4.8	The plane types L^+ and L^- in the M^4 after the application of uniform deformations F_{L^+} and F_{L^-}	96
4.9	The four plane types in the M^4	97
4.10	The M^4 Mixture rules	98
4.11	A schematic of the intermediate planar unit cells for the modulated martensite described by $(B^-, B^+, L^-, B^-, B^+, B^-, B^+, L^-, B^-, B^+)$	99
4.12	The transformed planar unit cells for the modulated martensite described by $(B^-, B^+, L^-, B^-, B^+, B^-, B^+, L^-, B^-, B^+)$	100
4.13	The planar unit cells after it has been fully deformed	101
4.14	The modulated structure $(B^-, B^+, L^-, B^-, B^+, B^-, B^+, L^-, B^-, B^+)$	101
4.15	The two types of unit cells according to <i>Balandraud and Zanzotto (2007)</i>	103
4.16	The modulated martensite k_5 as interpreted by the M^4 and <i>Balandraud and Zanzotto (2007)</i>	105
4.17	The modulated martensite k_7 as interpreted by the M^4 and <i>Balandraud and Zanzotto (2007)</i>	106
4.18	The modulated martensite k_9 as interpreted by the M^4 and <i>Balandraud and Zanzotto (2007)</i>	107
4.19	The M^4 relative error in λ_1	108
4.20	The M^4 relative error in λ_2	108
4.21	The M^4 relative error in λ_3	109
4.22	The M^4 relative error in $\ S\ $	109

4.23	The M^4 relative error in \mathcal{E}	109
4.24	Kinematic compatibility of $\mathbf{k}_7^{(2)}$ compared to B19 and other modulated structures found	112
4.25	The modulated martensite $\mathbf{k}_7^{(2)}$	113
4.26	$\mathbf{k}_7^{(2)}$ path	114
4.27	B2 \leftrightarrow $\mathbf{k}_7^{(2)}$ transition at $\theta = 2.52$ and $\theta = 2.52$	115
4.28	Energy for constrained and unconstrained modulated structures	116
B.1	A one dimensional lattice with reference lattice parameter a . The region enclosed by the black square is an example of an essential unit cell. The gray region is an example of a non-essential unit cell comprising of two atoms.	142
B.2	The 2_m path of the biatomic chain.	146
B.3	A segment of a one dimensional biatomic crystal.	147
B.4	A figure showing paths bifurcating from GSBPPs. Solid and dashed lines indicate stable and unstable solutions respectively.	149
B.5	A figure showing six special symmetry-breaking bifurcation points found along the \mathbf{u}_{24}^{2m} path.	150
B.6	A figure showing a path bifurcating away from a SSBBP (shown in plum).	153
B.7	All paths bifurcating from SSBBPs along the \mathbf{u}_{24}^{2m} path (shown in red). Note that none of these paths were found to be stable.	154
B.8	Simulation result using a 10 atom non-essential unit cell.	160
B.9	Simulation result using a 14 atom non-essential unit cell.	160
C.1	A figure showing six special symmetry-breaking bifurcation points found along the principal path.	162

List of Tables

3.1	Fitted parameters for the Au-47.5 atomic %Cd Morse EIP model reproduced here from <i>Guthikonda and Elliott (2011)</i>	72
3.2	Bifurcation temperatures of modulated structures found in the BFB study of AuCd using the model developed by <i>Guthikonda and Elliott (2011)</i> . Structures marked as “stable” imply the path contains a stable segment. . . .	81
3.3	Relative energy values (at $\theta = 1.5$) of the various structures found in the BFB of AuCd. The values are shown relative to the energy of the ground state martensite B19.	83
3.4	Constrained energy values of B19, \mathbf{k}_5 , \mathbf{k}_7 , and \mathbf{k}_9 . $\Delta\mathcal{E}$ gives the increase in energy between the constrained and unconstrained energy values.	86
4.1	The M^4 mixture rules.	97
4.2	The ratio of middle eigenvalues λ_2/λ_2^{B2} for various structures at $\theta = 1.5$. The structure $\mathbf{k}_7^{(2)}$ was found using the M^4 and is the most compatible structure to B2.	112
4.3	The energy difference (relative to unconstrained B19) for constrained and unconstrained structures at $\theta = 1.5$. $\mathbf{k}_7^{(2)}$ is energetically more favorable upon the application of a kinematic compatibility constraint.	116
B.1	Parameters used in the LJ potential of <i>Dobson et al. (2008)</i> . r_{cut} is the cutoff distance and a is the lattice parameter of the essential unit cell.	145

Chapter 1

Introduction to shape memory alloys

The motivation for this work is to understand the mechanism that is responsible for the fascinating properties exhibited in shape memory alloys (SMAs). Shape memory alloys such as NiTi, AuCd, CuAlNi, AgCd, and TiPd are special materials that exhibit the *shape memory effect* which is the material's ability to fully recover deformations of up to 10% strain upon temperature cycling past a transition temperature. At high temperatures, these materials also exhibit *pseudo-elasticity* which is the ability of the material, through a hysteretic process, to completely recover its (unloaded) shape after the application and removal of large loads. The fascinating properties of SMAs can be attributed to the material undergoing a *martensitic phase transformation* (MPT) where the material changes its crystallographic structure (*Bhattacharya (2003)*).

A martensitic phase transformation occurs when the underlying crystal structure becomes unstable at a particular thermo-mechanical load and the material reconfigures into a new stable structure. These transformations are often temperature-induced and may occur over a wide range of temperatures depending on factors such as material composition. A martensitic phase transformation is a first-order, diffusionless, solid-to-solid phase transformation that involves the coordinated displacement of the atoms within the crystal as the crystal structure transforms from one lattice to another at some critical temperature. The high temperature crystal structure is usually a high symmetry structure and is known as the *austenite* phase whereas the low temperature structure has a lower symmetry than the austenite phase and is referred to as the *martensite* phase. In some MPTs the unit cell of the low-symmetry phase may only consist of a single unit cell of the high-symmetry phase (for example, BaTiO₃ (*Megaw (1947)*)). For other MPTs, the unit cell of the low-symmetry phase may consist of two, three, or more unit cells of the high-symmetry phase (for example, Per-

ovskites (*Unoki and Sakudo (1967)*)). These latter cases are called period-extending MPTs. The first observation of the shape memory effect dates back to 1932 when a Swedish researcher named Arne Olander observed the shape memory property in AuCd alloys (*Olander (1932)*). Olander used electrochemical techniques on AuCd alloys to identify the B2 cubic phase and B19 orthorhombic phase as the high temperature and low temperature phases, respectively. This was later confirmed by *Bystrom and Almin (1947)* using X-ray diffraction methods on AuCd alloys of different compositions.

Aside from the shape memory effect and pseudo-elasticity, other macroscopic properties such as specific heat and dielectric susceptibility may change at structural phase transitions (*Otsuka and Wayman (1999)*). As such, it is important to study the mechanism behind phase transitions in order to accurately account for the behavior of SMAs in a given application.

1.0.1 Shape memory effect and pseudo-elasticity

The properties of shape memory effect and pseudo-elasticity are illustrated using Fig. 1.1, which is obtained from *Shaw (1997)*. This figure shows the effect of load and temperature on the deformation of a polycrystalline thin wire of NiTi. Here, δ/L is the elongation of the wire from its original length (i.e., nominal strain), P/A_0 is the load per reference cross-sectional area (i.e., nominal stress) and T is the temperature. The shape memory effect is shown in the figure from segments ① to ④ in which the material is initially loaded (segment ① to ②) and unloaded (② to ③). At this point the wire is seemingly “permanently” deformed as is indicated by the non-zero strain at zero load. The wire is then heated (segments ③ to ④) and recovers its shape once heated past a transition temperature (somewhere between segments ③ and ④). It then maintains its shape upon cooling (④ to ①). Also shown in the figure is a schematic depiction of the microstructure of the material that is responsible for the shape memory effect. Starting again from point ①, the material is in its self-accommodated microstructure of twinned martensite. The material is then detwinned through the application of load to obtain the deformed wire (②). Upon heating, the NiTi wire undergoes a phase transition to a high-symmetry austenite phase (shown in red) and recovers its original shape. The self-accommodated structure (of state ①) is then obtained after the wire has been cooled. This transformation of the austenite to a microstructure with no net change in shape or volume is known as *self-accommodation* (*Bhattacharya (2003)*). Pseudo-elasticity is depicted in Fig. 1.1 in segments ⑤ to ⑩. Pseudo-elasticity refers to the fact that energy is dissipated via a hysteresis loop as shown in the figure. At high temperature the material initially behaves as a linearly elastic material

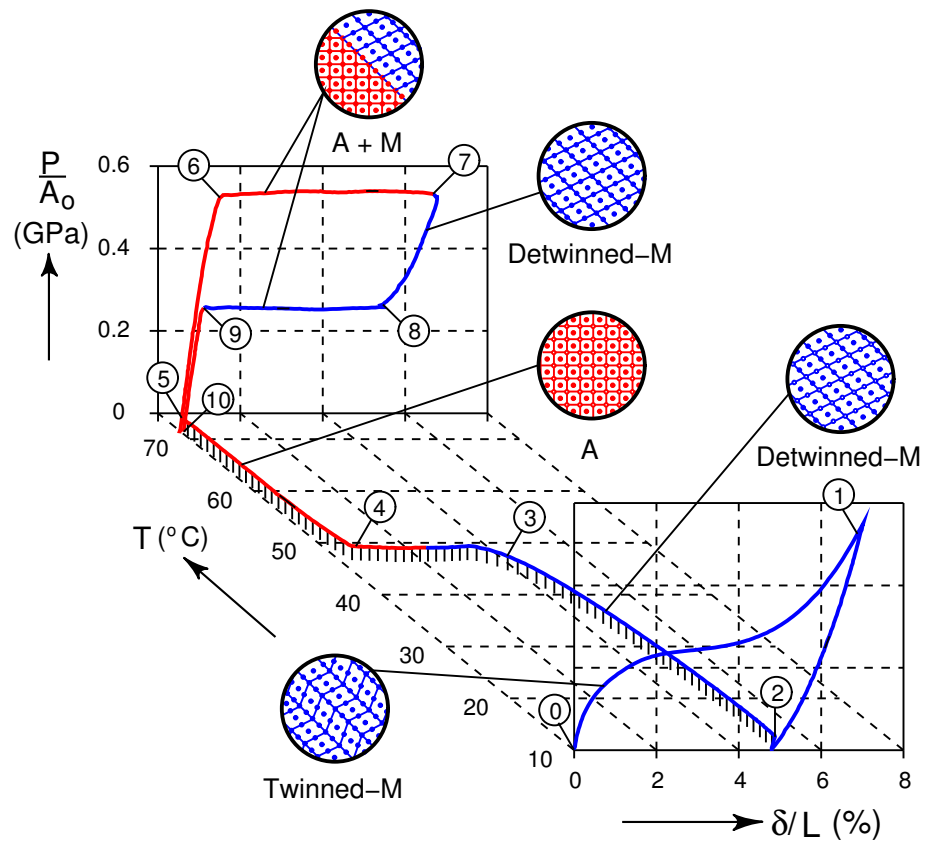


Figure 1.1: Experiment on NiTi thin wire showing the properties of shape memory and pseudo-elasticity (from *Shaw* (1997)).

(⑤ to ⑥). Here the microstructure of the wire remains in the high-symmetry austenite phase. The wire then undergoes a large elongation with no appreciable increase in load shown between segment ⑥ and ⑦. In this stage, the material is a combination of both the austenite and martensite phases as the wire undergoes a stress-induced phase transition up to point ⑦ where the material consists completely of detwinned martensite. Upon removal of load (segments ⑦ to ⑩) the material reaches a critical stress (point ⑧) where a phase transition to the combination of austenite and martensite phases occurs somewhere between points ⑧ and ⑨. During this stage, the wire again undergoes a large deformation with no appreciable change in load. Finally, between points ⑨ and ⑩ the NiTi wire is again in the austenite phase and behaves as a linearly elastic material.

1.0.2 Shape memory alloy applications

Martensitic phase transformations produce material property changes that have often been used for technological purposes. This section will present several examples of how SMAs have been used in various applications to great success.

1.0.2.1 Military applications

The most common of the SMAs to be used in various applications is a Nickel-Titanium alloy known as Nitinol. For example, one of the first applications of SMAs was for the repair of leaking titanium hydraulic pipes in the wing box of jet fighters between the 1960s and 1970s (*Kauffman and Mayo* (1997)). This particular Nitinol product was manufactured by the Raychem Corporation under the product name Cryofit™ and is a “shrink-to-fit” pipe coupler used to solve the problem of coupling hydraulic-fluid lines in the F-14 jet fighter built by the Grumman Aerospace Corporation. A schematic of the Cryofit™ coupler can be seen in Fig. 1.2.

1.0.2.2 Medical applications

Another early use of Nitinol can be found in orthodontic bridge wires (Fig. 1.3). Prior to Nitinol, orthodontic wires were mostly composed of stainless steel. However, George B. Andreasen, D.D.S of the University of Iowa noticed that the large recoverable strain combined with a low elastic modulus of Nitinol made it ideal for orthodontic applications (*Kauffman and Mayo* (1997)).

One revolutionary use of SMAs has been in the area of orthopedic surgery. In 1989, the U.S. Food and Drug Administration approved the use of Mitek Surgical Products’ Mitek

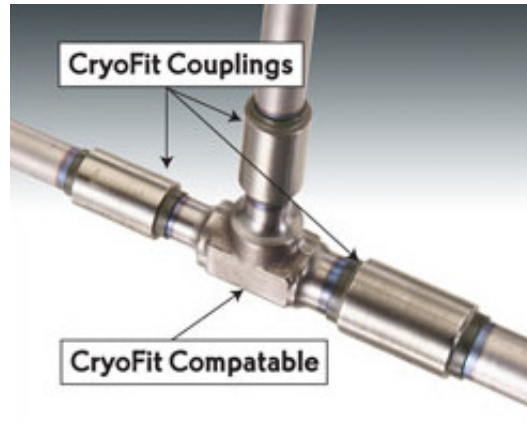


Figure 1.2: The Cryofit™ coupler is made from an alloy of the SMA NiTi. The coupler has an inside diameter that is slightly larger than the tube outside diameter. As the coupler warms, it recovers its shape and generates a reliable metal to metal seal. Figure obtained from <http://www.aerofit.com>.



Figure 1.3: A schematic of orthodontic braces. Figure obtained from <http://www.bracesquestions.com>.

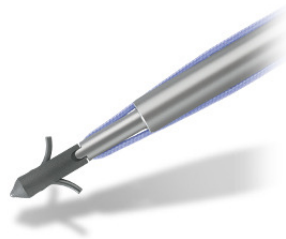


Figure 1.4: Mitek Surgical Products' Mitek Anchor is used to treat torn ligaments and tendons.

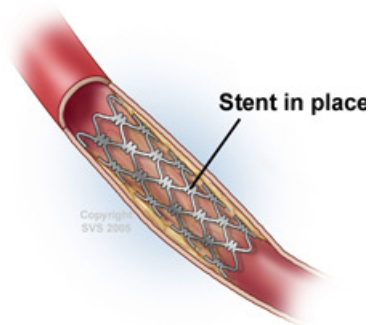


Figure 1.5: A schematic of a cardio-vascular stent. Below the transition temperature, the stent may be contracted for easy insertion into the blood vessels. The original shape is recovered as the stent is warmed to body temperature. Figure obtained from <http://www.vascularweb.org>.

Anchor (see Fig. 1.4) to treat torn ligaments and tendons. Prior to the Mitek Anchor, the most common way to treat torn ligaments and tendons was to surgically attach the muscles to the bone. This surgical procedure involves screws and staples and is quite invasive to the body. Mitek's Anchor is a fraction of the size of the older devices and requires only tiny incisions through which the compressed Anchor is squeezed through. The Anchor subsequently returns to its designed shape after it warms to body temperature (*Kauffman and Mayo (1997)*).

Another application found in the medical industry is the use of SMAs as vascular stents (Fig. 1.5) which are used to prevent the collapse of blood vessels after angioplasty procedures. This application seems to be one of the most important applications of SMAs (for example, Medtronic Corporation posted a second quarter 2012 earning of \$210 million on vascular stents alone).

Another medical use for Nitinol is in the design for a blood filter that can be set in a vein to trap blood clots without surgery (*Pelton et al. (2000)*). The Simon-Nitinol filter, which



Figure 1.6: Shape memory alloys have been used in eyeglass frames due to their ability to sustain large recoverable strains. Figure obtained from <http://www.marchon.com>.

was patented by Morris Simon, M.D. of Beth Israel Medical Center replaces the process of surgically implanting blood filters in order to trap blood clots (pulmonary embolisms). Both the vascular stents and the Simon-Nitinol filter take advantage of the SME in the following way. First, the material is deformed into a desired shape. After the desired shape has been reached, the material is then cooled, compressed, and inserted into a large vein via a catheter. When the stent or filter is warmed to body temperature, the SME causes it to spring back into its original shape.

SMA has also been used as flexible eyeglass frames released under the trademark Flexon by the Marchon company (<http://marchon.com>). The SMA is used in the frame's bridge, top bar, and temples where flexibility is most needed while allowing the frame to retain its original shape (Fig. 1.6).

1.0.2.3 Engineering applications

Shape memory alloys have also been researched for use as solid-state heat engines. In fact, an example of this application is marketed as an educational toy under the product name Thermobile™. Thermobile™ is an educational toy that demonstrates the conversion of heat energy to mechanical energy. The toy works in the following manner: one end of the toy is inserted into warm water. SMA wire loops encircling pulleys then contract, exerting torque on the pulleys causing them to rotate. Currently, the use of SMAs as solid-state heat engines is still being developed (see for example, *Browne et al.* (2012)). A couple of issues that still need to be resolved include the increase in power output as well as the increase in lifespan (number of transformation cycles before failure) of the SMA (*Kauffman and Mayo* (1997)). Interestingly a patent exists (United States Patent number: 5442914 filed in 1993) for a shape memory alloy heat engine.

The unique properties of SMAs have also made them ideal for certain applications in ocean

engineering (*Buehler and Wang (1968)*). Memory Metals, Inc. of Stamford, Connecticut is currently studying the use of corrosion-resistant SMA seals for use in geothermal brine wells (*Kauffman and Mayo (1997)*).

The consideration of SMAs for various other engineering applications is an active area of research. For example, Beta Phase, Inc. is researching SMAs for use as electrical switches to open or close electrical connections between components. Boeing Corporation is also researching SMAs as use for “hingeless” actuators to replace complex mechanical devices such as motors, pulleys, and gears (*Kauffman and Mayo (1997)*).

1.0.2.4 Safety applications

A fire sprinkler system has been developed by Battelle Columbus Laboratories in Ohio in which a SMA is used as a heat sensor to drastically decrease the response time from fire to water release. SMAs have also been used as automatic hot water shut-off valves. The Memry Corporation has developed an anti-scald valve that shuts off the release of water when the temperature of the water rises dangerously high (*Kauffman and Mayo (1997)*).

1.0.3 Current limitations of shape memory alloys

Despite the use of SMAs in a variety of applications such as the ones discussed in the preceding subsection, there are various challenges that still remain in the design of SMAs. This subsection gives a brief discussion on some of these challenges.

1.0.3.1 Predicting transition temperatures

One challenge in designing SMAs for a particular application is in predicting the temperature at which the transition occurs. One of the main reasons for this is that the transition temperature is highly sensitive to the composition of the alloy. Experimental results of this effect can be seen in Fig. 1.7 obtained from *Duerig et al. (2011)*. The experiment determined that while the transition temperature for NiTi changes little at around 48 to 49 percent Nickel content, it drastically changes by about 200 degrees Celsius from 49 to 51 percent Nickel content. Additionally, the transition temperature for a virgin SMA of a given composition changes with each load cycle until it eventually stabilizes after several cycles (*Churchill and Shaw (2008)*). While this issue can be addressed by pre-cycling the alloy, the pre-cycling also serves to decrease the useable lifespan of the SMA.

Another difficulty in predicting the transition temperature of an SMA is that the transition temperature upon heating does not equal the transition temperature upon cooling. That is,

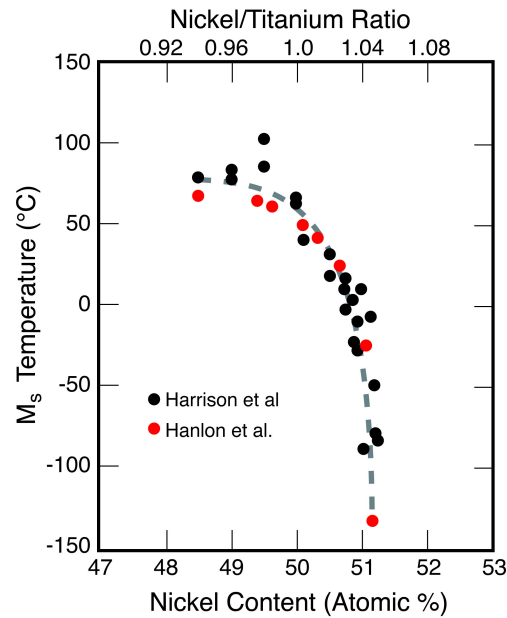


Figure 1.7: The transition temperature of SMAs is highly sensitive to composition. Small changes in composition between 49 to 51 percent Nickel content produce a drastic change in transition temperature of the SMA NiTi. Figure is obtained from *Duerig et al. (2011)*.

the transformation does not occur at one fixed temperature but over a wide range of temperature. The transformation begins at one temperature (known as the start temperature) and is completed at another temperature (known as the finish temperature). Additionally, there is a difference between the transition temperature upon cooling from austenite to martensite and upon heating from martensite to austenite. A schematic of this effect can be seen in Fig. 1.8. The figure shows the response of a generic SMA as it is heated and cooled. Points along the curve have been marked where martensite and austenite formation has finished and started (M_f , A_f , M_s , and A_s). This effect is commonly referred to as the hysteresis in the SMA response and is generally defined as the difference between the temperatures at which the material is 50 percent transformed to austenite upon heating and 50 percent transformed to martensite upon cooling (*Melton et al. (1988)*). It is also commonly observed that the hysteresis depends on the composition of the alloy (*Duerig et al. (2011)*). For typical uses, it is desirable to minimize the width of the hysteresis.

1.0.3.2 Reversibility and fatigue

Both reversibility and fatigue are the results of nano-scale damage to the material as it is loaded and unloaded. Reversibility refers to the fact that SMAs do not *perfectly* recover

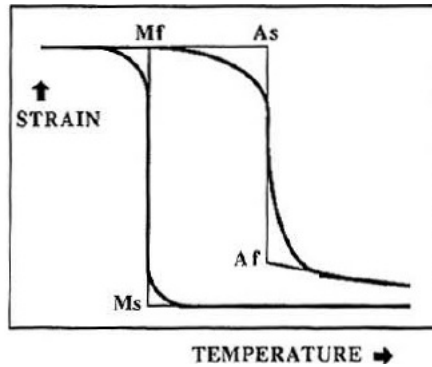


Figure 1.8: A schematic showing transformation temperature hysteresis in a typical SMA (Melton *et al.* (1988)).

their initial shape upon heating or cooling. That is, there is a permanent strain of the material upon loading and unloading. This permanent strain is attributed to the occurrence of permanent slip during a shape change and the lack of reversible movement of partial dislocations *Tsuzaki et al.* (1995). However, it is unclear how much of the permanent strain results with each cycle, or how composition affects the permanent strain. A more thorough understanding of the mechanisms behind reversibility and fatigue of SMAs may allow for the design of a more ideal SMA with a longer usable lifespan.

1.0.3.3 Environmental response

Shape memory alloys also display a sensitivity to environmental stimuli such as exposure to acidic media. For example, for Ni-Mn-Ga alloys in acidic environments, the martensite is significantly more chemically reactive than the austenite (*Gostin et al.* (2012)). This can be attributed to a much higher density of surface defects in terms of a large number of twin boundaries. Additionally, SMAs also show a sensitivity to precipitate crack initiation. Precipitates are often used in order to improve the mechanical properties of the SMA but have the undesirable effect of increasing the number of sites for crack formation (*McCluskey et al.* (2011)).

1.1 Mechanisms behind the shape memory effect and pseudo-elasticity

In order to develop an understanding of the shape memory effect and pseudo-elasticity, it is necessary to start at the nanometer length scale with a description of the arrangement of individual atoms in the material and then to consider the material's behavior at a set of successively larger length scales. We start with a description of the basic building blocks in terms of which SMAs are described.

1.1.1 Lattice-scale description

SMA are materials comprised of atoms arranged to form crystals. A crystal structure is described by (1) its associated *Bravais lattice*, which describes the long range order and translational periodicity of the atomic arrangement throughout space, and (2) its *lattice basis*, which describes the short range order and arrangement of atoms about each Bravais lattice site.

1.1.1.1 Crystal structure

The notion of Bravais lattices and lattice basis will be discussed here in order to provide the fundamental language that will be used to discuss the crystal structure of SMAs. The presentation here will follow the work presented in *Bhattacharya (2003)*, *Tilley (2006)*, and *Elliott et al. (2006a)*.

1. *Bravais lattice*

A Bravais lattice $\mathcal{L}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{o})$ is an infinite set of points in three-dimensional space generated by the periodic translation of a single point ‘ \mathbf{o} ’ through three linearly independent vectors. The three vectors are known as *lattice vectors* and are denoted here as $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. So, if ‘ \mathbf{o} ’ is an arbitrary choice of origin in three-dimensional space, then

$$\mathcal{L}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{o}) = \{\mathbf{x} \mid \mathbf{x} = \sum_{i=1}^3 \ell^i \mathbf{e}_i + \mathbf{o}; \ell^1, \ell^2, \ell^3 \in \mathbb{Z}\}. \quad (1.1)$$

The lattice vectors $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ and the choice of lattice center ‘ \mathbf{o} ’ then define a unit cell of a Bravais lattice. Additionally, the choice of lattice vectors to describe a Bravais lattice is non-unique but belongs to a class of linearly independent lattice vector sets that generate all lattice points when the linear combinations are formed as in Eq. (1.1). For example, if $\mathcal{L}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{o})$ describes a Bravais lattice, then there is another set of lattice vectors $(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3)$, and origin \mathbf{o}' such that $\mathcal{L}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{o}) = \mathcal{L}(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3, \mathbf{o}')$. Often, the unit cell is chosen for convenience and to reveal the underlying symmetry of the crystal.

Associated with each set of lattice vectors $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, is a set of *reciprocal lattice* vectors $(\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3)$ defined by

$$\mathbf{e}^i \cdot \mathbf{e}_j = \delta_{ij}, \quad i, j = 1, 2, 3, \quad (1.2)$$

where δ_{ij} is the Kronecker delta. From this, the metric tensors G^{ij} and G_{ij} can be constructed such that $\mathbf{e}^i = G^{ij}\mathbf{e}_j$ and $\mathbf{e}_i = G_{ij}\mathbf{e}^j$ as follows:

$$G^{ij} = \mathbf{e}^i \cdot \mathbf{e}^j, \quad G_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j. \quad (1.3)$$

The reciprocal lattice vectors can be used to define the first *Brillouin zone* which is the primitive unit cell in reciprocal space (i.e., using reciprocal lattice vectors). It can be shown that the vibrational behavior of a crystal can be completely described by consideration of a set of plane wave perturbations associated with each point within the first Brillouin zone.

2. Lattice basis

In order to describe any periodic arrangement of atoms, the notion of a *multi-lattice* will be defined. A multi-lattice is a periodic arrangement of atoms described as a collection of a *finite number of inter-penetrating Bravais lattices defined by the same lattice vectors*. Any periodic arrangement of atoms can be described as a multi-lattice. A multi-lattice consisting of M distinct mono-atomic sub-lattices with the same set of lattice vectors ($\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$) is called an M -lattice.

An M -lattice is described by its lattice vectors ($\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$) as well as the relative *fractional position vectors* \mathbf{P}_α (Sands (1993)) of each sub-lattice with respect to a “skeletal lattice”. The fractional position vectors are defined as follows:

$$\mathbf{P}_\alpha = P_\alpha^i \mathbf{e}_i, \quad P_\alpha^i \in [0, 1), \quad i = 1, 2, 3, \quad \alpha = 0, 1, \dots, M - 1. \quad (1.4)$$

This allows for the identification of each sub-lattice α that is embedded into the skeletal lattice and gives the offset of the sub-lattice origin with respect to the origin of the skeletal lattice. A unit cell for the M -lattice is given by the unit cell of the skeletal lattice and contains M atoms. An example of this description is given in Fig. 1.9 where a two-dimensional 4-lattice is shown. The 4-lattice is described by sub-lattices $\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2,$ and \mathcal{L}_3 with lattice vectors ($\mathbf{e}_1, \mathbf{e}_2$) and fractional position vectors $\mathbf{P}_0, \mathbf{P}_1, \mathbf{P}_2,$ and \mathbf{P}_3 that give the offset of each sub-lattice with respect to the skeletal lattice origin. The 4-lattice crystal is shown on the top right hand corner of the figure denoted by Ω with the shaded region in Ω showing the 4-lattice unit cell. The reference position of the α -th atom in the ℓ -th unit cell is given by

$$\mathbf{X}_{(\ell, \alpha)} = \mathbf{X}_\ell + \mathbf{P}_\alpha, \quad \mathbf{X}_\ell = \ell^i \mathbf{e}_i, \quad \mathbf{P}_\alpha = P_\alpha^i \mathbf{e}_i, \quad i = 1, 2, 3, \quad (1.5)$$

where \mathbf{X}_ℓ gives the position vector of the ℓ -th skeletal lattice unit cell.

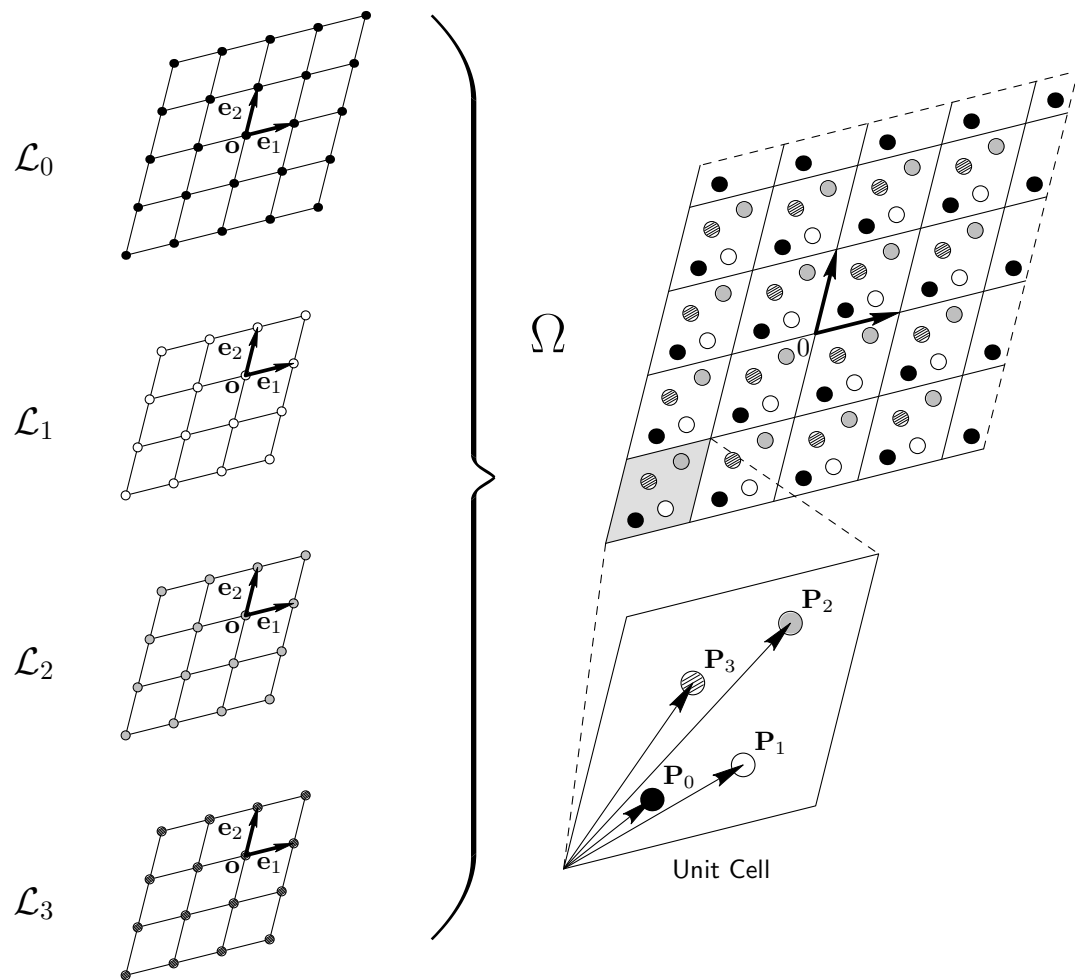


Figure 1.9: An example of a 4-Lattice (figure obtained from *Elliott et al. (2006a)*).

1.1.1.2 Cauchy-Born kinematics: Quasi-uniform deformation

In this work, allowable deformations of an M -lattice are those that may be described as another M -lattice. Thus, the current atomic positions can be written as

$$\mathbf{x}_{(\ell,\alpha)} = \mathbf{X}_{(\ell,\alpha)} + \mathbf{u}_{(\ell,\alpha)} \quad (1.6)$$

$$= \mathbf{F}(\mathbf{X}_{(\ell,\alpha)} + \mathbf{S}_\alpha), \quad (1.7)$$

where $\mathbf{u}_{(\ell,\alpha)}$ in Eq. (1.6) gives the displacement away from the reference configuration of atom α in unit cell ℓ . The alternative definition of the current atomic positions (Eq. (1.7)) is given in terms of a uniform deformation tensor \mathbf{F} and the sub-lattices' reference "internal atomic shift" vectors \mathbf{S}_α . This system of describing crystalline deformations is called *Cauchy-Born* (CB) kinematics (see for example, *Born and Huang (1954)*, *Elliott et al. (2006a)*, *Elliott et al. (2006b)*).

1.1.1.3 Energy density for perfect crystals

The modeling of SMAs using an energy density has been a subject of active study for many decades. In 1937 L. D. Landau developed a phenomenological theory of second-order phase transitions that was later generalized by Devonshire to include first order transitions (including MPTs) (*Devonshire (1949)*, *Devonshire (1951)*). This theory has successfully modeled the properties of known phase transitions (see, for example *Dove (1993)*). Landau theory is based on a phenomenological expression for the free energy as a Taylor expansion in the order parameter (*Cowley (1980)*). This is in contrast to the approach taken in this work where the free energy is explicitly derived from an atomistic description of the crystal. As such, this section will present a discussion of the free energy density function for a perfect crystal that will follow the presentation given in *Elliott et al. (2006a)* and *Bhattacharya (2003)*.

In order to obtain the energy density for a crystal, certain assumptions regarding the atomic interactions must be made. The most accurate description of atomic interactions is through the quantum mechanical description of atoms (see for example, *Eck et al. (1999)*). Here, the atomic interactions are described through the coupled nuclei and electron interactions where the spatial distribution of the electrons is first calculated through the Schrödinger equation. The force on the nuclei by the electrons is then described by the Hellmann-Feynmann theorem which states that once the spatial distribution of the electrons are known, then the forces in the system can be calculated using concepts in classical electrostatics (*Cancès et al. (2003)*). Unfortunately, for all but the simplest systems, the full quantum me-

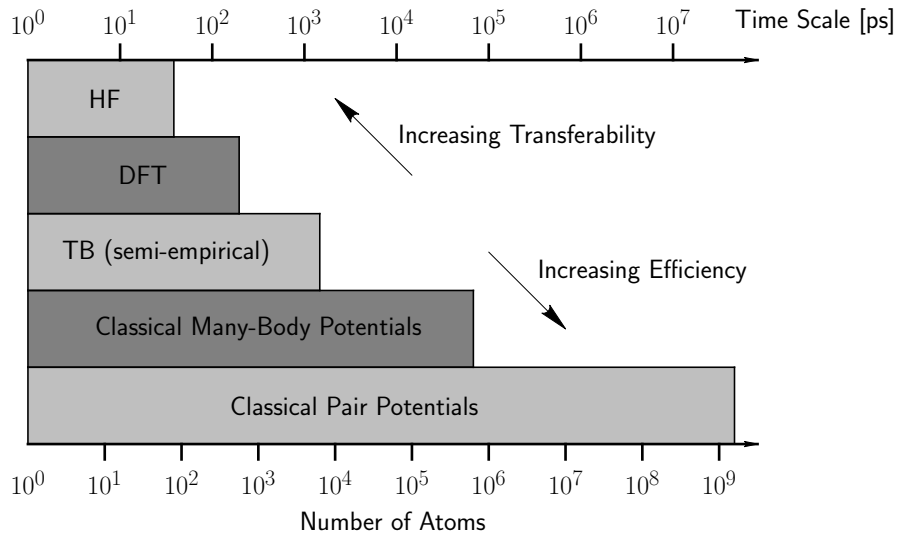


Figure 1.10: A chart showing the performance of various descriptions of atomic interactions used in the modeling of materials. Here, HF is the Hartree Fock method, DFT is Density Function Theory, TB is the Tight Binding method. The figure is obtained from the KIM project (E. B. Tadmor, R. S. Elliott, J. P. Sethna, R. E. Miller, and C. A. Becker. Knowledgebase of Interatomic Models (KIM). <http://openkim.org>, 2011).

chanical system contains a vast number of *degrees of freedom* (DOFs) and implementation of such a simulation is impractical for realistic systems containing (upwards of) 10^{23} atoms. As such, instead of a full quantum mechanical system, simplifying assumptions regarding atomic interactions can be used. The next level of approximation gives rise to methods such as Density Functional Theory (e.g., *Zayak et al. (2003)*). This method does better in that it is less computationally expensive than the full quantum mechanical description and as such is typically used in simulations involving up to 10^3 atoms (*Steinhauser (2007)*). This trend continues where greater approximation gives rise to methods that are less computationally expensive. See Fig. 1.10 showing the performance of various methods typically used in the modeling of materials. The figure is obtained from the KIM project (E. B. Tadmor, R. S. Elliott, J. P. Sethna, R. E. Miller, and C. A. Becker. Knowledgebase of Interatomic Models (KIM). <http://openkim.org>, 2011).

In this work, atoms in a crystal Ω are assumed to interact through a set of atomic potentials that satisfy the following conditions: (1) the interactions between the atoms are described by forces between the atoms and are assumed to interact pair-wise¹ (i.e., the force on atom i is described by the sum of the forces exerted on atom i by atom j as j ranges through all atoms in the crystal), (2) the forces on the atoms are derived from a potential function given by the net sum of the free energy contributions of individual atoms $\psi_{(\ell,\alpha)}(\theta)$ in the

¹Note that the use of pair interactions is not essential to anything that is done in this work.

system where θ denotes the temperature, and (3) the free energy contribution $\psi_{(\ell,\alpha)}(\theta)$ of an atom depends only on the positions of the surrounding atoms and the temperature. The resulting free energy of a finite crystal will then be a function of the atomic displacements $\mathbf{u}_{(\ell,\alpha)}$ from their reference positions $\mathbf{X}_{(\ell,\alpha)}$ and the temperature θ . The free energy density \mathcal{E} is then given by

$$\mathcal{E}(\mathbf{u}, \theta) = \frac{1}{nV} \sum_{(\ell,\alpha) \in \Omega} \psi_{(\ell,\alpha)}(\theta), \quad \mathbf{u} \equiv \{\mathbf{u}_{(\ell,\alpha)} | (\ell, \alpha) \in \Omega\}, \quad (1.8)$$

where \mathbf{u} is the vector of all atomic displacements, $V = \mathbf{e}_1 \cdot (\mathbf{e}_2 \times \mathbf{e}_3)$ is the reference unit cell volume and n is the number of unit cells in the crystal. The total sum of forces $\Phi_i^{(\ell,\alpha)}$ on atom (ℓ, α) at some configuration $(\tilde{\mathbf{u}}, \tilde{\theta})$ in the \mathbf{e}_i ($i = 1, 2, 3$) direction on each atom due to all other atoms in Ω is given by

$$\Phi_i^{(\ell,\alpha)} \equiv \left. \frac{\partial nV\mathcal{E}}{\partial u^i_{(\ell,\alpha)}} \right|_{(\tilde{\mathbf{u}}, \tilde{\theta})}, \quad (1.9)$$

and the stiffness $K_{ij}^{(\ell,\ell',\alpha,\alpha')}$ between atoms (ℓ, α) and (ℓ', α') in Ω is given by

$$K_{ij}^{(\ell,\ell',\alpha,\alpha')} \equiv \left. \frac{\partial^2 nV\mathcal{E}}{\partial u^i_{(\ell,\alpha)} \partial u^j_{(\ell',\alpha')}} \right|_{(\tilde{\mathbf{u}}, \tilde{\theta})}. \quad (1.10)$$

1.1.1.4 Cauchy-Born energy density

Substituting CB kinematics into the total energy density given in Eq. (1.8) and introducing the notation $\mathbf{S} = (\mathbf{S}^0, \mathbf{S}^1, \dots, \mathbf{S}^{N-1})$, for the internal atomic shift vectors results in the CB energy density given by

$$\tilde{\mathcal{E}}(\mathbf{F}, \mathbf{S}, \theta) \equiv \mathcal{E}(\mathbf{u}_{(\ell,\alpha)}(\mathbf{F}, \mathbf{S}), \theta). \quad (1.11)$$

Equilibrium configurations $(\mathring{\mathbf{F}}, \mathring{\mathbf{S}}, \mathring{\theta})$ of the crystal are given by extrema of the total energy density function as reflected in the following equations:

$$\left. \frac{\partial \tilde{\mathcal{E}}}{\partial F_j^i} \right|_{(\mathring{\mathbf{F}}, \mathring{\mathbf{S}}, \mathring{\theta})} = 0, \quad \text{and} \quad \left. \frac{\partial \tilde{\mathcal{E}}}{\partial S^i} \right|_{(\mathring{\mathbf{F}}, \mathring{\mathbf{S}}, \mathring{\theta})} = 0. \quad (1.12)$$

One solution to the equilibrium equations is a rigid-body translation of the crystal. Often, only solutions containing zero translations are desired. One method to eliminate these rigid-body modes is to set the shift of one atom in the crystal to zero (say, $\mathbf{S}^0 = \mathbf{0}$). However, this constraint artificially ‘‘pins’’ one sub-lattice in space which may be undesirable for

branch-following and bifurcation studies (for an example of this, see *Jusuf (2010)*). Other methods to remove the translation DOFs exist. Two such methods known as *Projection* and *Phantom Energy* methods are detailed in *Jusuf (2010)*.

1.1.1.5 Homogenized continuum energy density

In many SMA modeling approaches, such as those discussed in *Bhattacharya (2003)*, the internal shifts in the CB model are eliminated to obtain the *homogenized continuum* (HC) energy density that depends only on the uniform deformation \mathbf{F} . This may be physically reasonable since the dynamic evolution of the internal atomic shifts is often of a time scale that is much shorter than the evolution of \mathbf{F} . The HC energy density is defined as

$$\bar{\mathcal{E}}(\mathbf{F}, \theta) \equiv \tilde{\mathcal{E}}(\mathbf{F}, \mathbf{S}(\mathbf{F}, \theta), \theta), \quad (1.13)$$

where $\mathbf{S}(\mathbf{F}, \theta)$ is obtained implicitly by solving the shift equilibrium equations given in Eq. (1.12)₂. With the energy density defined, equilibrium configurations $(\hat{\mathbf{F}}, \hat{\theta})$ are given by extrema of the total energy density function and therefore must satisfy the following equation:

$$\left. \frac{\partial \bar{\mathcal{E}}}{\partial F_j^i} \right|_{(\hat{\mathbf{F}}, \hat{\theta})} = 0. \quad (1.14)$$

The HC energy density provides a connection between a material's state of deformation and its energy. A fundamental assumption in continuum mechanics is that, "this response is intrinsic to the material and should therefore be independent of the frame of reference used to describe the motion of the material" (*Tadmor et al. (2012)*). This requirement manifests itself primarily through the concept of *material frame-indifference and symmetry*. Frame-indifference and material symmetry are complicated subjects that involve subtle arguments regarding the nature of frames of reference. The presentation below is meant only to provide some background of the constraint on the free energy of a material. As such, many subtle details have been left out in order to present the main idea of the two requirements. A more extensive discussion on the principles of frame-indifference and material symmetry can be found in texts such as *Malvern (1969)*, *Holzappel (2000)*, and *Tadmor et al. (2012)*.

The HC energy density is required to satisfy the following two properties:

1. *Frame-indifference*

The property of frame-indifference of the energy density is argued from a physical perspective. Simply stated, the energy of the crystal remains invariant after a change of frame². That is,

$$\bar{\mathcal{E}}(\mathbf{F}, \theta) = \bar{\mathcal{E}}(\mathbf{QF}, \theta), \quad \text{for all } \mathbf{Q} \in SO(3). \quad (1.15)$$

The principle of frame-indifference is also called the principle of objectivity. It should be noted that the whole discussion is within the framework of classical mechanics and does not consider Einsteinian relativity.

2. Material symmetry

Let $\{\mathbf{e}_i \in \mathbb{R}^3 \mid i = 1, 2, 3\}$ be lattice vectors of a crystal Ω . It is reasonable to require that the choice of lattice vectors does not affect the energy of the crystal if it describes the same crystal structure. The second requirement for the energy density of a crystal is to reflect this property. Let $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ be the lattice vectors for Ω . Denote the *point group of the lattice* \mathcal{P} as follows:

$$\mathcal{P} = \{\mathbf{R} \in SO(3) \mid \mathbf{R}\mathbf{e}_i = \mu_{ij}\mathbf{e}_j, \text{ for some } \mu_{ij} \in \mathbb{Z}^{3 \times 3}, \det(\mu_{ij}) = 1\}, \quad (1.16)$$

where $\mathbb{Z}^{3 \times 3}$ is the set of all 3×3 matrices with integer entries.

Then the property of material symmetry requires that³

$$\bar{\mathcal{E}}(\mathbf{FR}, \theta) = \bar{\mathcal{E}}(\mathbf{F}, \theta), \quad \forall \mathbf{R} \in \mathcal{P}. \quad (1.17)$$

A physical interpretation for this condition says that the HC free energy reflects the fact that the properties of a crystalline solid are identical in crystallographically equivalent directions. For example, consider an experiment that measures the energy of a crystal at any given state. The crystal is initially at some reference configuration described by the deformation gradient \mathbf{I} and temperature θ . The temperature is fixed throughout this experiment. In the first case, deform the crystal so that the deformation state is described by the deformation gradient \mathbf{F} . Evaluating the energy of the

²Most discussions regarding frame-indifference interpret the application of a rigid-rotation as equivalent to a change of frame. This distinction serves to highlight the subtle arguments involved in the discussion of frame-indifference. For a more in-depth discussion, refer to *Tadmor et al. (2012)*.

³Here, the presentation of the concept of material symmetry is limited to invariance with respect to the point group of the lattice. This is only a subset of the full set of invariances. The full set include other deformations such as *lattice invariant shears*. However, these ideas will not be presented here. For a more thorough discussion on the topic, see *Bhattacharya et al. (2004)*.

crystal as in Eq. (1.13) gives $\bar{\mathcal{E}}(\mathbf{F}, \theta)$ as the value of the energy at this deformation state. After computing the energy, let the crystal return to its initial configuration described by (\mathbf{I}, θ) . Now, deform the crystal so that it is at a state described by (\mathbf{FR}, θ) , where \mathbf{R} corresponds to a deformation gradient in the point group of the crystal \mathcal{P} . This can be done in the following way: first, rotate the crystal in such a way that the rotation corresponds to the deformation gradient $\mathbf{R} \in \mathcal{P}$; the crystal is then at a state described by (\mathbf{R}, θ) . After rotating the crystal, deform the crystal by a deformation that corresponds to the deformation gradient \mathbf{F} , so that the crystal is finally at a state described by (\mathbf{FR}, θ) . Computing the energy for the crystal at this state gives the energy to be $\bar{\mathcal{E}}(\mathbf{FR}, \theta)$. By material symmetry, since \mathbf{R} is an element of the point group of the crystal, the energies computed in the two cases must be equal. Thus, $\bar{\mathcal{E}}(\mathbf{FR}, \theta) = \bar{\mathcal{E}}(\mathbf{F}, \theta)$.

1.1.1.6 Stability of crystalline solids

Two concepts arise in the discussion of the stability of crystalline solids—*dynamic stability* and *thermodynamic stability*. Dynamic stability of a configuration is defined as stability with respect to small perturbations about an equilibrium configuration of the crystal. Additionally, we shall see later in this section that stable configurations will correspond to local minima of the energy density. Thermodynamic stability considers the ability of the system to sample all of the possible dynamically stable states. A thermodynamically stable state implies a ground-state (global energy minimum). At this point, we shall restrict our attention to dynamic stability criteria.

Recall that equilibrium configurations are given as extrema in the energy density of the crystal (Eqs. (1.12) in the case of the CB energy density or Eq. (1.14) in the case of the HC energy density). These extrema may correspond to either a local maximum, local minimum, or a saddle point. The stability of a configuration is related to which of these three types of equilibrium points characterizes the configuration in question.

This section will briefly detail the notion of the dynamic stability of a crystalline solid and also the criteria that will be used to detect a change in a material's stability: Phonon stability and Cauchy-Born stability. The presentation here is meant to provide a broad overview of the stability criteria used in this work. As such, details such as their implementation will not be discussed here. For a more thorough presentation of the various notions of crystalline stability, refer to *Elliott et al. (2006a)*.

1. *Phonon stability*

The phonon stability criterion takes into account all atomic degrees of freedom and

investigates the linearized stability of a crystal under bounded perturbations of all wavelengths. This section will present a derivation of the phonon stability criterion. Recall that $\mathcal{E}(\mathbf{u}, \theta)$ is the free energy density for the finite crystal Ω where $\mathbf{u} \equiv \{\mathbf{u}_{(\ell, \alpha)} | (\ell, \alpha) \in \Omega\}$ is the vector of all atomic displacements. The equilibrium configuration $\mathring{\mathbf{u}}$ of the crystal satisfies

$$\left. \frac{\partial \mathcal{E}}{\partial u^{i, (\ell, \alpha)}} \right|_{\mathring{\mathbf{u}}} = 0 \quad \begin{cases} \ell \in \Omega, \\ \alpha = 0, 1, \dots, M-1, \\ i = 1, 2, 3, \end{cases} \quad (1.18)$$

where α indexes the sub-lattices embedded in the skeletal lattice (or equivalently the atoms in the unit cell) and i refers to the i -th spatial direction. At this point, we restrict attention to equilibrium configurations that satisfy Eq. (1.7). The stability of an equilibrium configuration of a crystal is evaluated by considering the linearized equations of motion of the system about its equilibrium configuration $(\mathring{\mathbf{u}}, \mathring{\theta})$ with periodic boundary conditions which are given by

$$m_\alpha \delta \ddot{u}^{i, (\ell, \alpha)} = - \sum_{(\ell', \alpha') \in \Omega} G^{ij} K_{jk}^{\circ(\ell, \ell', \alpha, \alpha')} \delta u^{k, (\ell', \alpha')}, \quad (1.19)$$

where G^{ij} is given by Eq. (1.3) and $K_{ij}^{\circ(\ell, \ell', \alpha, \alpha')} \equiv \left. \frac{\partial^2 nVE}{\partial u^{i, (\ell, \alpha)} \partial u^{j, (\ell', \alpha')}} \right|_{\mathring{\mathbf{u}}, \mathring{\theta}}$ is referred to as the global stiffness matrix. m_α is the mass of atom α and $(\ddot{\cdot}) = \partial^2(\cdot)/\partial t^2$ indicates two derivatives with respect to time. The linearized equations of motion given in Eq. (1.19) are a set of $3Mn \times 3Mn$ linear ordinary differential equations with exponential solutions of the form

$$\delta u^{j, (\ell, \alpha), (r)} = \Delta u^{j, (\ell, \alpha), (r)} \exp(-i\omega^{(r)}t), \quad r = 1, 2, \dots, 3Mn, \quad (1.20)$$

where $i = \sqrt{-1}$, $\Delta u^{j, (\ell, \alpha), (r)}$ is the r -th normal mode (referred to as a *phonon* in the condensed matter physics literature) in the j -th direction of atom α in unit cell ℓ . $\omega^{(r)}$ is the associated phonon frequency of the r -th normal mode. Every solution to Eq. (1.19) can be expressed as a linear combination of the $3Mn$ independent phonon modes given in Eq. (1.20). By substituting Eq. (1.20) into Eq. (1.19), it can be shown that the phonon modes and phonon frequencies are given by the eigenvectors and square roots of the eigenvalues of $G^{ij} K_{jk}^{\circ(\ell, \ell', \alpha, \alpha')}$, respectively, and as such satisfy the

following equation:

$$\sum_{(\ell', \alpha') \in \Omega} G^{ik} \overset{\circ}{K}_{kj}^{(\ell, \ell', \alpha, \alpha')} \Delta u^{j, (\ell', \alpha'), (r)} = (\omega^{(r)})^2 \Delta u^{i, (\ell, \alpha), (r)}, \quad (1.21)$$

for any $r = 1, 2, 3, \dots, 3Mn$. Further, the first three normal modes correspond to rigid-body translation modes with $\omega^{(r)} = 0$ for $r = 1, 2, 3$. The application of periodic boundary conditions helps to eliminate boundary effects and restricts normal mode vibrations to only those characteristic of the bulk material. Additionally, the use of periodic boundary conditions eliminates unbounded perturbations (such as bending and uniform deformations) from consideration in regard to stability. The crystal is considered stable if, for any $\epsilon > 0$ there exists a δ such that the solutions to the linearized equations of motion given in Eq. (1.19) satisfy

$$\|\delta \mathbf{u}^{(\ell, \alpha)}(0)\|, \|\delta \dot{\mathbf{u}}^{(\ell, \alpha)}(0)\| < \epsilon, \Rightarrow \|\delta \mathbf{u}^{(\ell, \alpha)}(t)\|, \|\delta \dot{\mathbf{u}}^{(\ell, \alpha)}(t)\| < \delta, \quad (1.22)$$

for all $t > 0$, where $\delta \mathbf{u}^{(\ell, \alpha)}(0)$ and $\delta \dot{\mathbf{u}}^{(\ell, \alpha)}(0)$ are displacements and velocities at time $t = 0$ and $\|\cdot\|$ is any norm for $\mathbf{u}_{(\ell, \alpha)}$. An equivalent condition for stability of the conservative system under consideration is: *solutions will remain bounded if and only if all eigenvalues of the global stiffness matrix $G^{ik} \overset{\circ}{K}_{kj}^{(\ell, \ell', \alpha, \alpha')}$ are positive.* Since the eigenvalues of $G^{ik} \overset{\circ}{K}_{kj}^{(\ell, \ell', \alpha, \alpha')}$ are the squares of the phonon frequencies, the crystal is phonon stable if

$$(\omega^{(r)})^2 > 0, \quad r = 4, 5, \dots, 3Mn. \quad (1.23)$$

Generally, diagonalization of the global stiffness matrix $G^{ik} \overset{\circ}{K}_{kj}^{(\ell, \ell', \alpha, \alpha')}$ to determine the phonon frequencies is a time consuming computation due to the large dimension of the system. Fortunately, the periodicity conditions imply that the stiffness matrix is “block-circulant” (Davis (1994)) which ensures that a block-diagonalization to $3M \times 3M$ blocks exists which greatly simplifies the calculations to obtain the phonon frequencies. It can be shown (see for example, Elliott *et al.* (2006a)) that the $3M \times 3M$ dynamical stiffness matrix $\mathbb{K}_p^j(\mathbf{k}, \alpha, \alpha')$ is given by

$$\mathbb{K}_p^j(\mathbf{k}, \alpha, \alpha') = (m_\alpha m_{\alpha'})^{-1/2} \sum_{\ell' \in \mathbb{Z}^3} G^{jk} K_{kp}^{(0, \ell', \alpha, \alpha')} \exp \{ -i\mathbf{k} \cdot (\mathbf{X}_{(\ell', \alpha')} - \mathbf{X}_{(0, \alpha)}) \}. \quad (1.24)$$

The *wavevector* \mathbf{k} is a vector in reciprocal space and in this context describes a per-

turbation in a particular crystalline direction with a wavelength that is related to the magnitude of the wavevector. Often the terms, “zone boundary” and “gamma point” are used to indicate wavevectors on the boundary (e.g, $\mathbf{k} = \frac{1}{2}(110)_{\text{cubic}}$) and origin ($\mathbf{k} = 0$) of the first Brillouin zone, respectively. Physically, wavevectors on the zone boundary describe the shortest wavelength perturbations of the crystal whereas wavevectors near the gamma point describe perturbations in the long wavelength limit.

The Phonon stability of a crystal can be evaluated using Eq. (1.24) by first choosing a wavevector \mathbf{k} and then obtaining the eigenvalues of the dynamical stiffness matrix $\mathbb{K}_p^j(\mathbf{k}, \alpha, \alpha')$. The stability criterion then states that the crystal configuration is phonon stable if (and only if) all the eigenvalues of $\mathbb{K}_p^j(\mathbf{k}, \alpha, \alpha')$ are positive for every wavevector \mathbf{k} in the first Brillouin zone. Of special interest is when a perturbation with a particular wavevector yields an imaginary (negative squared frequency) eigenvalue. Then the reference crystal is dynamically unstable when subjected to perturbations with a wavelength and direction given by the corresponding wavevector. It can be shown that in order to become *phonon stable*, the crystal may undergo a phase transition into a configuration described by a combination of the reference state and the normal mode associated with the unstable phonon mode. These configurations will in general break the translational symmetry of the primitive unit cell of the reference crystal. This means that a new definition of the primitive unit cell is necessary to describe the deformed crystal. Additionally, the number of atoms in the deformed unit cell will contain an integer multiple of the number of atoms in the reference unit cell⁴.

The result of a phonon frequency calculation is often given in terms of the phonon dispersion relation. The phonon dispersion relations of a crystal describe the frequency response of the crystal to plane wave perturbations over a range of wavelengths, in a particular crystalline direction. Fig. 1.11 shows an example of a phonon dispersion relation obtained from *Minkiewicz et al. (1967)*. The quantity “phonon energy” in the figure can be translated to frequency (refer to *Minkiewicz et al. (1967)*) for more details). In principle, to ensure phonon stability, Eq. (1.10) must be computed for an infinite number of wavevectors (in all directions at all wavelengths). However, for practical reasons, only a finite number need to be calculated due to the continuity of the curves describing the frequency response.

2. Cauchy-Born stability

⁴The exact multiple is associated to the wavelength of the unstable normal mode

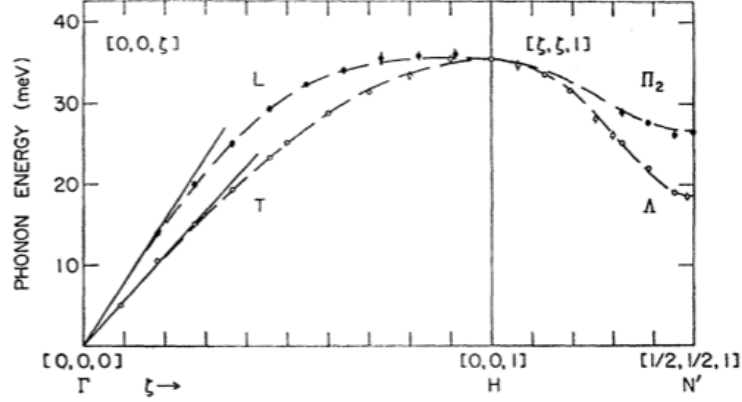


Figure 1.11: The room-temperature phonon dispersion curves of α -Fe along the $[0, 0, \zeta]$ and $[\zeta, \zeta, 1]$ directions.

The Cauchy-Born (CB) stability criterion is based on local minimization of the CB energy density function (Eq. (1.11)) and addresses the stability with respect to all CB kinematic compatible perturbations. Refer to *Elliott et al. (2006a)* for a discussion on the overlap and comparison between the phonon-stability and CB-stability criteria. Recall that in the CB model the current atomic positions are given in terms of the uniform deformation \mathbf{F} of the crystal and the “internal shifts” \mathbf{S} of the atoms and are given in Eq. (1.7). An equilibrium configuration $(\mathring{\mathbf{F}}, \mathring{\mathbf{S}})$ is CB-stable if it is a local minimum of the energy density $\tilde{\mathcal{E}}(\mathbf{F}, \mathbf{S}, \theta)$. Equivalently, $(\mathring{\mathbf{F}}, \mathring{\mathbf{S}})$ is CB-stable if the second derivative of $\tilde{\mathcal{E}}$ is positive definite with respect to all non-zero symmetric right stretch tensor perturbations $\delta\mathbf{U}$ and all non-zero internal shift perturbations $\delta\mathbf{S}$. That is,

$$[\delta\mathbf{U}, \delta\mathbf{S}] \begin{bmatrix} \mathbf{R}^\top \frac{\partial^2 \tilde{\mathcal{E}}}{\partial \mathbf{F} \partial \mathbf{F}} \mathbf{R} & \mathbf{R}^\top \frac{\partial^2 \tilde{\mathcal{E}}}{\partial \mathbf{F} \partial \mathbf{S}} \\ \frac{\partial^2 \tilde{\mathcal{E}}}{\partial \mathbf{S} \partial \mathbf{F}} \mathbf{R} & \frac{\partial^2 \tilde{\mathcal{E}}}{\partial \mathbf{S} \partial \mathbf{S}} \end{bmatrix}_{\mathring{\mathbf{F}}, \mathring{\mathbf{S}}} \begin{bmatrix} \delta\mathbf{U} \\ \delta\mathbf{S} \end{bmatrix} > 0, \quad \forall (\delta\mathbf{U}, \delta\mathbf{S}) \neq \mathbf{0}, \quad (1.25)$$

where $\delta\mathbf{U}$ is a symmetric right stretch tensor perturbation and $\mathbf{F} = \mathbf{R}\mathbf{U}$. This implies that an equilibrium configuration $(\mathring{\mathbf{F}}, \mathring{\mathbf{S}})$ is stable if the eigenvalues of the matrix in Eq. (1.25) are strictly greater than zero (after those eigenvalues associated with rigid-body motion are eliminated).

1.1.1.7 Multi-welled SMA energy landscape

An *energy landscape* refers to the surface defined by the energy density function. Typical modeling of SMAs involves defining a free energy landscape that is “multi-welled”. The term multi-well refers to the feature of the free energy where at certain temperatures, the

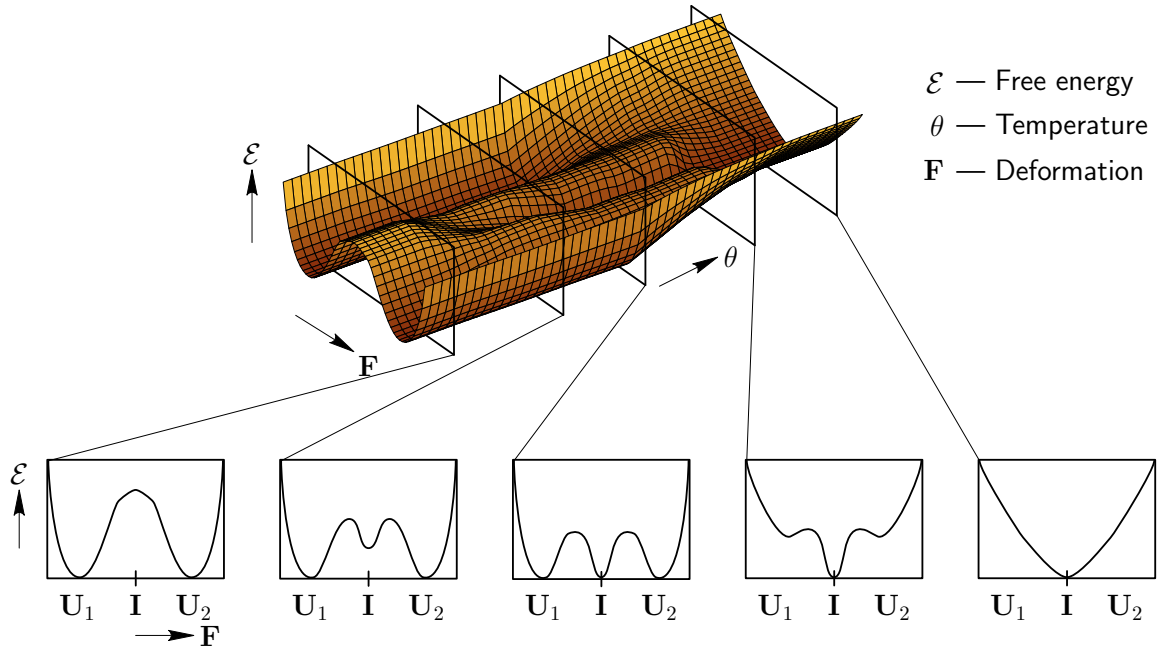


Figure 1.12: A schematic of an energy landscape for a SMA (figure obtained from *Elliott (2007)*). The actual landscape of a typical SMA would be much more complex.

free energy may consist of multiple (and therefore non-unique) dynamically stable equilibria (local minima). The basis for the assumption regarding the multi-welled nature of the free energy landscape of SMAs is the experimental observation of distinct crystal structures of the material⁵. The observation of multiple crystal structures in the material would imply that the structures are dynamically stable and hence are local minima of the energy landscape. With the existence of distinct local minima on the energy landscape established, the picture for the free energy landscape of SMAs is completed by considering the thermodynamic stability of the structures as the temperature is varied. In order to facilitate this discussion, consider a schematic of the free energy landscape for a SMA, given in Fig. 1.12 (obtained from (*Elliott (2007)*)). Note that this schematic greatly simplifies the actual energy landscape of a typical SMA and is presented here only as an illustrative tool. In the figure, \mathcal{E} , \mathbf{F} , and θ represent the free energy density, deformation gradient, and temperature, respectively. The variation of the free energy with temperature is shown at the bottom of the figure where \mathbf{I} , \mathbf{U}_1 , and \mathbf{U}_2 correspond to the austenite phase and two martensite variants (discussed later in this section). At a fixed temperature, the equilibrium deformation states (\mathbf{I} , \mathbf{U}_1 , and \mathbf{U}_2) correspond to extrema in the free energy landscape as shown at the bottom of Fig. 1.12.

⁵The distinct crystal structures are the austenite and variants of martensite above and below the transition temperature, respectively.

Here, a *thermodynamically stable* configuration (at a given temperature) is defined as a structure that is a global minima in energy (i.e., a ground-state structure). As such, the thermodynamically stable states will be those states most likely to manifest themselves in the thermodynamic limit (an implicit assumption here is that the system has sufficient time to sample all possible local minima). It is important to note that while thermodynamically stable states are often experimentally observed, the converse may not be true. That is, experimentally observed states may not correspond to global minima in energy as meta-stable states of higher energy may be observed in experiments. The mechanism for the stabilization of meta-stable states may vary and this issue will be revisited in the discussion of *modulated martensites* later in this chapter. For now, assume that a thermodynamically stable state implies an experimentally observable state *and vice versa*.

Now, assume that there is a transition temperature θ_c where the material will transition from austenite to martensite. Then for $\theta > \theta_c$, the austenite phase corresponds to a global minima of the free energy density. Similarly, for $\theta < \theta_c$, the martensite is a global minimum of the free energy density of the material. Moreover, since the austenite configuration is a global minimum at $\theta > \theta_c$ and the martensite configuration is a global minima at $\theta < \theta_c$, then at the transition temperature, the two configurations are both global minima of the energy (in fact, the transition temperature can be defined such that it is the temperature where the energies of the austenite and martensites are equal). For this reason, two “wells” of equal energy in the landscape must exist at the transition temperature. At this point, the broad picture of the free energy landscape of SMAs is clear—a free energy landscape for a SMA will exhibit minima that evolve to reflect the ability of the material to transition from one observed configuration to another as a function of the thermo-mechanical loading.

1.1.1.8 A more detailed description of a SMA energy landscape

It is important to note that Fig. 1.12 greatly simplifies the nature of the energy landscape and is presented only to facilitate the discussion regarding the multi-welled nature of a SMA energy landscape. The remaining discussion in this subsection will present a more realistic depiction of some aspects of the energy landscape starting from the HC free energy and will follow the presentation given in *Bhattacharya (2003)*. The requirement that the free energy is frame-indifferent results in a more complicated characterization of the wells in the energy landscape than that which is depicted in Fig. 1.12. Furthermore, to describe MPTs in their simplest setting, only homogenous perfect crystals of shape memory material will be considered. In this manner, only perfect crystal configurations and their energies need to be taken into account.

Assume that a crystalline material Ω is describable through a HC free energy density as defined in Eq. (1.13). Additionally, assume that the reference configuration of the crystal at the transition temperature θ_c is the high-symmetry austenite phase described by lattice vectors $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$, with free energy given by $\mathcal{E}(\mathbf{I}, \theta)$, where $\mathbf{I} \in \mathbb{R}^{3 \times 3}$ is the identity tensor. Since austenite is a global minimum at temperatures greater than the transition temperature, then the HC energy of the crystal must satisfy,

$$\bar{\mathcal{E}}(a(\theta)\mathbf{I}, \theta) \leq \bar{\mathcal{E}}(\mathbf{F}, \theta), \quad \forall \mathbf{F} \in \mathbb{R}^{3 \times 3}, \quad \det(\mathbf{F}) > 0, \quad \theta \geq \theta_c, \quad (1.26)$$

where $a(\theta)\mathbf{I}$ describes the deformation of the austenite allowing for thermal expansion⁶. Since the energy is frame-indifferent,

$$\bar{\mathcal{E}}(a(\theta)\mathbf{I}, \theta) = \bar{\mathcal{E}}(a(\theta)\mathbf{R}, \theta), \quad \forall \mathbf{R} \in SO(3). \quad (1.27)$$

This implies that not only is the austenite with deformation gradient $a(\theta)\mathbf{I}$ a global minimum but any deformations with deformation gradients $a(\theta)\mathbf{R} \in SO(3)$ are also global minima. This is in contrast to Fig. 1.12, where only one unique minimum associated to the austenite phase exists. However, if we identify the set of deformations that are different only by rigid-body rotations \mathbf{R} as equivalent deformations, then the set of deformations can be partitioned into disjoint equivalence classes. Furthermore, if we choose the deformations that have $\mathbf{R} = \mathbf{I}$ as representatives of each equivalence class, then by the polar decomposition theorem, this set of deformations will belong to the set of positive definite, symmetric *right stretch tensors* $\{\mathbf{U} \in \mathbb{R}^{3 \times 3} \mid \det(\mathbf{U}) > 0, \mathbf{U} = \mathbf{U}^\top\}$. In this manner, an interpretation of Fig. 1.12 is that it is a depiction of the energy landscape as a function of right stretch tensor and temperature. If the rotations are included in the depiction of the energy landscape, a pictorial depiction of the energy wells can be given by Fig. 1.13. This visualization is similar to that found in *Bhattacharya (2003)*. In the figure, the red circle indicates states of minimum energy (that is, energy minimizing deformation gradients). The identity deformation gradient \mathbf{I} is indicated as the dot which lies on the circle because of Eq. (1.26). Deformations corresponding to rotations must also lie on the circle because of Eq. (1.27).

Let \mathcal{P}_A be the point group of the austenite phase. As in *Bhattacharya (2003)*, the austenite structure is henceforth assumed to have strictly greater symmetry than the martensite structure. That is, if $(\mathbf{e}'_1, \mathbf{e}'_2, \mathbf{e}'_3)$ are lattice vectors for the martensite structure, then the point group of the martensite \mathcal{P}_M is assumed to be a proper subgroup of the point group

⁶More precisely, $a(\theta) \in \mathbb{R}$ is a function of temperature. That is, $a(\theta) \in [1, \infty)$ for $\theta \geq \theta_c$.

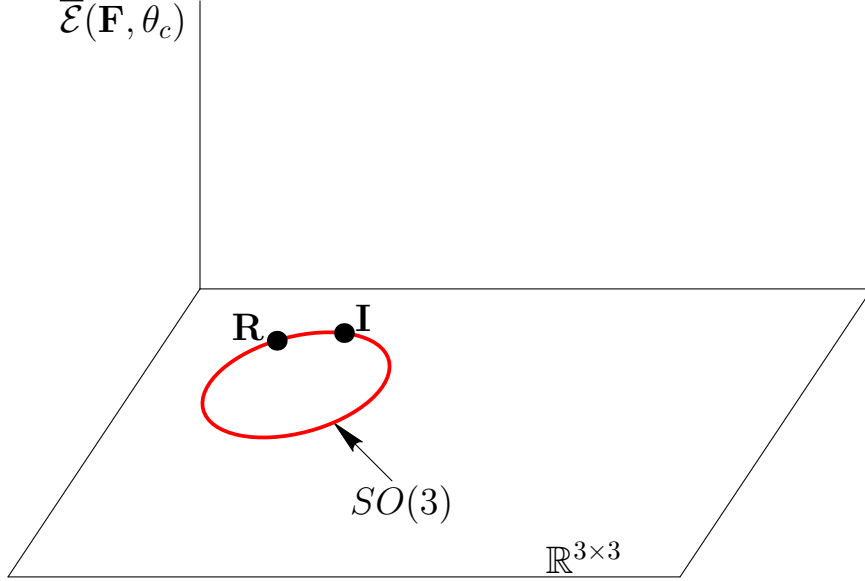


Figure 1.13: A schematic of the energy well showing the austenite phase as a local minimum state. Here, \mathbf{I} is the identity map. $SO(3)$ is schematically described by the circle. Frame-indifference implies that $SO(3)$ must also be states of local minimum energy.

of the austenite \mathcal{P}_A . As was argued in the beginning of this subsection, the fundamental postulate of materials undergoing MPTs (see for example, *Toledano and Toledano (1987)* and *Cowley (1980)*) is that there exists another state of minimum energy corresponding to a martensite structure at $\theta = \theta_c$. That is, there exists $\mathbf{U}_1 \neq \mathbf{I}$ such that

$$\bar{\mathcal{E}}(\mathbf{U}_1, \theta_c) = \bar{\mathcal{E}}(\mathbf{I}, \theta_c) \leq \bar{\mathcal{E}}(\mathbf{F}, \theta_c), \quad \forall \mathbf{F} \in \mathbb{R}^{3 \times 3}, \quad \det(\mathbf{F}) > 0. \quad (1.28)$$

Since the material transitions from austenite to martensite at $\theta < \theta_c$, then it must be that $\mathbf{U}_1 \in \mathbb{R}^{3 \times 3}$ is a deformation gradient such that $\mathbf{e}'_i = \mathbf{U}_1 \mathbf{e}_i$, for $i = 1, 2, 3$. That is, \mathbf{U}_1 corresponds to the deformation gradient that transforms the austenite lattice vectors into the martensite lattice vectors. This postulate results in a multi-well energy landscape where at $\theta = \theta_c$, there now exists another minimum at \mathbf{U}_1 (and $\mathbf{R}\mathbf{U}_1$ due to frame-indifference) corresponding to the martensite well as depicted in Fig. 1.14. Furthermore, since $\mathbf{U}_1 \neq \mathbf{I}$, then it is easy to verify that $\mathbf{R}\mathbf{U}_1 \notin SO(3)$, for any $\mathbf{R} \in SO(3)$. Therefore, the well containing \mathbf{U}_1 must lie on a disjoint circle.

Symmetry of the austenite structure implies the existence of *martensite variants* with deformation gradients \mathbf{U}_i , $i > 1$. This is due to the property of material symmetry that implies that there must also be wells at $\mathbf{U}_j = \mathbf{U}_1 \mathbf{R}_j$, where $\mathbf{R}_j \in \mathcal{P}_A$ (*Bhattacharya (2003)*). These are the remaining variants of the martensite (*James and Kinderlehrer (1989)*). However,

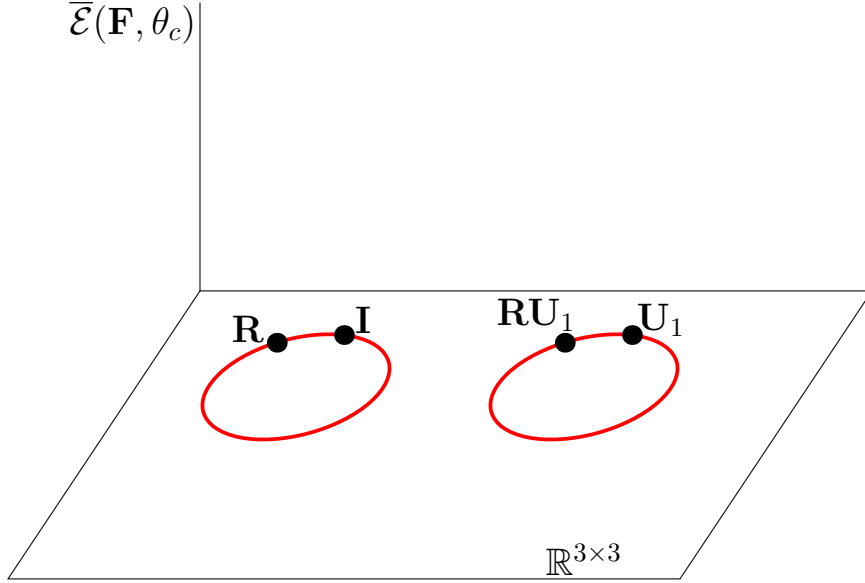


Figure 1.14: A schematic showing the multi-well nature of the energy landscape at the transition temperature. Here \mathbf{U}_1 is the deformation gradient that takes the lattice vectors describing the austenite structure to the lattice vectors describing the martensite structure.

not all \mathbf{R}_j will give a new \mathbf{U}_j . The number of distinct \mathbf{U}_j (and therefore the number of martensite variants) is given by $|\mathcal{P}_A|/|\mathcal{P}_M|$ where $|\mathcal{P}_A|$ is the order of the austenite point group and $|\mathcal{P}_M|$ is the order of the martensite point group.

For example, consider a hypothetical phase transformation involving a high-symmetry austenite structure described by a square two-dimensional lattice into a low-symmetry martensite structure described by a rectangular two-dimensional lattice. The unit cell of the square lattice is shown on the left hand side of Fig. 1.15 marked by A . It is easy to verify that the point group of a square has four elements, whereas the point group of a rectangular structure has two elements giving two martensite variants M_1 and M_2 with unit cells shown on the right side of Fig. 1.15. Note that the rectangular martensite structures labeled M_1 and M_2 can be formed from the square lattice by two deformations \mathbf{U}_1 and \mathbf{U}_2 that are only different by an element in the point group of A (a rotation by 90 degrees). Due to material symmetry, the variants M_1 and M_2 are equal in energy. Thus, at the transition temperature the energy landscape consists of three wells, as depicted in Fig. 1.16. This figure, like in Fig. 1.13 and Fig. 1.14 shows minimum energy deformations. Note that in Figs. 1.13–1.16 the dependence of the energy as a function of temperature is not shown. Instead, consider Fig. 1.16 as the energy minima plotted as a function of deformation gradient at the transition temperature.

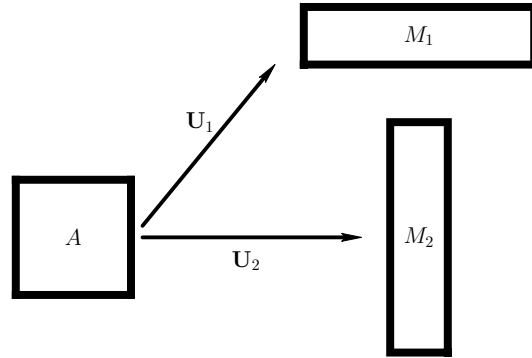


Figure 1.15: A schematic of a two-dimensional square lattice austenite (A) phase transformed into two martensitic variants (M_1 and M_2) via deformations \mathbf{U}_1 and \mathbf{U}_2 .

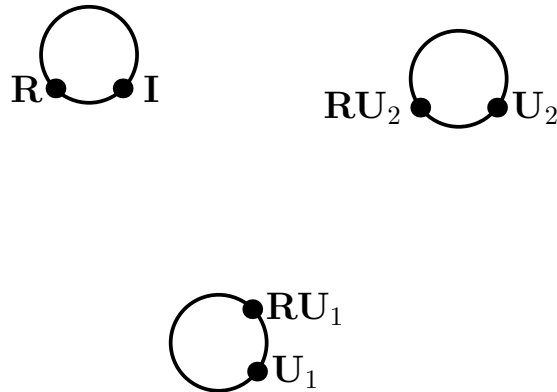


Figure 1.16: A schematic of the energy wells at the transition temperature accounting for frame-indifference. The solid lines denote states of minimum energy where \mathbf{I} , \mathbf{U}_1 , and \mathbf{U}_2 represent deformation gradient matrices that take the austenite phase to itself and the two martensitic variants, respectively. \mathbf{R} denotes a rigid-rotation.

1.1.2 Micro-scale description

In the previous section a nano-scale description of a typical SMA was discussed. Starting from a lattice-level description of a SMA, a homogenous continuum free energy density $\bar{\mathcal{E}}(\mathbf{F}, \theta)$ was derived. The HC energy then defines an energy landscape that describes the dependence of the energy on the deformation gradient \mathbf{F} and temperature θ . The multi-well characteristic feature of an SMA energy landscape was also discussed where at the transition temperature the landscape consists of energy minimum deformations corresponding to the austenite structure and martensite variants.

In this section, we will consider the behavior of SMAs at the micro-scale level. At this scale the material is treated as a continuum and microstructures consisting of combinations of the austenite phase and martensite variants are described.

The macroscopic behavior of SMAs is intimately tied to the fine needle-like martensite micro-scale structures (henceforth referred to as *microstructure*) that develop in these materials and a large literature now exists to help explain phenomena at this scale. Successful continuum approaches that take into account the different martensite variants and their possible spatial arrangements have been investigated by *James* (1986), *Bhattacharya* (1991), and *Hane and Shield* (1999). These methods are based on a global energy minimization approach that explains the details of SMA microstructures.

1.1.2.1 Twinning

First, consider the microstructure labeled “Twinned-M” in Fig. 1.1. This microstructure is comprised of martensite variants separated by planar interfaces and is one of the possible microstructures of a SMA material (*Otsuka and Wayman* (1999)). According to the theory of martensite (see for example, *Bhattacharya* (2003)) the microstructures arise from the material’s ability to accommodate continuous deformations while allowing for jumps in deformation gradients. The following discussion will briefly present the features of this theory.

With no loss of generality, assume that the material has only two martensite variants M_1 and M_2 with deformation gradients \mathbf{U}_1 and \mathbf{U}_2 . Note that since $\theta < \theta_c$, then the energy landscape has only two wells at \mathbf{U}_1 and \mathbf{U}_2 (and hence $\mathbf{R}\mathbf{U}_1$ and $\mathbf{R}\mathbf{U}_2$ due to frame-indifference).

A simplified representation of a twinned martensite microstructure is given in the left hand side of Fig. 1.17 where the right hand side of the figure is a magnified image of the region shaded in gray. This microstructure is said to be comprised of *twinned martensite* where

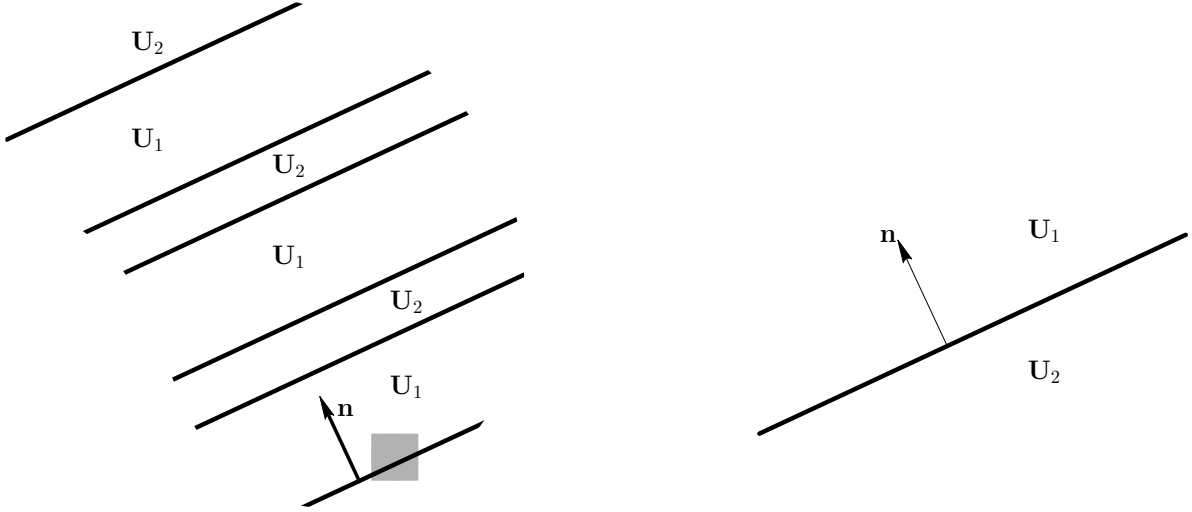


Figure 1.17: A deformation involving two variants of martensite separated by an interface with normal \mathbf{n} that is defined through Eq. (1.29).

a *twin* is defined as a planar defect in a crystal such that the lattice on one side of the interface can be obtained by a rotation or simple shear of the lattice on the other side of the interface (*Bhattacharya (2003)*). Now, consider the interface having unit normal \mathbf{n} shown in the enlarged segment of Fig. 1.17. The deformation $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is continuous with deformation gradient $\nabla^+ \phi(\mathbf{X}) = \mathbf{U}_1$ and $\nabla^- \phi(\mathbf{X}) = \mathbf{U}_2$ above and below the interface, respectively. Equating the tangential derivatives at the interface leads to the *Hadamard jump condition* (*Wechsler (1954)*) given by

$$\mathbf{R}_2 \mathbf{U}_2 - \mathbf{R}_1 \mathbf{U}_1 = \mathbf{a} \otimes \mathbf{n}, \quad (1.29)$$

for some rotation matrices \mathbf{R}_1 and \mathbf{R}_2 in $SO(3)$ and vector $\mathbf{a} \in \mathbb{R}^3$. Eq. (1.29) is henceforth referred to as the *twinning equation* and reflects the kinematic condition that requires the body to remain unbroken (i.e., continuous deformation mapping). Note that, the condition allows for a jump in deformation gradient. However, for the body to remain unbroken, the jump in the deformation gradient cannot be arbitrary and must in fact be described by Eq. (1.29) with a planar interface with normal \mathbf{n} . Fig. 1.18 is a representation of this condition. Recall that the energy wells are represented as circles with \mathbf{U}_1 , \mathbf{U}_2 as deformation gradient matrices that take the austenite lattice vectors to the martensite variant lattice vectors. In the figure, \mathbf{R}_1 and \mathbf{R}_2 are rotation matrices satisfying Eq. (1.29). A straight line joins two wells if an interface can be formed (i.e., if a solution to Eq. (1.29) exists). One question that may arise here is whether it is possible for a deformation to occur that involves the austenite phase and a single martensite variant separated by an interface.

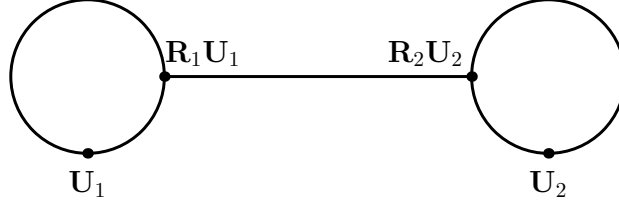


Figure 1.18: A schematic representation of a twin. The energy wells of two martensite variants is shown with \mathbf{U}_1 and \mathbf{U}_2 as the deformation gradient matrices that take the austenite phase to the martensite variants M_1 and M_2 , respectively. \mathbf{R}_1 and \mathbf{R}_2 are rotation matrices satisfying the twinning equation (Eq. (1.29)) (Figure obtained from *Bhattacharya (2003)*).

Ball and James (1987) showed that it is not generally possible to form an exact austenite-martensite interface in most martensitic materials since the kinematic compatibility would not generally be satisfied.

1.1.2.2 Energy minimizing deformations for uniform displacement boundary conditions

In the preceding subsection the twinning equation that describes a twinned microstructure is explained as a consequence of the Hadamard jump condition, which is a kinematic condition that allows for a jump in deformation gradient while maintaining the requirement that the body remain unbroken. The theory also predicts the unlikelihood of a material to form an interface between the austenite phase and a single martensite phase. This corresponds well with experimental observation involving twinned martensites.

In this section, the effect of boundary condition on the microstructure of SMAs will be presented. The essential idea is that given a general boundary condition, the material can accommodate the boundary condition while at the same time keeping its energy at or close to the minimum by forming a deformation composed of a mixture of the different energy wells.

The presentation of the theory here will follow that detailed in *Ball and James (1987)*. The main question addressed by *Ball and James (1987)* is the following: can a microstructure comprising of twinned martensites be formed while keeping the energy at a minimum? To illustrate the formation of fine mixture of martensite, consider an example involving a special deformation gradient of the form

$$\mathbf{F}_\lambda = \lambda \mathbf{U}_1 + (1 - \lambda) \mathbf{U}_2, \quad \lambda \in [0, 1], \quad (1.30)$$

where \mathbf{U}_1 and \mathbf{U}_2 are energy minimizing deformation gradients (corresponding to martensite variants) of a homogenous perfect crystal Ω at $\theta < \theta_c$. Assume that the homogenous

perfect crystal Ω of shape memory material is described using a HC free energy density $\bar{\mathcal{E}}(\mathbf{F}, \theta)$. With no loss of generality, assume that $\bar{\mathcal{E}}(\mathbf{U}_j, \theta) = 0$, for $j = 1, 2$. To simplify the discussion, let the boundary condition imposed on the boundary $\partial\Omega$ be a uniform boundary condition $\mathbf{x} = \mathbf{F}_\lambda \mathbf{X}$, for $\mathbf{X} \in \partial\Omega$. Here, the analysis can be separated into two cases: case (1) $\lambda = 0$ or $\lambda = 1$ and case (2) $\lambda \in (0, 1)$. To minimize the energy for case (1), the material assumes the corresponding variant deformation since $\bar{\mathcal{E}}(\mathbf{U}_j, \theta) = 0$ for $j = 1, 2$. These cases correspond to the detwinned martensite microstructure shown in Fig. 1.1 labeled “Detwinned-M”. In case (2), $\lambda \in (0, 1)$ and hence

$$\bar{\mathcal{E}}(\mathbf{F}_\lambda, \theta) > \bar{\mathcal{E}}(\mathbf{U}_j, \theta) = 0, \quad j = 1, 2, \quad (1.31)$$

since by assumption, $\mathbf{R}\mathbf{U}_j$, $j = 1, 2$ are the only energy minimizing deformation wells at $\theta < \theta_c$. Therefore, if the deformation of Ω corresponds to the homogenous deformation corresponding to the boundary value, (i.e., $\mathbf{F}_\lambda \mathbf{X}$, for all $\mathbf{X} \in \Omega$), then it must be that the total energy of the body is strictly greater than zero. This can be seen from the following calculation:

$$\text{Total Energy} = \int_{\Omega} \bar{\mathcal{E}}(\mathbf{F}_\lambda, \theta) dV = \text{Vol.}(\Omega) \bar{\mathcal{E}}(\mathbf{F}_\lambda, \theta) > 0. \quad (1.32)$$

However, the energy of the body can be minimized (therefore obtaining the solution to the minimization boundary value problem subject to this particular boundary condition) by making a fine scale mixture of \mathbf{U}_1 and \mathbf{U}_2 in the following manner:

First, for any positive integer n , construct the deformation ϕ^n shown in Fig. 1.19, where \mathbf{n} is obtained from the twinning equation and $\lambda \in [0, 1]$ is obtained from the boundary condition in Eq. (1.30). Furthermore, it is easy to show that since $n \geq 1$, the spacing between the interfaces decreases with increasing n . This means that as $n \rightarrow \infty$, the lamellas consisting of the martensite variants become finer and finer. In this manner, ϕ^n approaches the homogenous deformation gradient \mathbf{F}_λ as⁷ $n \rightarrow \infty$. Additionally, the energy is kept at a minimum at every n since the deformation is composed of energy minimizing deformations. However, for any fixed n , the boundary condition is never satisfied exactly and instead is always jagged. In order to remedy this issue, an *interpolation layer* is introduced as shown in gray in Fig. 1.20.

Since by definition the gradient in the interpolation layer is not \mathbf{U}_1 or \mathbf{U}_2 , the energy of the layer is not zero. However, it can be shown that $t \propto \frac{1}{n}$ where t is the thickness of the

⁷This statement uses the word “approaches” in a formal way. To be precise, *Ball and James* (1987) showed that ϕ^n converges to \mathbf{F}_λ weakly.

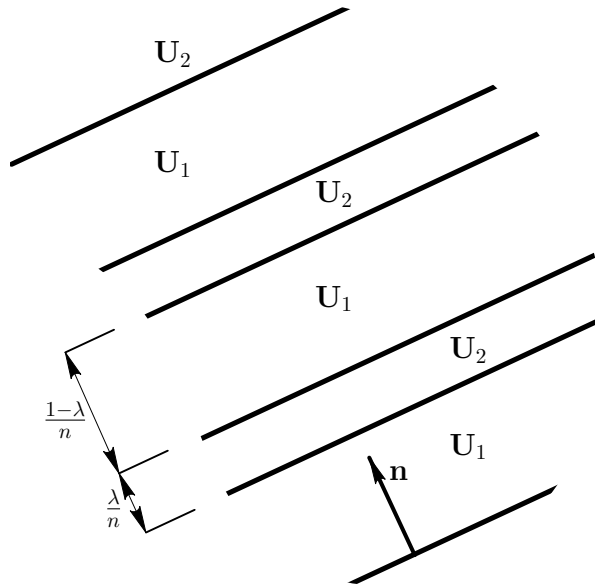


Figure 1.19: A deformation consisting of alternating gradients (from *Bhattacharya* (2003)).

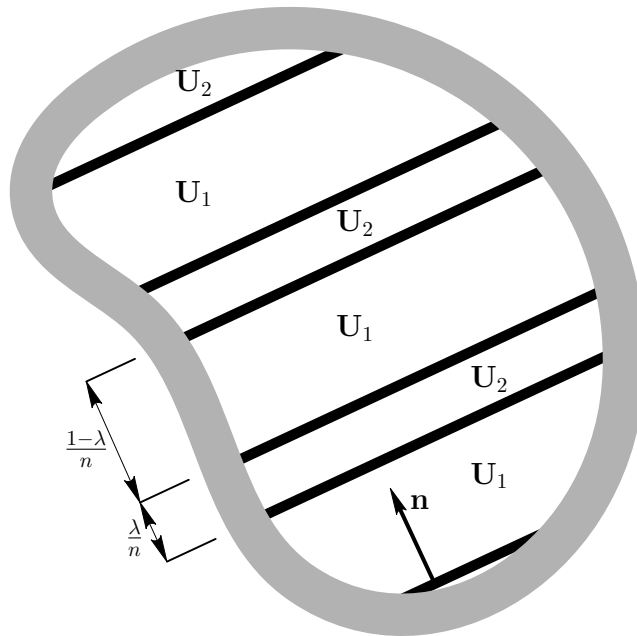


Figure 1.20: A deformation with alternating gradients and an interpolation layer shown in gray.

interpolation layer. Therefore, as n becomes large, the total energy tends to zero. Even in this simple example we see that by configuring the crystal to deformations consisting of a fine-scale mixture of alternating layers of \mathbf{U}_1 and \mathbf{U}_2 with phase fraction λ , the boundary condition is satisfied and the energy is kept low⁸ in the limit of $n \rightarrow \infty$.

The fineness n is a measure of the ‘closeness’ to minimum that the energy will theoretically be. As n becomes large the energy of the system tends to its infimum. Note that in theory the microstructure will be infinitely fine. In reality there is a limit to the fineness of the microstructure. However, the answer to why the microstructure is not observed to be infinitely fine has not been completely determined (*Bhattacharya (2003)*). One possible answer to this question is the introduction of an interfacial energy in the twin-boundary (*Semenovskaya and Khachatryan (1992)*, *Ball and James (1987)*) in addition to the bulk energy. The total energy of the system then is a sum of the bulk elastic energy and the total interfacial energy (which are shown to be proportional to n^{-1} and n , respectively, in *Bhattacharya (2003)*). Therefore, as n becomes large, the bulk energy goes down but the interfacial energy goes up. This balancing of the bulk and interfacial energy determines the length scale of the minimizing microstructure.

1.1.2.3 Austenite-martensite interface

In the high temperature phase the SMA will completely be comprised of its high symmetry austenite phase. At low temperatures the microstructure is comprised of twinned (or detwinned) martensite depending on the boundary condition. This subsection will present an explanation for the last microstructure that was mentioned regarding Fig. 1.1. This microstructure is called the austenite-martensite interface or habit plane and is schematically shown in Fig. 1.21.

Consider a typical microstructure of a SMA at the transition temperature shown in Fig. 1.21. In this figure, three crystal structures are shown with the austenite phase on the left hand side and two martensite variants on the right hand side. The austenite phase is explicitly assumed to be the reference configuration and so \mathbf{I} , \mathbf{U}_1 , and \mathbf{U}_2 denote the deformation gradients that take the austenite lattice vectors to themselves and to the two martensite variants, respectively. $\lambda \in [0, 1]$ corresponds to the phase fraction of each martensite variant while \mathbf{m} and \mathbf{n} define the austenite-martensite interface normal, and martensite variants interface normal, respectively. As mentioned in the previous section, it is not possible to form an exact austenite-martensite interface in most martensitic materials. Instead, one observes an

⁸Interestingly, while the resulting deformation is shown to be an energy minimizing deformation, it is not a state of equilibrium (*Healey and Miller (2007)*).

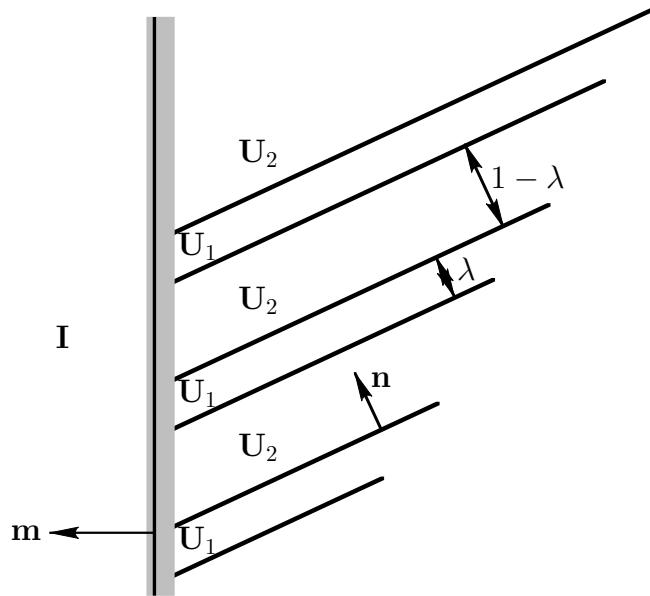


Figure 1.21: A typical austenite-martensite interface. I , U_1 , and U_2 represent the austenite and the two martensite variants, respectively. $\lambda \in [0, 1]$ corresponds to the phase fraction of each martensite variant, while \mathbf{m} and \mathbf{n} define the austenite-martensite interface normal and twinned martensite interface normal, respectively.

interface such as that shown in Fig. 1.21 (shown in gray in the figure) where an austenite phase is separated by an interface from fine twins of variants of martensite. Furthermore, it is observed from experiments (*Sun et al. (1999)*) that the interface is not sharp.

The main idea behind the explanation for the austenite-martensite interface is to first consider the uniform deformation corresponding to the austenite structure to be the boundary condition on the boundary of the martensite domain. The martensite variants on the right-hand side of the barrier then form a microstructure as described in the previous subsection. The balance of the interfacial energy on the twin boundaries and the bulk energy will then determine the length-scale of the lamellas which will limit the fineness of the twins. Since the lamellas will not be infinitely fine, the boundary condition will not be satisfied exactly and an interpolation region is introduced (shown in gray in Fig. 1.21). *Ball and James (1987)* showed that the deformation gradients in this interpolation region do not cause the energy to blow up if and only if: (1) the deformations I , U_1 , and U_2 belong to the energy wells and (2) the average deformations on both sides are compatible. That is, the deformation in Fig. 1.21 is shown to be a member of an energy minimizing sequence of

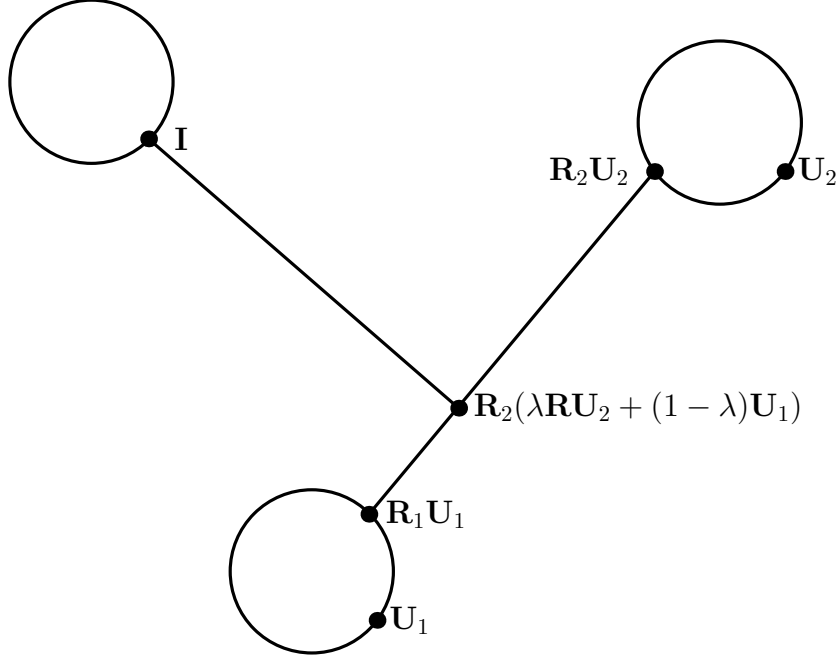


Figure 1.22: Schematic representation depicting solution to the austenite-martensite interface equation (from (Bhattacharya (2003))).

deformations if

$$\mathbf{R}\mathbf{U}_2 - \mathbf{U}_1 = \mathbf{a} \otimes \mathbf{n} \quad (1.33)$$

$$\mathbf{R}_2(\lambda\mathbf{R}\mathbf{U}_2 + (1 - \lambda)\mathbf{U}_1) = \mathbf{I} + \mathbf{b} \otimes \mathbf{m}, \quad (1.34)$$

for some rotations $\mathbf{R}_1, \mathbf{R}_2$, $\mathbf{R} = \mathbf{R}_2^\top \mathbf{R}_1$, some vectors $\mathbf{a}, \mathbf{b}, \mathbf{m}$, and \mathbf{n} , and some scalar $\lambda \in [0, 1]$. The microstructures that satisfy the above equations are said to have a *rank-one connection between the austenite deformation gradient and the average deformation gradient of the twinned martensite*. Equations (1.33) and (1.34) are known as the *austenite-martensite interface equations* and fully define (through the phase fraction λ and interface directions \mathbf{n} and \mathbf{m}) the microstructure seen at the transition temperature. Again, note that the fineness of the twinned microstructure in the martensite domain is not determined by this theory (in fact, the theory would imply an infinitely fine twin).

In short, the nature of microstructure in SMAs is a consequence of the geometric condition of allowing for a jump in the deformation gradient while maintaining the requirement that the body remains unbroken. The geometric condition is schematically shown in Fig. 1.22. Furthermore, it can be shown that such a condition can be related to an energy minimizing sequence of deformation. This then provides an argument for why one might expect to actually observe such geometric configurations in nature.

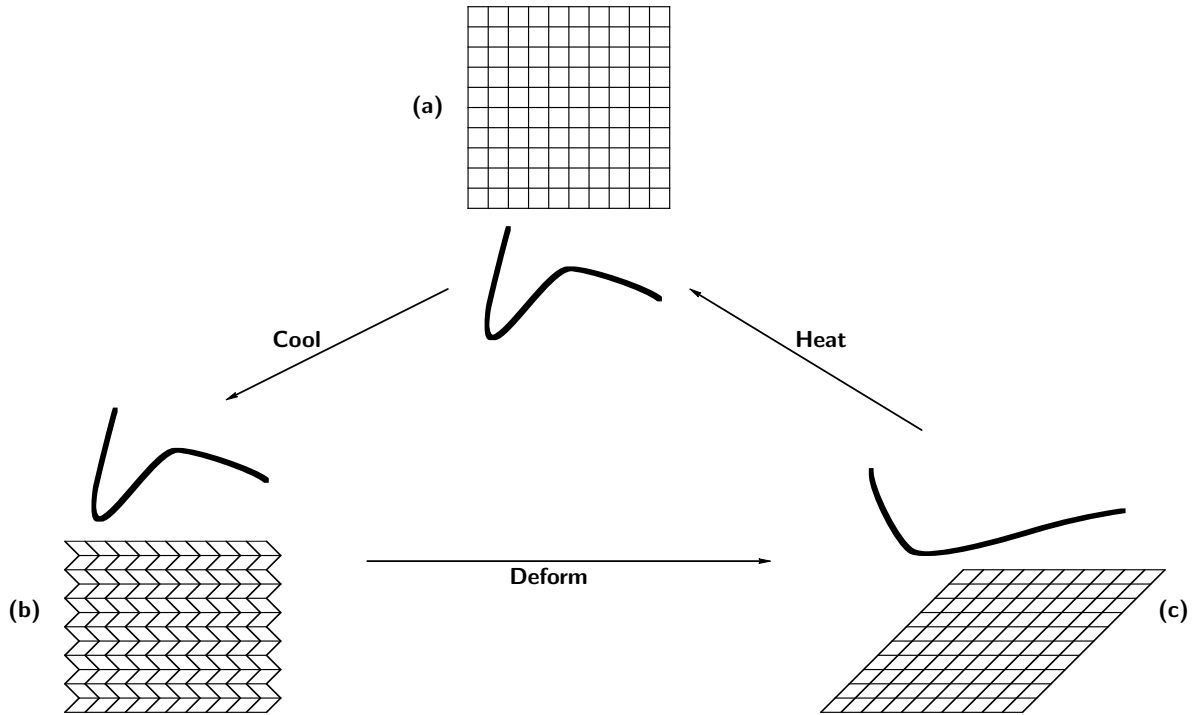


Figure 1.23: A schematic description of the shape memory effect. A SMA wire retains its shape upon cooling ((a) to (b)). The wire can recover its shape ((c) to (a)) upon heating past a critical temperature after it has been deformed while cool ((b) to (c)). Possible microstructure of the SMA is shown alongside the wire at each stage. The austenite phase is shown in (a) which transforms to a self-accommodated microstructure upon cooling (b). One martensite variant is preferred upon loading (c) to obtain a particular microstructure. Figure obtained from *Bhattacharya* (2003).

1.1.3 Macro-scale description

In the previous sections we saw that the characteristic multi-welled feature of an SMA energy landscape naturally results in deformations involving microstructures. In the macro-scale the mechanism behind the shape memory effect is the ability for the material to form microstructures that accommodate a given deformation. Consider the example in Fig. 1.23. Starting with the (austenite) reference configuration of the wire (shown in Fig. 1.23 (a)) at some temperature $\theta > \theta_c$, the wire is cooled past the transition temperature but retains its macroscopic shape via self-accommodation from the austenite to twinned martensite microstructure as shown in Fig 1.23 (b). The wire is deformed into the shape shown in Fig. 1.23 (c). To accommodate this deformation, the material can either: (1) distort its crystal lattice or (2) rearrange the variants to form a different microstructure. Recall that the microstructure is comprised of martensite variants that have the same energy. It is therefore favorable for the material to accommodate the deformation by changing its mi-

crostructure as opposed to distorting its crystal lattice. In other words, the material can change its shape to align with a given deformation without changing its energy. Furthermore, since both configurations have the same energy, there is no reason for the wire to return to its original configuration and the wire is now seemingly permanently deformed. As the wire is heated the martensite variants undergo a MPT into the austenite phase and the wire returns to its reference configuration. This whole process is that associated with the shape memory effect.

The process of pseudo-elasticity is related to the ability of the material to accommodate boundary conditions by undergoing a MPT. At high temperatures the material is in its reference austenite phase. Upon application of a load the material undergoes a stress-induced MPT from austenite to detwinned martensite characterized by the large deformation of the material with no appreciable increase in load. Removal of load results in the material transitioning back into the austenite phase at which point the reference configuration is fully recovered.

With this explanation, it is clear that the shape memory effect and pseudo-elasticity is heavily tied to the MPTs and the formation of microstructure, which is facilitated by the existence of non-unique low energy states (martensite variants) at low temperatures and a unique low energy state (austenite) at high temperatures. Note that MPTs are accompanied by latent heat of transformation resulting in non-uniform temperature fields and heat transfer (in both directions) that among other factors results in hysteresis.

1.2 Modulated martensites

In the preceding section we have presented a theory that has been used to great success to explain the existence and features of microstructures in SMAs. In the discussion of the theory we have argued that the shape memory effect and pseudo-elasticity phenomena in SMAs exist mainly because the material is able to form microstructures in order to accommodate a thermo-mechanical load. This ability to form microstructure is explained by the theory as the formation of energy-minimizing mixtures of phases that are each energy minima along the energy landscape of the material. In this section a special type of martensitic phase transformation relevant to this work will be discussed. As will be shown, this special type of martensitic phase transformation gives rise to a complication in the application of the above theory for a subset of SMAs—those that exhibit *modulated martensites* (MMs). In describing MPTs a term known as *lattice correspondence* is used to relate the unit cells in the initial and final structures. In essence, a lattice correspondence is used to relate the minimum number of essential unit cells of the high-symmetry structure to the essential unit

cell of the low-symmetry structure (*Bowles and Wayman (1972)*). Once the corresponding unit cells have been identified the lattice correspondence then describes the uniform deformation that transforms the essential skeletal unit cell of the martensite into the (possibly) non-essential unit cell of the austenite. If the transformation is between structures which both have the same number of atoms in the respective essential unit cells, then the correspondence will describe the uniform deformation that transforms one essential cell (martensite) to another essential cell (austenite). However, if the essential unit cells of the two structures contain different numbers of atoms, then the correspondence will necessarily relate non-essential unit cells. What is implied here is that it is not necessary that an essential unit cell of the martensite structure have a one-to-one correspondence to the essential unit cell of the austenite structure. For a material with a correspondence that uses non-essential unit cells the MPT is said to be period-extending. If it takes twice the number of atoms in the essential unit cell of the austenite structure to describe the essential unit cell of the martensite structure, then we say that the MPT is a period-doubling MPT. Most common MPTs are period-doubling (e.g., NiTi, AuCd). Period-preserving phase transformations are rare in SMAs and are most often found in materials that have interesting electromagnetic properties (e.g., MPT is BaTiO₃). One particular case of MPTs is a type of period-extending MPT that takes the austenite structure into a structure known as modulated martensite.

Since the middle of the twentieth century an increasing number of examples have been described in which X-ray scattering patterns exhibit features other than the regular arrangement of main Bragg reflections. Among these features are groups of reflections that are usually weak and arranged adjacent to some of the main reflections (see for example, Fig. 1.24). These reflections, which subdivide the unit cell in reciprocal space, are called *satellites* and can be explained by a periodic distortion of the basic structure. The basic structure which gives rise to the main reflections is thus modified by a modulating function that has a period of several unit cells of the basic structure. The basic structure is called the *ideal/base* structure and the periodically distorted structure is referred to as the *modulated* structure. For example, *Tanisaki (1961)* observed in NaNO₂ adjacent reflections amongst the main Bragg reflections which were later interpreted by *Tanisaki (1963)* and *Yamada et al. (1963)* as the basic crystal structure with a periodic distortion of the atomic positions. Modulated structures are commonly found in β -phase alloys (*Warlimont and Delaey (1974)*) where the austenite can be described as body-centered cubic (providing that atomic species are ignored). The modulated martensites are then composed of a periodic shuffling of (110)_{cubic} planes in the $[\bar{1}10]_{\text{cubic}}$ direction. Fig. 1.24 gives an example of a diffraction

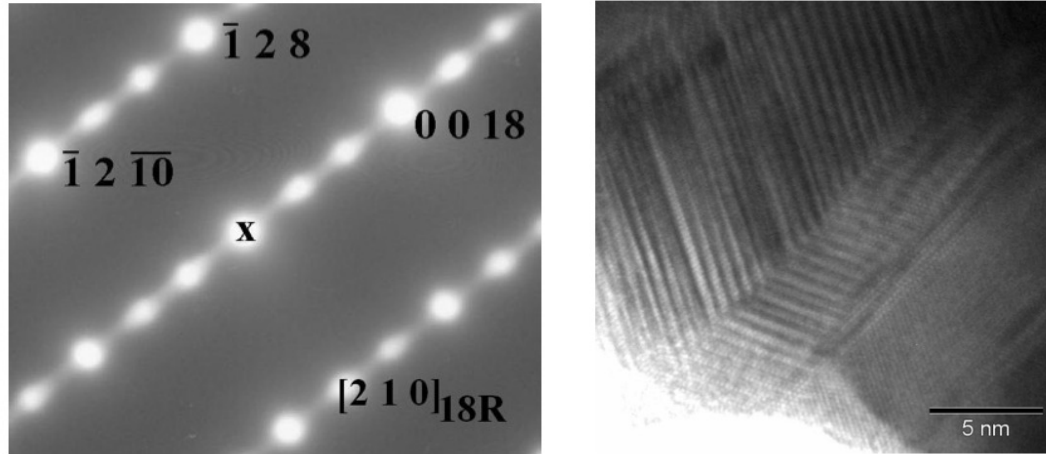


Figure 1.24: An example of modulated martensite found in Cu-Al-Ni alloys (*Torres et al. (2007)*) showing diffraction pattern corresponding to $[210]_{18R}$ zone axis (left figure) and twins of the 18R martensite (right figure).

pattern for a modulated phase found in Cu-Al-Ni alloys obtained by *Torres et al. (2007)*. It is useful to view the phase transition of β -phase alloys into modulated martensites from the dispersion relation perspective. This is because one can view the cubic to modulated martensite transformations as the onset of dynamical instabilities when subjected to perturbations given by wavevectors in the $(110)_{\text{cubic}}$ direction⁹. The modulated martensite is then described by a combination of the reference BCC state and the normal mode associated with the unstable phonon mode (which in this case will be shuffling of $(110)_{\text{cubic}}$ planes). This observation will be used later in this work.

Several other examples of materials with modulated martensites have been experimentally observed. For example, one of the most common materials that have been observed to have modulated phases are Ni-Mn-Ga alloys where a BCC austenite is indeed observed if one were to disregard the atomic species. Additionally, copper based SMAs studied by *Balandraud and Zanzotto (2007)* have been found to have modulated martensites. Other types of materials such as Fe-Pd (*Khachaturyan et al. (1991)*) and NaNO_2 (*Tanisaki (1961)*) have been observed to have modulated phases. Further, *Bohm (1983)* observed in LiAlSiO_4 that modulated phases most commonly nucleate at phase transitions and that the actual modulation that is observed depends on the composition of the alloys. *Pons et al. (2000)* observed modulated martensites in near stoichiometric Ni-Mn-Ga alloys that correspond to 10, 14, and 18 $(110)_{\text{cubic}}$ planes, depending on the composition of the alloys. The following fig-

⁹The wavelength of the perturbations are commensurate with the periodicities of the observed modulated martensite.

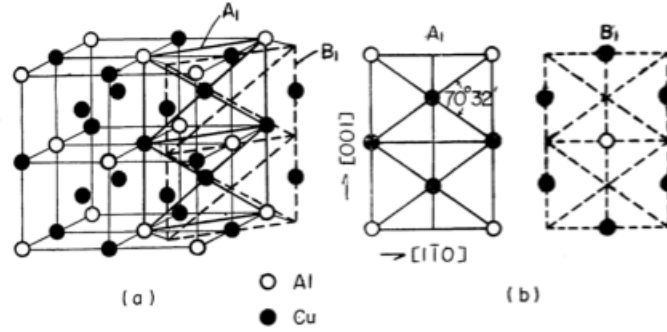


Figure 1.25: A figure of the austenite phase in Cu-Al alloys *Nishiyama and Kajiwara* (1963). The austenite phase can be considered to be an alternating stacking arrangement of 'A₁' and 'B₁' planes.

ures from *Nishiyama and Kajiwara* (1963) provide an example of a MPT in Cu-Al alloys where the austenite phase is B2 cubic and can be considered to be an alternating stacking arrangement of 'A₁' and 'B₁' planes (Fig. 1.25). The modulated martensite phase is formed by a periodic shuffling of the (110)_{cubic} planes (Fig. 1.26). Additionally, there are materials that have been observed to not only have modulated martensites of a single modulation but instead exhibit multiple observable modulations. For example, in the diffraction study of Ni-Fe-Ga alloys performed by *Hamilton et al.* (2007) the material was thermally induced to undergo a MPT from a BCC austenite phase to a 10M (modulation of five unit cells) structure. However, *Hamilton et al.* (2007) also observed a 14M (modulation of seven unit cells) martensite during thermal cycling under constant tensile stress. In both the 10M and 14M martensites the base structure was found to be a face-centered-tetragonal (FCT) structure (see Fig. 1.27). Another example of a material that exhibit several modulated martensites can be found in *Bak and von Boehm* (1980) where the rare-earth metal CeSb (not a SMA) was observed to undergo several phase transitions between various commensurate phases. Starting from a FCC structure at temperatures greater than 16.1 K, the system passes through phases with periodicities of 13, 7, 18, and 11 as the temperature is lowered. At the lowest temperature in the experiment CeSb is found to have a martensite phase with period equal to 4 times the unit cell of the FCC austenite structure.

Modulated martensites are important because they occur in materials with important technological applications. Many of these materials are ferromagnetic, ferroelectric, or multiferroic and are candidates for heat engines, microactuators, and microsensors. Additionally, modulated martensites may be part of a mechanism for improving kinematic compatibility since many materials that are highly compatible have been shown to exhibit modulated phases. For example, Ti-Ni-X (with $X \in \{Cu, Pd, Pt, Au\}$) alloys have been shown to have

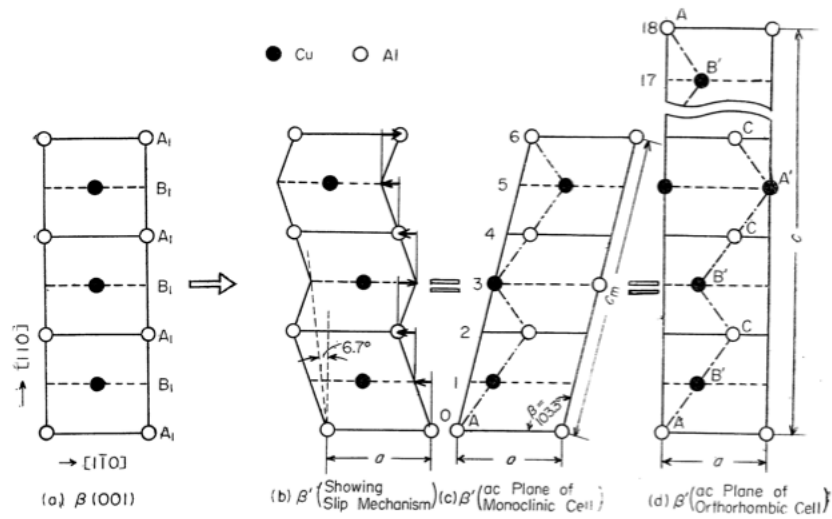


Figure 1.26: An example of modulated martensites in Cu-Al alloys *Nishiyama and Kajiwara* (1963). (a) shows the austenite phase, (b) shows the glide occurring at the first two layers in one direction and then a third layer in the opposite direction. (c) and (d) represent other mechanisms to form other modulated martensites.

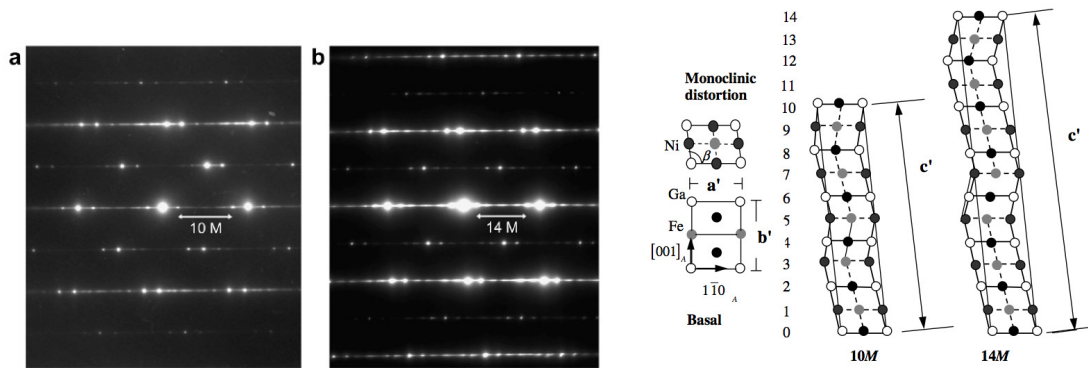


Figure 1.27: An example of a structure exhibiting two modulations of modulated martensites (*Hamilton et al. (2007)*). The left side of the figure displays a diffraction pattern for the 10M structure and the 14M structure marked (a) and (b), respectively. The right side of the figure displays a schematic of the modulated martensites composed of five (10M) and seven (14M) unit cells of the FCT base martensite. The 10M martensite was thermally induced without applied stress, while the 14M martensite was stress-induced during thermal cycling under constant tensile loads.

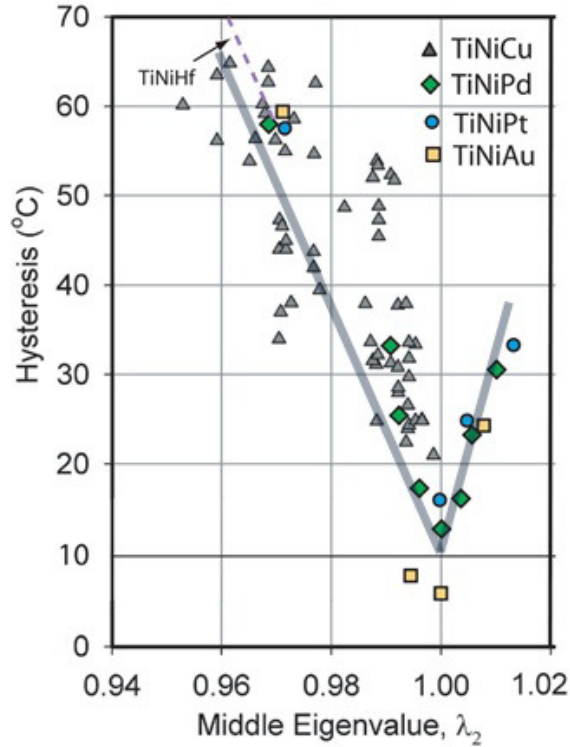


Figure 1.28: A plot of hysteresis vs λ_2 for various materials obtained from *Zhang et al.* (2009).

modulated phases (*Potapov et al.* (2003)) and are highly kinematically compatible as shown in Fig. 1.28 obtained from *Zhang et al.* (2009). Kinematic compatibility is desirable since it has been shown to lead to low hysteresis, high reversibility, and a high fatigue lifespan.

A complication arises when describing MPTs in materials that exhibit modulated martensites. Recall that the micro-scale theory used to explain the existence of microstructures in SMAs depends on *a priori knowledge of the energy minimizing deformations*. That is, in each of the special microstructures discussed above the solution to the variational problem is obtained by minimizing the energy from a set of allowable deformations that are each assumed to be energy minima (i.e., the allowable deformations being the deformations that describe the austenite and martensite variants). Further, this allowable set of deformations is assumed to be fixed and known a priori as part of the material's crystallographic description. However, in view of the existence of modulated structures, where the modulated martensites are comprised of a (periodically altered) base martensite, it is unclear as to which structures should be included in the set of allowable deformations. The above examples of Ni-Fe-Ga and CeSb alloys highlight materials that not only possess one stable modulated martensite but are in fact observed to have stable martensites of several different

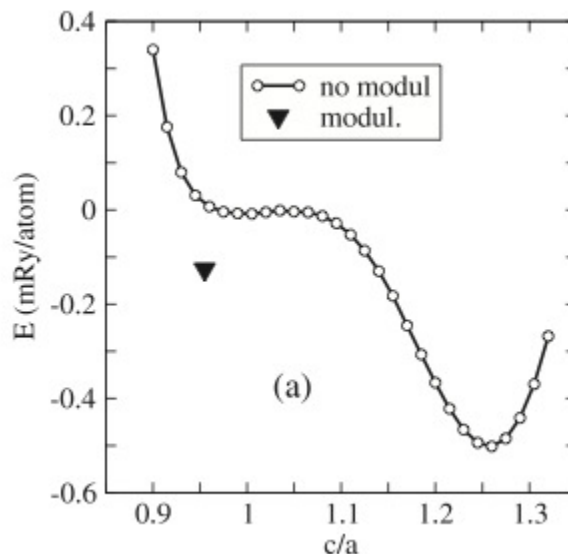


Figure 1.29: Despite common belief, modulated martensites are not global minima in energy but are instead meta-stable states of higher energy (Zayak *et al.* (2003)).

wavelengths. However, it is still unclear as to whether stable martensite phases of *all* modulations can exist in the same material or what factors prohibit some modulations from being observed in experiments. It is also unclear whether the ability of certain materials to undergo a MPT into martensites of multiple modulations is inherently special to those materials or if that ability is a general characteristic of all materials that undergo a MPT.

Another complication in the application of the micro-scale theory of microstructure formation arises in materials that exhibit modulated martensite. This is because commonly modulated martensites are considered as thermodynamically stable structures that correspond to ground states of the energy landscape (at a particular thermo-mechanical loading). The basis for this assumption is due to the experimental observations of modulated martensites in the material. However, DFT calculations of a Ni_2MnGa alloys by Zayak *et al.* (2003) showed that modulated martensites are meta-stable states of higher energy and that in fact, it is the base martensite that is the global minima in energy¹⁰. Fig. 1.29 is a plot obtained from Zayak *et al.* (2003) showing the energy as a function of c/a ratio. The modulated phase has a c/a ratio of around 0.975 and is computed to be a higher energy than the no modulation base martensite that has a c/a ratio of around 1.25. Assuming this result to be true, then one question that can be asked is what is the mechanism that stabilizes the formation of modulated martensites in certain SMAs? Additionally, how does one describe the formation of microstructures involving modulated martensites? Presently, there

¹⁰Note that Zayak *et al.* (2003) never evaluated the dynamic stability of the modulated structures.

are two competing theories to explain the experimental observation of meta-stable modulated martensites: (1) the adaptive martensite hypothesis and (2) modulated martensites as periodic stacking of close-packed planes.

1.2.1 Adaptive martensite viewpoint of modulated martensites

One idea that has come up in the literature is that of *adaptive martensites* (see for example, *Khachaturyan et al.* (1991) and *Kaufmann et al.* (2011)) which views modulated martensites as *nano-twinned tetragonal martensite*. The adaptive martensite hypothesis (AMH) requires that the ground state martensite is of tetragonal symmetry since non-tetragonal martensites do not twin on $(110)_{\text{cubic}}$ planes (recall that in all cases of interest here modulated martensites are characterized as periodic shuffling of $(110)_{\text{cubic}}$ planes). The inspiration behind the adaptive martensite perspective for the formation of modulated phases is based on the consideration of a homogenous nucleation of a martensitic domain at a MPT and a close examination of the austenite-martensite interface that occurs at such a nucleation¹¹. Moreover, researchers such as *Khachaturyan et al.* (1991) realized that at the atomic level modulated martensites could be constructed when lamellae of the twin-related variants reach a size (λ in Fig. 1.21) comparable with the interatomic distance. At this scale the microscale structure of the martensitic plate becomes a nanoscale structure where the martensite plate is composed of a nano-inhomogeneous lattice. This nano-inhomogeneous lattice can be described as an ideal homogeneous lattice whose atomic structure is related to the atomic structure of the base martensite lattice by an appropriate shuffling of crystal planes plus a certain distortion of these planes and their interplanar distances. The lattice shuffling and distortion imitates the homogenous nucleation of a nano-twinned ground-state tetragonal martensite within the austenite domain (*Khachaturyan et al.* (1991)). This leads to the idea that the modulated martensites can be viewed as a meta-stable structure composed of thermodynamically stable base martensite that is stabilized by the requirement of a kinematically compatible interface at the habit plane. This kinematic compatibility condition is very much analogous to the rank-one condition discussed above in the micro-scale description of a MPT. In this manner, we can view the mechanism that gives rise to a modulated structure as the nano-scale analog to the micro-scale rank-one condition. That is, the mechanism is described as an atomic-scale twinning—nano-twinning. The hypothesis behind the adaptive martensite perspective is then guided by the application of the rank-one compatibility between the base martensite and austenite at the discrete level to explain the mechanism behind the observation of modulated martensites in crystallographic

¹¹This region has so far been treated as a “fuzzy” continuum interface (see for example, a typical depiction of this interface as the gray region in Fig. 1.21) where the interface is not considered at an atomic level.

experiments.

To illustrate the adaptive martensite hypothesis (AMH) further consider a MPT with a large crystal-lattice mismatch. In such a MPT *Khachaturyan et al.* (1991) argues that the ideal martensite phases cannot homogeneously nucleate because the elastic energy generated by a coherent homogenous nucleus of the martensite phase is too high. This elastic energy barrier completely blocks the nucleation of homogenous particles of the ideal martensitic phase via thermal fluctuations. The adaptive martensite viewpoint then hypothesizes that given a material with low twin boundary energy and low twin boundary interaction energy, the material can transform by bypassing such a nucleation and instead of a homogenous domain of the base martensite, a domain of low lattice mismatch modulated martensite may nucleate. This modulated martensite is henceforth called the adaptive martensite because its structure is realized by the kinematic compatibility to the austenite and not solely by its position on the energy landscape. The adaptive martensite may then be viewed as a metastable phase whose free energy is slightly higher than that of the stable base martensite but whose nucleation, unlike the nucleation of the base martensite, does not substantially raise the elastic energy. Thus, the nucleation of the adaptive martensite can occur in the same way as the classical nucleation of a phase with a low crystal-lattice mismatch.

There have been several works that have used the adaptive martensite theory to describe modulations that are consistent with those observed. For example, *Lovey* (1987) is able to explain the transformation from a BCC to an 18R structure (a modulation of 18 (110)_{cubic} planes) found in a crystallographic analysis of Cu-Al and Cu-Zn-Al alloys. In this particular work, *Lovey* (1987) calculated the “stacking fault density”, which is a measure of the deviation away from a perfect FCT structure, for a 9R (or 18R) modulated structure using the rank-one kinematic compatibility condition. *Lovey* (1987) found that this theoretically computed stacking fault density was in good agreement to stacking fault densities found using experimental crystallographic analysis. Another similar example can be found in *Krystian and Pichl* (2000) where a transformation of a BCC structure to 9R modulated structure was found in crystallographic analysis of Lithium. In their work, *Krystian and Pichl* (2000) explicitly computed the deformation gradient \mathbf{U} as a composition of two deformations that first take the BCC structure to an FCC/HCP structure and then take the FCC/HCP structure to the 9R structure observed in experiments. Having obtained the deformation gradient, *Krystian and Pichl* (2000) then verified that in fact this experimentally computed deformation gradient does indeed form a rank-one connection to the austenite (BCC) phase.

Despite the successes of the AMH there are shortcomings in the adaptive martensite viewpoint. The AMH does not give any insight into the stability of different modulations of modulated martensites. That is, the adaptive martensite hypothesis cannot provide information regarding the dynamic stability (local minima) or the thermodynamic stability (global minima) of modulated martensites. Perhaps more importantly, the AMH cannot explain modulated martensites with non-tetragonal base martensite. This is because for anything but a tetragonal base martensite the twinning interface between the variants does not occur along the modulating direction (i.e., along $(110)_{\text{cubic}}$ planes). For example, an orthorhombic base martensite will twin along the $(100)_{\text{cubic}}$ planes. Therefore, an alternative explanation must be used in order to explain the existence of modulated martensites that have a non-tetragonal base martensite. This leads us to the second viewpoint on the experimental observation of modulated martensites discussed in the next subsection.

1.2.2 Modulated martensites as periodic stacking of close-packed planes

Another idea for the experimental observation of modulated martensites is to interpret the phase transformation taking the BCC austenite (ignoring atomic species) to a martensite composed of a regular stacking of close-packed planes. The compaction of the BCC lattice then occurs through a two-step process where the alternate $(110)_{\text{cubic}}$ planes of the BCC crystal first undergo an in-plane contraction and then experience a suitable relative sliding of adjacent planes and an adjustment of the interplanar distance to attain a three-dimensional close-packed structure. The final stacking sequence produced by this “zig-zag” sliding defines the modulated martensite. In this manner, *Balandraud and Zanzotto (2007)* identify a ‘fundamental planar unit cell’ as shown in Fig. 1.30 (obtained from *Balandraud and Zanzotto (2007)*). According to *Balandraud and Zanzotto (2007)*, the “zig-zag” stacking sequence is determined by the compatibility conditions at the austenite-martensite interface. This is similar to the adaptive martensite hypothesis where the pattern of the long-period stacking of modulated martensites is interpreted as nano-twinning. To demonstrate the idea, *Balandraud and Zanzotto (2007)* first derived the micro-scale deformation gradient that is at the root of all bcc-to-close-packed transformations from the two-step mechanism described above. Next, several modulated structures were analyzed to determine their stacking sequence and applied to obtain the deformation gradient associated to the modulated martensites studied. *Balandraud and Zanzotto (2007)* found that in fact, it is, “close enough” to compatibility suggesting that a habit plane can be formed comprising of austenite and variants of modulated martensite.

One important aspect that distinguishes this work from the AMH is that by introducing the

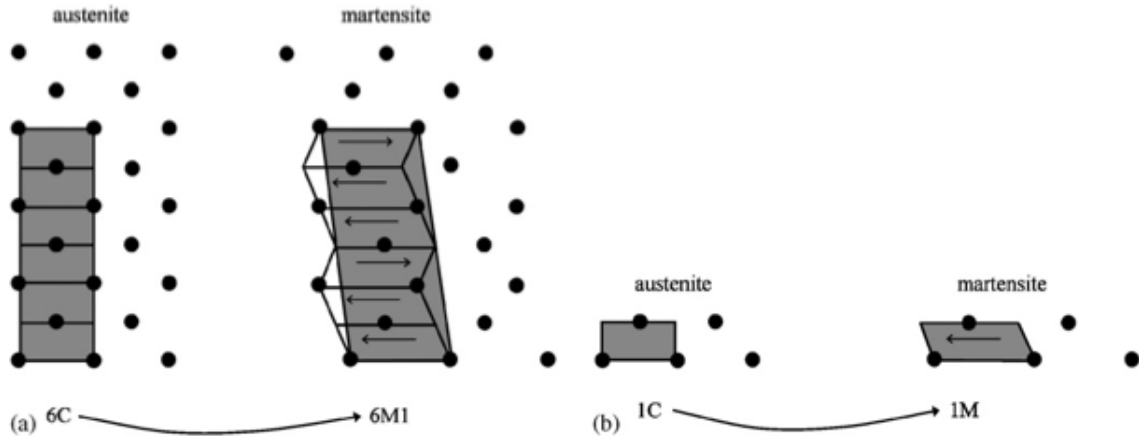


Figure 1.30: (a) A schematic of the transformation from BCC to a modulated structure 6M. (b) The fundamental unit that composes a modulated martensite is composed of two adjacent layer of uniformly compacted $(110)_{\text{cubic}}$ BCC planes *Balandraud and Zanzotto* (2007).

fundamental planar unit cell as described above *Balandraud and Zanzotto* (2007) remove the limitation of requiring the presence of a tetragonal ground-state martensite. In fact in some sense, *Balandraud and Zanzotto* (2007) have replaced the role of the ground-state martensite with the fundamental planar unit-cell in the calculation of kinematic compatibility. However, this viewpoint does share two shortcomings with the AMH, in that (1) it does not have the ability to identify which modulated structures are local minima and (2) it does not provide any information regarding the global energy ordering of the modulated structures.

1.3 Probing the energy landscape of a typical SMA

Note that while the AMH and the viewpoint described in *Balandraud and Zanzotto* (2007) are both useful in explaining the experimental observation of modulated martensites, they do not give any insight into the energy characteristics of the different structures examined. Therefore, a careful examination of the features of the atomistic energy landscape of a typical SMA may yield further understanding that could be useful in interpreting what modulated martensites are, why they form, and what role they play in SME and PE.

The main objective for this work is to probe the energy characteristics of a typical SMA using empirical atomistic simulation. It is our belief that modulated martensites are natural features of the energy landscape of SMAs and further, that modulated martensites of arbitrary modulations can be found as equilibrium points on the energy landscape of the material. In short, this work seeks to verify the hypothesis that *a material does not have one*

unique martensite structure but instead has multiple stable martensites with modulations of various wavelengths. The modulation that ultimately is observed will depend on a variety of factors such as boundary conditions, geometrical constraints, and defects, and how these factors affect the energy landscape of the material. Several works can be found that may lend credence to this hypothesis. For example, *Kaufmann et al. (2011)* have experimentally observed a material where multiple modulations were found to occur within the same sample. They explain the relationship between the modulations observed as a branching of twinned structures within the domain. *Hamilton et al. (2007)* observed that Ni-Fe-Ga alloys exhibit several modulations of martensite depending upon whether the material is mechanically loaded during thermal cycling. Specifically, *Hamilton et al. (2007)* observed a modulated 10M structure upon a stress-free thermal cycling of the material and a 14M structure upon thermal cycling under a constant tensile load suggesting that the observed modulation of modulated martensites depends on the applied load on the material. This suspicion was also noted by *Khachaturyan et al. (1991)* who stated that, “a typical feature of the adaptive martensite is sensitivity of its crystal structure to the applied stress”. It is therefore conceivable that there is a coupling between the stress fields associated with the nucleation of martensitic domains and the observed modulation during a martensitic phase transformation. In this manner, certain regions within the austenite domain can provide a stress field that will bias the nucleation of different modulations of martensite.

1.4 Outline

In order to probe the energy landscape of the SMA model used in this work a technique called the *Branch-Following and Bifurcation (BFB) method* will be heavily used. As such, Ch. 2 of this document will briefly detail a generic branch-following and bifurcation method for systems with a given potential energy function. Here, the computation of *equilibrium paths* corresponding to crystal structures of a given symmetry group will be detailed. The detection and computation of critical points known as *symmetry-breaking bifurcation points* will also be discussed. These bifurcation points will be of special interest since multiple paths of different symmetry may cross at these bifurcation points. These points will then correspond to material instabilities where a particular phase will transition into another. Next, a technique known as the *Lyapunov-Schmidt-Koiter (LSK) decomposition* will be discussed. This technique will be useful in order to compute paths bifurcating away from bifurcation points. These paths are important since they may correspond to modulated structures.

Chapter 3 will detail the particular model that will be used in this work. Once the details of

the model have been presented the result of the BFB study of that model will be discussed. We will see that several structures corresponding to modulated phases of commonly observed modulations are natural features of the energy landscape of the model. Details regarding the modulated phases such as kinematic compatibility, energy differences, and the effect of a kinematic compatibility constraint will be presented in this chapter. In Ch. 4 a concept that we have called the *Modulated Martensite Mixture Model* (M^4) will be presented. This new interpretation of modulated martensites is developed in view of the analysis of the various structures found in the BFB study discussed in Ch. 3. Using only a small set of input data, M^4 is capable of accurately predicting the energy, lattice constants, and structural details of an arbitrary modulated martensite. This is demonstrated by comparing the M^4 predictions to atomistic computational results. Finally, the predictive capabilities of the M^4 will also be used in order to predict the existence of a structure that is highly kinematically compatible to B2. This prediction is then verified via an atomistic simulation of the predicted structure. This document concludes in Ch. 5 which provides a summary and conclusions drawn from this work.

Chapter 2

Branch-following and bifurcation method

The purpose of this chapter is to briefly detail the mathematics behind the main tool that will be used in this work. The tool of importance is known here as the Branch-Following and Bifurcation (BFB) method and is a combination of several techniques that together can be used to obtain equilibrium solutions to systems that are described by potential functions (*Allgower (1980)*). Examples of where the BFB method may be useful can be found in several applications such as chemistry, mathematics, structural mechanics, modern physics, biology etc (for some examples, see *Jusuf (2010)* or *Allgower (1980)*). Relevant to this work, the BFB method is ideally suited in the study of materials since in many cases materials can be described by defining an energy function that depends on the degrees of freedom in the problem (for example, lattice parameter) and the given boundary condition (for example, tractions). In these types of problems, this method can be used to compute the evolution of the degrees of freedom as the boundary condition is varied.

The first part of this chapter will introduce the concept of branch-following. This technique can be used to obtain general *equilibrium* solutions for systems defined by potential functions. The next part of the chapter deals with the detection and computation of *critical points*. A subset of these critical points are called *symmetry-breaking bifurcation points* and are of special interest since these are points where two (or more) paths of distinct symmetries cross. Next, a technique known as the Lyapunov-Schmidt-Koiter (LSK) decomposition will be presented. This technique will be used in this work to aid in the determination and computation of paths branching away from bifurcation points. Note that this chapter does not seek to present in detail the various topics mentioned in the preceding paragraphs.

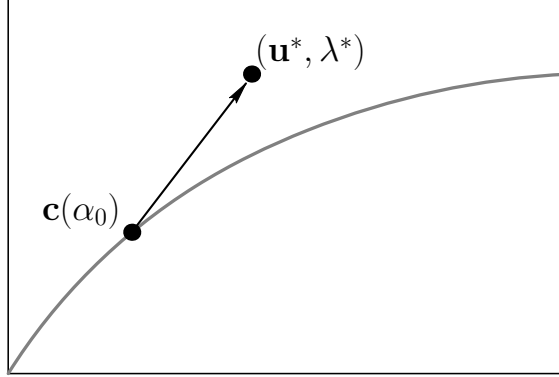


Figure 2.1: Starting from $\mathbf{c}(\alpha_0)$ a predictor step is taken onto $(\mathbf{u}^*, \lambda^*)$.

For a more detailed presentation on these topics refer to *Jusuf (2010)*.

2.1 Branch-following

A branch-following or continuation method deals with curves that can be parameterized with respect to a general parameter λ . These curves are generally described as solutions to nonlinear systems of the form

$$\mathbf{F}(\mathbf{u}, \lambda) = \mathbf{0}, \quad (2.1)$$

where $\mathbf{F}(\mathbf{u}, \lambda) : \mathbb{R}^N \times \mathbb{R} \rightarrow \mathbb{R}^N$. Eq. (2.1) is commonly referred to as an *equilibrium equation*. In this work, \mathbf{F} will be derived from a potential function \mathcal{E} and Eq. (2.1) can be written as

$$\mathbf{F}(\mathbf{u}, \lambda) \equiv \nabla_{\mathbf{u}} \mathcal{E}(\mathbf{u}, \lambda) = \mathbf{0}. \quad (2.2)$$

It can be shown (*Allgower (1980)*) that Eq. (2.2) defines a continuous curve of solutions, commonly known as an *equilibrium path* and is here denoted by $\mathbf{c}(\alpha) \equiv (\mathbf{u}(\alpha), \lambda(\alpha))$, where $\alpha \in \mathbb{R}$ can be viewed as an arclength-type parameter. The main idea behind the branch-following method is to obtain a discrete approximation to $\mathbf{c}(\alpha)$ by numerically computing a finite subset of points $\{\mathbf{c}(\alpha_i)\} \subset \mathbf{c}(\alpha)$, $i = 1, 2, \dots, b < \infty$. In this work, this is done using a *Predictor-corrector* algorithm. Predictor-corrector methods take in as an input an initial *known* solution point $\mathbf{c}(\alpha_0) \in \mathbf{c}(\alpha)$ and obtain a predicted point (shown as $(\mathbf{u}^*, \lambda^*)$ in Fig. 2.1). From the predicted point the corrector process iteratively updates the predicted point until the next point on the curve is obtained (shown as $\mathbf{c}(\alpha_1)$ in Fig. 2.2). The process is repeated using the new point on the curve as a new input to the predictor-

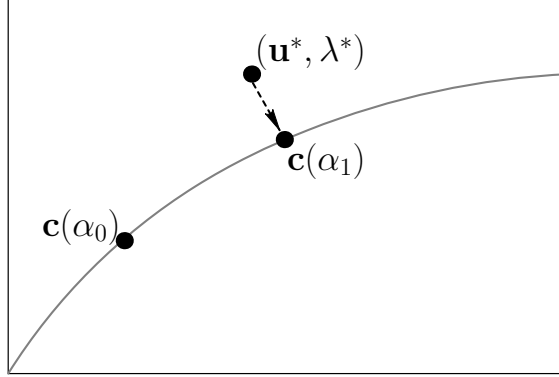


Figure 2.2: Starting from $(\mathbf{u}^*, \lambda^*)$ the corrector process iteratively solves for $\mathbf{c}(\alpha_1)$.

corrector process. In this way a sequence of points along $\mathbf{c}(\alpha)$ can be obtained. Several sophisticated schemes exist that can be used in the predictor-corrector process depending on the desired convergence rate and resolution for $\mathbf{c}(\alpha)$. For more details on the predictor-corrector method as well as the algorithm for a number of specific predictor-corrector methods, refer to *Jusuf* (2010).

2.1.1 Critical points

Critical points of \mathbf{F} are points $(\mathbf{u}_c, \lambda_c) \in \mathbb{R}^N \times \mathbb{R}$ that satisfy the following conditions:

$$\mathbf{F}(\mathbf{u}_c, \lambda_c) = \mathbf{0}. \quad (2.3)$$

$$\nabla_{\mathbf{u}}\mathbf{F}(\mathbf{u}_c, \lambda_c) \text{ is singular. That is, } \text{rank}(\nabla_{\mathbf{u}}\mathbf{F}(\mathbf{u}_c, \lambda_c)) < N. \quad (2.4)$$

Note that the above conditions imply that critical points are equilibrium points of \mathbf{F} such that $\nabla_{\mathbf{u}}\mathbf{F}$ contains at least one zero eigenvalue. Therefore, by keeping track of the eigenvalues of $\nabla_{\mathbf{u}}\mathbf{F}$ in the computation of $\{\mathbf{c}(\alpha_i)\}$ the presence of critical points between two successive equilibrium points $\mathbf{c}(\alpha_i)$ and $\mathbf{c}(\alpha_{i+1})$ can be detected. Again, there are subtleties in the implementation of this technique that will not be presented in this work.

Once the presence of a critical point has been detected, it may be desirable that the critical point be computed to a high degree of accuracy. There are a variety of methods, such as the *bisection* method, that can be used to compute a critical point once it has been detected. These methods are based on root finding algorithms for functions of a single argument. Refer to *Jusuf* (2010) for more on the detection and computation of critical points.

Critical points are classified in this work into the three categories: (1) turning points, (2) symmetry-preserving bifurcation points, and (3) symmetry-breaking bifurcation points. In order to discuss the concept of symmetry-preserving or symmetry-breaking the concept of

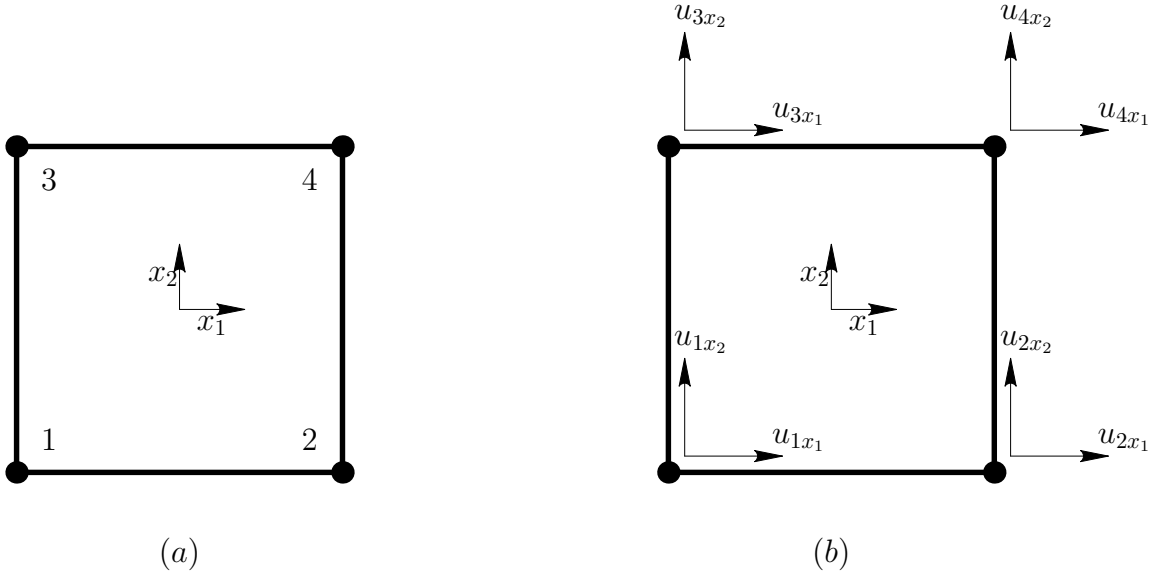


Figure 2.3: A square symmetric structure. (a) identifies nodes 1 to 4. (b) identifies vector degree of freedom at each node. \mathbf{u}_i is the displacement vector of node i .

the symmetry of a path will be discussed next.

2.1.2 Symmetry of an equilibrium path

Consider a system with symmetry described by the abstract group G . Let the DOFs of the system be described by a vector $\mathbf{u} \in \mathbb{R}^N$.

Definition. A *matrix representation* of a group G on \mathbb{R}^N is a homomorphism $\mathbf{T} : g \rightarrow \mathbf{T}(g)$ of G into $\mathbb{R}^{N \times N}$, where $g \in G$ and $\mathbb{R}^{N \times N}$ is the set of all non-singular linear transformations of \mathbb{R}^N into itself.

In general, the construction of a matrix representation of a group for a given problem depends on the DOFs of the problem (i.e., the physics of the problem). For example, suppose we are interested in a truss structure composed of identical bars as shown in the Fig. 2.3. The DOFs of the system are given by $\mathbf{u} = [u_{1x_1}, u_{1x_2}, u_{2x_1}, u_{2x_2}, u_{3x_1}, u_{3x_2}, u_{4x_1}, u_{4x_2}]^T \in \mathbb{R}^8$. The symmetry of the problem is given by $G = \{e, c_4, c_2, c_4^{-1}, \sigma_{13}, \sigma_{24}, \sigma_{x_2}, \sigma_{x_1}\}$ where e is the identity element, c_n is the clockwise rotation by $360/n$ degrees, and c_4^{-1} is the counter-clockwise rotation by 90 degrees. The operations σ_{13} and σ_{24} are mirrors about the diagonals, and σ_{x_1} and σ_{x_2} are mirrors about the horizontal and vertical axis. For a pictorial description of the operations of the symmetry group, refer to Fig. 2.4.

The action of the group on each nodal degree of freedom is a product of a permutation of

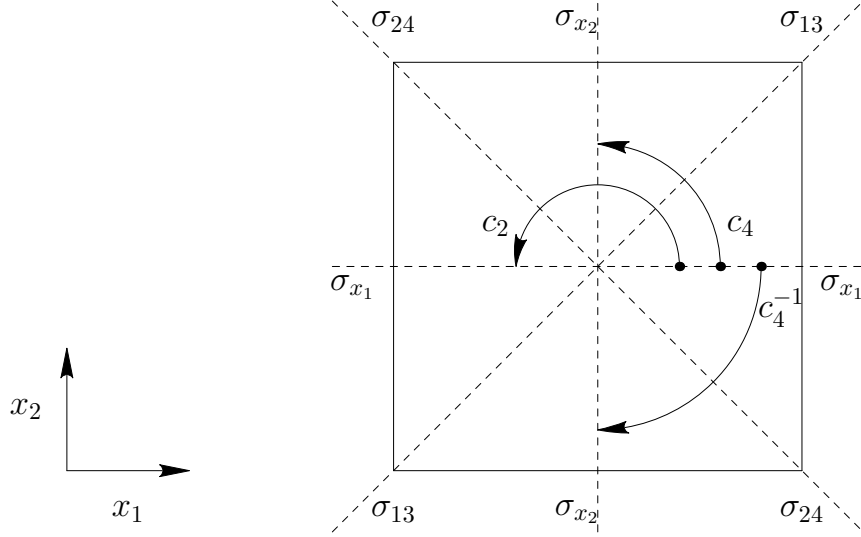


Figure 2.4: Symmetry of the square.

the nodes and the action of the group on the vector DOFs at each node. A representation of G in terms of permutation $\mathbf{T}_1(g) : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ is given by the following:

1. $\mathbf{T}_1(e) : [1, 2, 3, 4]^\top \mapsto [1, 2, 3, 4]^\top$.
2. $\mathbf{T}_1(c_4) : [1, 2, 3, 4]^\top \mapsto [3, 1, 4, 2]^\top$.
3. $\mathbf{T}_1(c_2) : [1, 2, 3, 4]^\top \mapsto [4, 3, 2, 1]^\top$.
4. $\mathbf{T}_1(c_4^{-1}) : [1, 2, 3, 4]^\top \mapsto [2, 4, 1, 3]^\top$.
5. $\mathbf{T}_1(\sigma_{13}) : [1, 2, 3, 4]^\top \mapsto [1, 3, 2, 4]^\top$.
6. $\mathbf{T}_1(\sigma_{24}) : [1, 2, 3, 4]^\top \mapsto [4, 2, 3, 1]^\top$.
7. $\mathbf{T}_1(\sigma_{x_1}) : [1, 2, 3, 4]^\top \mapsto [3, 4, 1, 2]^\top$.
8. $\mathbf{T}_1(\sigma_{x_2}) : [1, 2, 3, 4]^\top \mapsto [2, 1, 4, 3]^\top$.

For example, $\mathbf{T}_1(e) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$, and $\mathbf{T}_1(c_4) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$.

The components of a regular two-dimensional vector transform according to the vector transformation representation $\mathbf{T}_2(g) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$. For example,

$$\mathbf{T}_2(e) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } \mathbf{T}_2(c_4) = \begin{bmatrix} \cos(90^\circ) & -\sin(90^\circ) \\ \sin(90^\circ) & \cos(90^\circ) \end{bmatrix}.$$

The matrix representation $\mathbf{T}(g) : \mathbb{R}^8 \rightarrow \mathbb{R}^8$ in this particular case is then given by $\mathbf{T}(g) = \mathbf{T}_1(g) \otimes \mathbf{T}_2(g)$ for all $g \in G$. For example, in 2×2 block form,

$$\mathbf{T}(e) = \begin{bmatrix} \mathbf{T}_2(e) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_2(e) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{T}_2(e) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{T}_2(e) \end{bmatrix} \text{ and } \mathbf{T}(c_4) = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{T}_2(c_4) & \mathbf{0} \\ \mathbf{T}_2(c_4) & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{T}_2(c_4) \\ \mathbf{0} & \mathbf{T}_2(c_4) & \mathbf{0} & \mathbf{0} \end{bmatrix},$$

where $\mathbf{0}$ is the 2×2 matrix with zero entries. While the above example is specific to the square-symmetric problem it illustrates the process of the construction of a matrix representation. In general, knowledge of the symmetry of the problem and how it acts on the DOFs of the problem is sufficient to construct the matrix representation of a group.

Definition. An equilibrium path $\mathbf{c}(\alpha)$ is said to be a **G -symmetric equilibrium path** if

$$\mathbf{T}(g)\mathbf{u} = \mathbf{u}, \quad \forall g \in G, \quad \forall (\mathbf{u}, \lambda) \in \mathbf{c}(\alpha), \quad (2.5)$$

where $\mathbf{T}(g) : \mathbb{R}^N \rightarrow \mathbb{R}^N$ is the appropriate matrix representation of the element g belonging to the abstract finite symmetry group G on the carrier space \mathbb{R}^N .

Now that the concept of the symmetry of a path has been discussed, the three types of critical points will be presented next. For the remainder of this discussion it will be assumed that critical points detected and computed are found along a G -symmetric equilibrium path $\mathbf{c}(\alpha)$. This G -symmetric equilibrium path will be referred to as the *principal path*.

1. *Turning points*

Turning points are critical points with a single unique equilibrium path that passes through the critical point and correspond to points of local extrema with respect to the loading parameter λ of the solution path $\mathbf{c}(\alpha)$ (see Fig. 2.5). In physics and engineering applications a turning point can often signify a change in the stability of the solutions.

2. *Symmetry-preserving bifurcation points.*

Bifurcation points are critical points characterized by the presence of multiple solution paths intersecting at the critical point (see Fig. 2.6). Moreover, in the case of

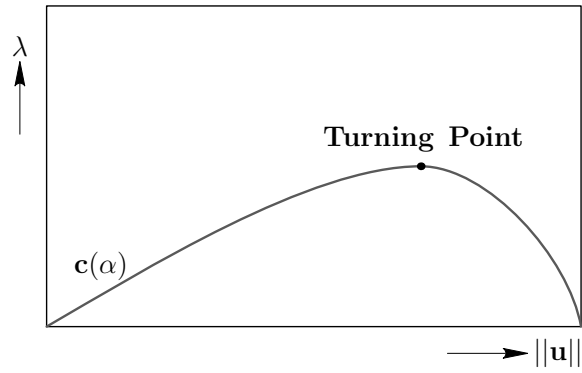


Figure 2.5: An example of a turning point found along the equilibrium path $c(\alpha)$.

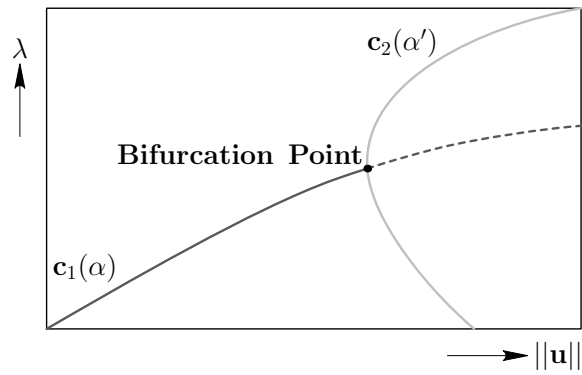


Figure 2.6: An example of a bifurcation point found along the equilibrium path $c_1(\alpha)$.

a symmetry-preserving bifurcation point the path bifurcating away from the principal path will be of the same symmetry group as the principal path. For example, referring to Fig. 2.6, if the principal path $c_1(\alpha)$ is G -symmetric and the bifurcating path $c_2(\alpha')$ is H -symmetric (for some symmetry groups G and H , respectively), then $G = H$. Additionally, it can be shown that symmetry-preserving bifurcation points are degenerate or structurally unstable in the sense that such a bifurcation point will be transformed into a turning point under any “generic” perturbation of the equilibrium equations (*Budiansky (1974)*). For this reason, symmetry-preserving bifurcation points will not be considered further in this work.

3. *Symmetry-breaking bifurcation points.*

Symmetry-breaking bifurcation points are bifurcation points where the symmetry of the path bifurcating away from the principal path is of a different symmetry group than the principal path. So, in the example shown in Fig. 2.6, if the principal path $c_1(\alpha)$ is G -symmetric and the bifurcating path $c_2(\alpha')$ is H -symmetric, then $G \neq H$. While in principle it may be that H is of higher symmetry than G , it will be assumed for the rest of this discussion that all paths bifurcating away from the principal path at a symmetry-breaking bifurcation point will be a path of lower symmetry.

2.2 Lyapunov-Schmidt-Koiter decomposition and asymptotic bifurcation analysis

In order to branch-follow a path bifurcating away from a bifurcation point it may be necessary to seed the numerical scheme with information regarding the bifurcating path (for example, the tangent of the bifurcating path at the bifurcation point). The LSK decomposition technique is a method to obtain an asymptotic expression for the path bifurcating away from the bifurcation point. The information obtained from the decomposition may be used to aid in the branch-following of a bifurcating path. The presentation here of the LSK decomposition technique follows that given in *Triantafyllidis and Peek (1992)* and *Elliott et al. (2002)*.

2.2.1 LSK decomposition

Assume that a bifurcation point has been detected along a principal path. Denote $(\mathbf{u}^c, \lambda^c)$ as the symmetry-breaking bifurcation point. Recall that, by definition, $(\mathbf{u}^c, \lambda^c)$ must satisfy conditions given in Eq. (2.3) and Eq. (2.4). In addition, let \mathbf{F} be obtained from a potential

function \mathcal{E} such that

$$\mathbf{F} \equiv \nabla_{\mathbf{u}} \mathcal{E}. \quad (2.6)$$

The equilibrium equations to be solved are then given by

$$\mathbf{F}(\mathbf{u}, \lambda) \equiv \nabla_{\mathbf{u}} \mathcal{E}(\mathbf{u}, \lambda) = \mathbf{0}. \quad (2.7)$$

The main idea for the LSK decomposition is to separate the equilibrium equations into two orthogonal subspaces, one of them being the null-space associated with the symmetry-breaking bifurcation point. Let the null-space \mathcal{N} be defined as follows

$$\mathcal{N} = \ker(\nabla_{\mathbf{u}} \mathbf{F}(\mathbf{u}^c, \lambda^c)) = \text{span}\{\phi_1, \dots, \phi_h\}, \quad (2.8)$$

where $\phi_i \in \mathbb{R}^N$, $i = 1, \dots, h$ forms an orthonormal basis of \mathcal{N} . The orthogonal complement to the null-space \mathcal{N}^\perp is defined as

$$\mathcal{N}^\perp = \{\mathbf{v} \in \mathbb{R}^N \mid \langle \mathbf{v}, \boldsymbol{\xi} \rangle = 0, \forall \boldsymbol{\xi} \in \mathcal{N}\}. \quad (2.9)$$

Furthermore, denote $h = \dim(\mathcal{N})$ and $N - h = \dim(\mathcal{N}^\perp)$. With this definition then, $\mathbb{R}^N = \mathcal{N} \oplus \mathcal{N}^\perp$ and the state vector \mathbf{u} can be decomposed as

$$\mathbf{u} = \hat{\mathbf{u}} + \boldsymbol{\xi} + \mathbf{v}, \quad (2.10)$$

where $\hat{\mathbf{u}}(\lambda)$ denotes the principal solution, $\boldsymbol{\xi} \in \mathcal{N}$, and $\mathbf{v} \in \mathcal{N}^\perp$. A projection operator $\mathbf{P} : \mathbb{R}^N \rightarrow \mathcal{N}$ can be defined as follows

$$\mathbf{P}^\top = [\phi_1 \dots \phi_h \ \mathbf{0}_1 \dots \mathbf{0}_{N-h}], \quad (2.11)$$

where $\mathbf{0}_i \in \mathbb{R}^N$, $i = 1, \dots, N - h$ denotes the zero vector. Then, by definition, it must be that $(\mathbf{I} - \mathbf{P}) \in \mathbb{R}^{N \times N}$ is such that $(\mathbf{I} - \mathbf{P}) : \mathbb{R}^N \rightarrow \mathcal{N}^\perp$.

Using the projection operators \mathbf{P} and $(\mathbf{I} - \mathbf{P})$ the equilibrium equations (Eq. (2.7)) can be projected into the two orthogonal subspaces as follows:

$$\mathbf{P} \mathbf{F}(\hat{\mathbf{u}} + \boldsymbol{\xi} + \mathbf{v}, \lambda^c + \Delta\lambda) = \mathbf{0} \in \mathcal{N}, \text{ and} \quad (2.12)$$

$$(\mathbf{I} - \mathbf{P}) \mathbf{F}(\hat{\mathbf{u}} + \boldsymbol{\xi} + \mathbf{v}, \lambda^c + \Delta\lambda) = \mathbf{0} \in \mathcal{N}^\perp. \quad (2.13)$$

In this manner, the original equilibrium equations are divided into two sets of equilibrium equations that must *simultaneously be satisfied* in order to satisfy Eq. (2.7).

Equation (2.13) defines a hyper-surface $\mathbf{v}(\boldsymbol{\xi}, \Delta\lambda) \in \mathcal{N}^\perp$ and represents a condition that must be satisfied in order for equilibrium to be achieved. $\mathbf{v}(\boldsymbol{\xi}, \Delta\lambda) \in \mathcal{N}^\perp$ can then be substituted into Eq. (2.12) to obtain the h -dimensional *bifurcation equation* $E_{,\boldsymbol{\xi}}(\boldsymbol{\xi}, \Delta\lambda) = \mathbf{0}$ given as follows

$$E_{,\boldsymbol{\xi}}(\boldsymbol{\xi}, \Delta\lambda) \equiv \mathbf{P}\mathbf{F}(\hat{\mathbf{u}} + \boldsymbol{\xi} + \mathbf{v}(\boldsymbol{\xi}, \Delta\lambda), \lambda^c + \Delta\lambda) = \mathbf{0}. \quad (2.14)$$

2.2.2 Asymptotic bifurcation analysis

Note that one solution to the bifurcation equation (with $(\boldsymbol{\xi}, \Delta\lambda) = (\mathbf{0}, \Delta\lambda)$, where $\Delta\lambda \neq 0$) is the principal path $\hat{\mathbf{u}}(\lambda^c + \Delta\lambda)$. To obtain the remaining solutions to the bifurcation equation (i.e., bifurcating paths), define a *bifurcation amplitude parameter* ζ which is the projection of a bifurcating path along the bifurcating tangent $\boldsymbol{\alpha}^{(1)}$ as follows:

$$\zeta \equiv \langle \boldsymbol{\alpha}^{(1)}, \boldsymbol{\xi}(\zeta) \rangle. \quad (2.15)$$

The bifurcating paths are parameterized by a power series in ζ where

$$\boldsymbol{\xi}(\zeta) = \boldsymbol{\alpha}^{(1)}\zeta + \frac{1}{2}\boldsymbol{\alpha}^{(2)}\zeta^2 + \dots, \text{ and} \quad (2.16)$$

$$\Delta\lambda(\zeta) = \Lambda_1\zeta + \frac{1}{2}\Lambda_2\zeta^2 + \dots \quad (2.17)$$

Furthermore, $\boldsymbol{\alpha}^{(i)}$, $i \in \mathbb{N}$, is defined such that

$$\langle \boldsymbol{\alpha}^{(i)}, \boldsymbol{\alpha}^{(1)} \rangle = \delta_{i1}, \quad \forall i \in \mathbb{N}, \quad (2.18)$$

where δ_{ij} is the Kronecker delta. Note that in Eq. (2.18), $\boldsymbol{\alpha}^{(1)}$ is defined to be of norm equal to one by convention. Furthermore, $\langle \boldsymbol{\alpha}^{(i)}, \boldsymbol{\alpha}^{(1)} \rangle = 0$, $\forall i > 1$ is a consequence of the definition of ζ .

The expansion of $\boldsymbol{\xi}(\zeta)$ (Eq. (2.16)) and $\Delta\lambda(\zeta)$ (Eq. (2.17)) is substituted into the bifurcation equation (Eq. (2.14)) which is then expanded with respect to ζ to investigate solution paths near the bifurcation point $(\mathbf{u}^c, \lambda^c) \leftrightarrow (\boldsymbol{\xi} = \mathbf{0}, \Delta\lambda = 0)$. Note that the first non-zero terms in the expansion are of $O(\zeta^2)$ due to equilibrium. At this point, two cases can be identified:

1. Asymmetric bifurcation

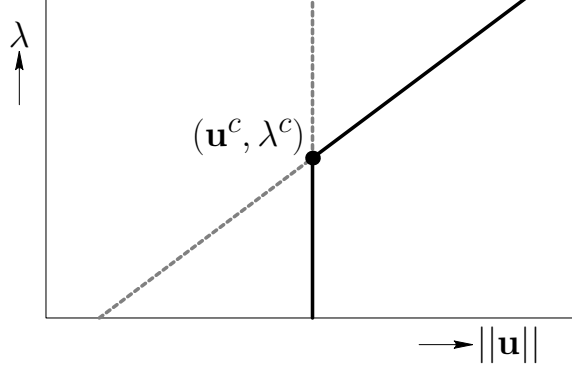


Figure 2.7: A schematic description of an asymmetric bifurcation path bifurcating from a bifurcation point $(\mathbf{u}^c, \lambda^c)$.

An asymmetric bifurcation is schematically shown in Fig. 2.7. Asymmetric bifurcations are characterized by the non-zero third order term in the expansion of the bifurcation equation. That is, an asymmetric bifurcation satisfies the following condition:

$$E_{,\xi\xi\xi} \neq \mathbf{0}, \quad (2.19)$$

where $[E_{,\xi\xi\xi}]_{ijk} = \frac{\partial^2 [E_{,\xi}]_i}{\partial \xi_j \partial \xi_k}$. This gives $O(\zeta^2)$ as the first non-zero term in the expansion of the bifurcation equation as follows:

$$E_{,\xi\xi\xi} \boldsymbol{\alpha}^{(1)} \boldsymbol{\alpha}^{(1)} + 2\Lambda_1 E_{,\xi\xi\lambda} \boldsymbol{\alpha}^{(1)} = \mathbf{0}, \quad (2.20)$$

where $[E_{,\xi\xi\lambda}]_{ij} = \frac{\partial^2 [E_{,\xi}]_i}{\partial \xi_j \partial \lambda}$. Eq. (2.20) can then be solved for $\boldsymbol{\alpha}^{(1)}$ and Λ_1 . These first order terms in the parameterization of $\boldsymbol{\xi}(\zeta)$ and $\Delta\lambda(\zeta)$ can be used to obtain the tangent to the bifurcation path that can then be used to numerically seed a branch-following algorithm.

The higher order terms in the parameterization of $\boldsymbol{\xi}(\zeta)$ and $\Delta\lambda(\zeta)$ can be solved for by using the higher order terms in the expansion of the bifurcation equation. These higher order terms can generically be written as follows, where at any given $O(\zeta^i)$:

$$\mathbf{B}\boldsymbol{\alpha}^{(i-1)} = \mathbf{b}^{(i)}, \quad \forall i \geq 3, \quad (2.21)$$

where

$$\mathbf{B} = E_{,\xi\xi\xi} \boldsymbol{\alpha}^{(1)} + \Lambda_1 E_{,\xi\xi\lambda}. \quad (2.22)$$

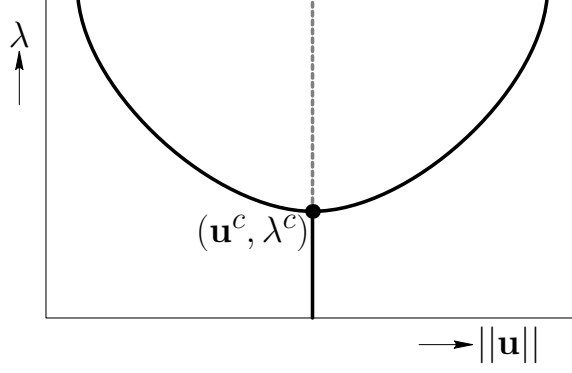


Figure 2.8: A schematic description of a symmetric bifurcation path bifurcating from a bifurcation point $(\mathbf{u}^c, \lambda^c)$.

Note that \mathbf{B} depends on $\alpha^{(1)}$ and in general, $\det(\mathbf{B}) \neq 0$. $\mathbf{b}^{(i)}$ depends on Λ_j and $\alpha^{(j)}$, $j < i$. For example,

(a) $O(\zeta^3)$:

$$\mathbf{b}^{(3)} = -\frac{1}{3} [\mathbf{E}_{,\xi\xi\xi\xi} \alpha^{(1)} \alpha^{(1)} \alpha^{(1)} + 3\Lambda_1 \mathbf{E}_{,\xi\xi\xi\lambda} \alpha^{(1)} \alpha^{(1)} + 3\Lambda_1^2 \mathbf{E}_{,\xi\xi\lambda\lambda} \alpha^{(1)} + 3\Lambda_2 \mathbf{E}_{,\xi\xi\lambda} \alpha^{(1)}],$$

(b) $O(\zeta^4)$:

$$\begin{aligned} \mathbf{b}^{(4)} = & -\frac{1}{4} [\mathbf{E}_{,\xi\xi\xi\xi\xi} \alpha^{(1)} \alpha^{(1)} \alpha^{(1)} \alpha^{(1)} + 6\mathbf{E}_{,\xi\xi\xi\xi} \alpha^{(1)} \alpha^{(1)} \alpha^{(2)} + 3\mathbf{E}_{,\xi\xi\xi} \alpha^{(2)} \alpha^{(2)} \\ & + 4\Lambda_1 \mathbf{E}_{,\xi\xi\xi\lambda} \alpha^{(1)} \alpha^{(1)} \alpha^{(1)} + 12\Lambda_1 \mathbf{E}_{,\xi\xi\xi\lambda} \alpha^{(1)} \alpha^{(2)} + 6\Lambda_1^2 \mathbf{E}_{,\xi\xi\lambda\lambda} \alpha^{(1)} \alpha^{(1)} \\ & + 6\Lambda_1^2 \mathbf{E}_{,\xi\xi\lambda\lambda} \alpha^{(2)} + 4\Lambda_1^3 \mathbf{E}_{,\xi\xi\lambda\lambda\lambda} \alpha^{(1)} + 6\Lambda_2 \mathbf{E}_{,\xi\xi\xi\lambda} \alpha^{(1)} \alpha^{(1)} + 6\Lambda_2 \mathbf{E}_{,\xi\xi\lambda} \alpha^{(2)} \\ & + 12\Lambda_1 \Lambda_2 \mathbf{E}_{,\xi\xi\lambda\lambda} \alpha^{(1)} + 4\Lambda_3 \mathbf{E}_{,\xi\xi\lambda} \alpha^{(1)}]. \end{aligned}$$

2. Symmetric bifurcation

A symmetric bifurcation is schematically shown in Fig. 2.8. Symmetric bifurcations are characterized by the third order term in the expansion of the bifurcation equation being equal to zero as reflected by the following equation:

$$\mathbf{E}_{,\xi\xi\xi} = \mathbf{0}. \quad (2.23)$$

In a symmetric bifurcation $O(\zeta^2)$ yields $\Lambda_1 = 0$ because of Eq. (2.23). Furthermore, at this stage, $\alpha^{(1)}$ is still undetermined and as such necessitates the use of $O(\zeta^3)$ terms in the expansion as given by the following equation:

$$\mathbf{E}_{,\xi\xi\xi\xi} \alpha^{(1)} \alpha^{(1)} \alpha^{(1)} + 3\Lambda_2 \mathbf{E}_{,\xi\xi\lambda} \alpha^{(1)} = \mathbf{0}, \quad (2.24)$$

where $[\mathbf{E}_{,\xi\xi\xi\xi}]_{ijkl} = \frac{\partial^3 [\mathbf{E}_{,\xi}]_i}{\partial \xi_j \partial \xi_k \partial \xi_l}$. Eq. (2.24) can be used to solve for $\alpha^{(1)}$ and Λ_2 . As

in the asymmetric case the higher order terms in the expansion of the bifurcation equation can be used to solve for each coefficient in the parameterizations of $\xi(\zeta)$ and $\Delta\lambda(\zeta)$. These higher order terms in the expansion of the bifurcation equation can, in general, be written as follows, where at any given $O(\zeta^i)$:

$$\mathbf{B}\alpha^{(i-2)} = \mathbf{b}^{(i)}, \quad \forall i \geq 4, \quad (2.25)$$

and where \mathbf{B} is given by

$$\mathbf{B} = E_{,\xi\xi\xi\xi} \alpha^{(1)} \alpha^{(1)} + \Lambda_2 E_{,\xi\xi\xi\lambda}. \quad (2.26)$$

Again, note that \mathbf{B} depends on $\alpha^{(1)}$ with $\det(\mathbf{B}) \neq 0$ and $\mathbf{b}^{(i)}$ depends on Λ_j and $\alpha^{(j)}$, $j < i$. For example, $O(\zeta^4)$:

$$\mathbf{b}^{(4)} = -\frac{1}{6}[E_{,\xi\xi\xi\xi\xi} \alpha^{(1)} \alpha^{(1)} \alpha^{(1)} \alpha^{(1)} + 6\Lambda_2 E_{,\xi\xi\xi\lambda} \alpha^{(1)} \alpha^{(1)} + 4\Lambda_3 E_{,\xi\xi\xi\lambda} \alpha^{(1)}].$$

The LSK decomposition and asymptotic bifurcation analysis of the equilibrium equations can be used to obtain an asymptotic expression for the bifurcating paths up to any arbitrary order. The BFB method combines the branch-following method (described in section 2.1) and the bifurcation point analysis using the LSK decomposition (described in section 2.2) to systematically obtain solutions to problems satisfying Eq. (2.2). Note that in an asymptotic bifurcation analysis quantities such as $E_{,\xi\xi\xi\lambda}$, $E_{,\xi\xi\xi\xi}$, or $E_{,\xi\xi\xi\xi\xi}$ are needed in Eq. (2.20) and Eq. (2.23). These quantities are computationally expensive to compute and as such, any knowledge of the form of these quantities will imply an increase in computational efficiency. One benefit in problems with symmetry is that the presence of symmetry will dictate the form of any quantities obeying the symmetry of the problem. This implication as well as other useful results from symmetry are discussed in Appendix A, beginning with the natural mathematical language that can be used to describe symmetry: group representation theory. Refer to Appendix B for an example of the application of the BFB method using a one-dimensional model of a phase transforming material. Additionally,

Chapter 3

Branch-following and bifurcation study of a typical shape memory alloy

In order to explore the features of modulated martensites a BFB study of a three-dimensional model of a phase-transforming material will be performed. The model used in this work was specifically developed by *Guthikonda and Elliott (2011)* to model AuCd, which is a known shape memory alloy. Gold-cadmium was one of the first shape memory alloys to be discovered. *Olander (1932)* used electrochemical techniques on AuCd to identify the austenite and martensite phases. Since then various studies have been performed on AuCd to verify that the austenite phase and the martensite phase of AuCd were found to be body-centered cubic (ignoring atomic species) and B19 orthorhombic, respectively (for example *Bystrom and Almin (1947)*). While the *Guthikonda and Elliott (2011)* model was tailored to model the behavior of AuCd, their results also demonstrated the model's capability to capture the general behavior of β -phase alloys. As such, even though this is labelled as a model for AuCd, we view the model as a generic model for the behavior of typical SMAs. Additionally, the B2 to B19 transformation implies a dynamic instability associated with wavevectors in the $(110)_{\text{cubic}}$ direction which is the primary mode in the phase transition into modulated martensites for β -phase alloys. For this reason, the *Guthikonda and Elliott (2011)* is an ideal model for exploring the existence of modulated martensites.

3.1 Problem description

This section will briefly detail the material model that was developed by *Guthikonda and Elliott (2011)* in order to capture the behavior of the SMA AuCd. The objective of this

section is to provide some background information on the material model that was used in the BFB study. As such, an extensive discussion of the development of the model will not be presented here. For a more detailed presentation of the model, refer to *Guthikonda and Elliott (2011)*. The remainder of this section will be organized as follows. The next subsection will introduce the *effective interatomic potential* model that describes a general class of constitutive models of a material. Following that, the Cauchy-Born kinematics will be reviewed. Once the atomic interactions have been determined the free energy density can be computed and the equilibrium equations determined. The last subsection will discuss the specifics of the model developed by *Guthikonda and Elliott (2011)*.

3.1.1 Effective interaction potential model

One way to model the behavior of a material is by a phenomenological thermomechanical *continuum* modeling to fit the relationship of the continuum concepts of stress, strain, and temperature. These types of “behavior fitting” are based on macroscopic testing and observation of the relationship between the relevant continuum level quantities (e.g., *Helm and Haupt (2003)*, *Savi et al. (2007)*). Another way to model the behavior is similar to that of the continuum level but is now based on the observations of finite-temperature X-ray diffraction (and higher resolution) experiments. This is a phenomenological thermomechanical *discrete* modeling theory and describes the relationship between the time-averaged equilibrium atomic structure of a material and the applied thermomechanical loads and temperature of a given specimen. This type of theory lies intermediate to the continuum level and the atomistic level in that it describes continuum quantities (e.g., stress, strain) in terms of atomistic quantities (e.g., atomic positions) (e.g., *Natsuki and Endo (2004)*, *Chen (2006)*). Relevant to this work is a type of a discrete modeling theory known as the Effective Interaction Potential (EIP) model.

EIP models are based on the hypothesis that the free energy of the system may be partitioned into individual atom contributions dependent only on the relative distances between the atomic nuclei and temperature (*Guthikonda and Elliott (2009)*). *Guthikonda and Elliott (2011)* developed an EIP model for the behavior of a perfect, infinite crystal of AuCd which will be used in this work. Before specific details of the EIP model are presented, it is instructive to review the concepts of the CB kinematics, free energy density, and equilibrium equations that will be used to solve for the equilibrium configuration of the crystal at a given thermomechanical loading.

3.1.2 Review: Cauchy-Born kinematics, free energy density, equilibrium equations

In order to describe a perfect, infinite crystal, a representative unit cell is chosen and its associated lattice vectors $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ (with reciprocal vectors $(\mathbf{e}^1, \mathbf{e}^2, \mathbf{e}^3)$) identified. The representative unit cell may or may not be the essential unit cell. If the representative unit cell happens to be an essential unit cell an M -lattice crystal will have M atoms in it. One method to capture period-extending phase transformations is to use a representative unit cell that consists of nM atoms where $n > 1$. Moreover, n can be chosen appropriately so that modulated phases up to a (predetermined) desired number of wavelengths can be captured. However, this method has the disadvantage of increased computational cost associated with simulating atoms in non-essential unit cells for paths where the crystal configuration can be described by an essential unit cell. To remedy this problem an additional level of sophistication can be added such that every path is computed using an essential unit cell provided that the presence of period-extending bifurcation points can be detected. This method utilizes the results from phonon stability calculations (see Ch. 1) at every equilibrium point to evaluate the dynamic stability of deformation modes associated with a prescribed set of wavevectors. The set of wavevectors in this case are those given by

$$\mathcal{K} = \left\{ \mathbf{k} = \frac{n}{c} (\mathbf{e}^1 + \mathbf{e}^2) \mid \frac{n}{c} \leq \frac{1}{2}, n, c \in \mathbb{Z} \right\} \quad (3.1)$$

and correspond to the primary mode for the transformation of BCC to modulated martensites¹. With the relevant wavevectors identified, a technique known as *Cascading Cauchy-Born* (Dobson *et al.* (2008)) can be used to optimize the choice of unit cells in the simulation. However, prior to the discussion on the Cascading Cauchy-Born technique a review of the details of the Cauchy-Born kinematics will be presented in order to provide the reader with some background.

3.1.2.1 Cauchy-Born kinematics

Deformations of the crystal can be described using Cauchy-Born (CB) kinematics (see, for example *Ericksen* (1984), *Bhattacharya* (2003), *Elliott et al.* (2006b)) where the reference position of the α -th atom in the $\ell = (\ell^1, \ell^2, \ell^3)$ unit cell is given by

$$\mathbf{X}_{(\ell, \alpha)} = \mathbf{X}_\ell + \mathbf{P}_\alpha, \quad (3.2)$$

¹The mode is associated with periodic shuffling of $(110)_{\text{cubic}}$ planes where the requirement that $\frac{n}{c} \leq \frac{1}{2}$ is due to the fact that wavevectors with $\frac{n}{c} > \frac{1}{2}$ can be mapped to wavevectors with $\frac{n}{c} \leq \frac{1}{2}$ (see for example, *Kittel* (1986) or *Ashcroft and Mermin* (1976)).

where $\mathbf{X}_\ell = \sum_{i=1}^3 \ell^i \mathbf{e}_i$ (\mathbf{e}_i are the crystal's lattice vectors and $\ell^i \in \mathbb{Z}$) gives the position vector of the ℓ -th skeletal unit cell. \mathbf{P}_α are the fractional position vectors which define the offset of the sub-lattices with respect to a skeletal lattice origin. Cauchy-Born kinematics describes the crystal's deformation in terms of a 3×3 symmetric right stretch tensor \mathbf{U} and a set of "internal atomic shift vectors" \mathbf{S}_α . Thus, in the current configuration the position of atom α in the ℓ -th unit cell is given by

$$\mathbf{x}_{(\ell,\alpha)} = \mathbf{U} (\mathbf{X}_{(\ell,\alpha)} + \mathbf{S}_\alpha). \quad (3.3)$$

Often, to eliminate rigid-body modes the shift of one sub-lattice in the crystal is set to zero (say, $\mathbf{S}_0 = \mathbf{0}$). However, this constraint artificially "pins" one sub-lattice in space which may be undesirable for branch-following and bifurcation studies. Other methods to remove the translation DOFs exist. Two such methods known as *Projection* and *Phantom Energy* methods are detailed in *Jusuf* (2010). Rigid-body rotations have already been eliminated through the use of the symmetric right stretch tensor \mathbf{U} instead of the general deformation gradient \mathbf{F} .

3.1.2.2 Cascading Cauchy-Born

Often an essential unit cell description is chosen for reasons such as computational efficiency. However, a non-essential description must be chosen when translation-symmetry breaking (i.e., period-extending) phase transformations take place. In general, the suitable choice of unit cell is not known a priori since it depends on the evolving deformation of the crystal, which in turn depends on the applied boundary condition. The Cascading Cauchy-Born (CCB) technique allows for the use of an essential unit cell at every step by identifying the loss of stability associated with a period-extending phase transformation and changing the unit cell as necessary to capture period-extending deformations. *Sorkin et al.* (submitted 2012) detail the CCB technique that can be used for a general period-extending phase transformation. However, the search for the existence of modulated martensites allows for a more targeted use of the CCB technique. This is due to the nature of the instabilities associated with modulated martensites as belonging to a family of wavevectors \mathcal{K} defined in Eq. (3.1). Therefore, the identification of period-extending bifurcation points associated with modulated martensites amounts to performing phonon stability analyses, with wavevectors from the aforementioned family, to detect the existence of period-extending bifurcation points where it is necessary to adopt a larger unit cell description of the crystal in order to capture the deformation of the bifurcating equilibrium path. The periodicity of that deformation is then c times the current unit cell size. Equivalently, this is done by

detecting the point on the equilibrium path whereby some eigenvalues $(\omega^{(r)})^2$ of the global stiffness matrix $G^{ij} K_{jk}^{\circ(\ell, \ell', \alpha, \alpha')}$ (see Eq. (1.19) in Ch. 1) becomes negative. That is, a period-extending bifurcation associated with modulated martensites is detected if $(\omega^{(r)})^2 = 0$ for some $\mathbf{k} \in \mathcal{K}$.

Once an instability along the principal path (in this case, the path is of B2 symmetry) associated with a particular wavevector $\mathbf{k} \in \mathcal{K}$ has been detected, the CCB technique allows for the construction of a new unit cell that can capture the corresponding periodic deformations. *Sorkin et al.* (submitted 2012) detail an algorithm that can be used to construct such a new unit cell. The algorithm is reproduced here in Algorithm 1. As is usual with lattice problems the set of new lattice vectors is not unique. Here, $\text{gcd}(\cdot, \cdot, \cdot)$ is the greatest common divisor function and $\mathbf{G}_1, \mathbf{G}_2$ and \mathbf{G}_3 are primitive lattice vectors of the B2 structure. The algorithm *lattice reduction* used in line 13 is detailed in *Arndt et al.* (2009). The algorithm takes as input, the lattice vectors of the primitive unit cell of the principal path equilibrium structure and begins by constructing two linearly independent primitive lattice vectors— \mathbf{G}_1^* and \mathbf{G}_2^* —that are both perpendicular to \mathbf{k} . Next, the vectors are “optimized” by the lattice reduction algorithm such that they are “the shortest and most orthogonal” of such vectors. This step yields two of the three lattice vectors in the new unit cell— \mathbf{G}_1^+ and \mathbf{G}_2^+ . The final step of the algorithm computes the third lattice vector \mathbf{G}_3^+ to be the shortest vector with the property that its projection along \mathbf{k} is equal to n . As a final note, the CCB algorithm only provides the lattice vectors in the new unit cell. Hence the positions of the simulated atoms within the unit cell must be computed separately.

3.1.2.3 Free energy density and equilibrium equations

Modeling SMAs often begins with defining the free energy density of the crystal. As mentioned in Ch. 1, the most accurate description of the free energy of the material would be obtained through the use of the quantum mechanical description of atomic interactions. However, this is computationally prohibitive for the purpose of this work. Here, we have used an EIP model that was demonstrated to exhibit the generic features of the behavior of SMAs (this will be discussed in a later section). The EIP model assumes that the crystal’s bulk free energy is given by the net sum of the pairwise atomic interactions between atoms in one unit cell (say $\ell = (0, 0, 0)$) and all other atoms in the crystal. Thus, the free energy density per unit reference volume of a crystal (at temperature θ) is given by

$$\mathcal{E}(\mathbf{u}, \theta) = \frac{1}{2V} \sum_{\alpha} \sum_{\ell', \alpha'} \phi^{\eta(\alpha)\eta(\alpha')}(r_{(0, \ell', \alpha, \alpha')}, \theta), \quad (3.4)$$

Algorithm 1 CCB algorithm reproduced here from *Sorkin et al.* (submitted 2012)

```

1: Input:  $\mathbf{G}_1, \mathbf{G}_2, \mathbf{G}_3, \mathbf{k}', n,$  and  $c$ 
2:  $k_i := \frac{c}{n} \mathbf{k}' \cdot \mathbf{G}_i, \quad i = 1, 2, 3$ 
3: if  $k_3 \neq 0$  then
4:    $\mathbf{G}_1^* := (-k_3 \mathbf{G}_2 + k_2 \mathbf{G}_3) / \text{gcd}(k_2, k_3)$ 
5:    $\mathbf{G}_2^* := (-k_3 \mathbf{G}_1 + k_1 \mathbf{G}_3) / \text{gcd}(k_1, k_3)$ 
6: else if  $k_2 \neq 0$  then
7:    $\mathbf{G}_1^* := (-k_2 \mathbf{G}_1 + k_1 \mathbf{G}_2) / \text{gcd}(k_1, k_2)$ 
8:    $\mathbf{G}_2^* := \mathbf{G}_3$ 
9: else
10:   $\mathbf{G}_1^* := \mathbf{G}_2$ 
11:   $\mathbf{G}_2^* := \mathbf{G}_3$ 
12: end if
13:  $(\mathbf{G}_1^+, \mathbf{G}_2^+) := \text{latticereduction}(\mathbf{G}_1^*, \mathbf{G}_2^*)$ 
14:  $\ell^i := \min_{(\ell^i | \ell^j k_j = c)} \|\ell^j \mathbf{G}_j\|$ 
15:  $\mathbf{G}_3^+ := \ell^i \mathbf{G}_i$ 
16: Return  $(\mathbf{G}_1^+, \mathbf{G}_2^+, \mathbf{G}_3^+)$ ;

```

where $\mathbf{u} = (\mathbf{U}, \mathbf{S}_0, \mathbf{S}_1, \dots, \mathbf{S}_{N-1})$ is the set of DOFs of an N atom representative unit cell. In principle, the summation in Eq. (3.4) must be performed over all the atoms in the infinite crystal but it is typical to impose a cutoff radius r_{cut} such that the summation only extends to all atoms in an Eulerian sphere of influence of radius less than r_{cut} (i.e., all atoms such that $r < r_{cut}$). $\eta(\alpha)$ gives the species (or type) of the α -th atom. $r_{(0, \ell', \alpha, \alpha')} = \|\mathbf{x}_{(\ell', \alpha')} - \mathbf{x}_{(0, \alpha)}\|$ gives the distance between two atoms in the crystal. The factor of $1/2$ in Eq. (3.4) arises from the fact that the pairwise interaction energy is equally shared by the two relevant atoms in each term of the sum.

The zero-stress equilibrium equations for the infinite crystal at temperature θ are given by the derivatives of Eq. (3.4),

$$\frac{\partial \mathcal{E}}{\partial \mathbf{u}} = \mathbf{0} \begin{cases} \frac{\partial \mathcal{E}}{\partial \mathbf{U}} = \mathbf{0}, \\ \frac{\partial \mathcal{E}}{\partial \mathbf{S}_\alpha} = \mathbf{0}, \quad \alpha = 0, 1, \dots, N-1. \end{cases} \quad (3.5)$$

The stability criteria used in this work will be the CB stability and Phonon stability criteria discussed in Ch. 1. The stability criteria ensure that the corresponding equilibrium configuration is observable (or physical). That is, if a structure becomes unstable as the temperature crosses a critical value, then the material must undergo a phase transition into a stable equilibrium configuration.

3.1.3 Morse EIPs with temperature dependent potential parameters

The EIP model that is exclusively used in this work is a Morse pair potential developed by *Guthikonda and Elliott (2011)*. The general form of the Morse potential is given by

$$\phi(r) = A \left\{ \exp \left[-2B \left(\frac{r}{\hat{r}} - 1 \right) \right] - 2 \exp \left[-B \left(\frac{r}{\hat{r}} - 1 \right) \right] \right\}, \quad (3.6)$$

where r is the distance between two interacting atoms and \hat{r} is the zero force distance. A and B are parameters and are related to the bond strength and bond stiffness, respectively. According to *Guthikonda and Elliott (2011)*, a cutoff distance of $r_{cut} = 20\text{\AA}$ is found to provide converged results for the summation in Eq. (3.4). In the case for AuCd, each of the atomic interactions is modeled with a separate free energy potential denoted $\phi^{\text{Au-Au}}$, $\phi^{\text{Cd-Cd}}$, and $\phi^{\text{Au-Cd}}$ which model the interactions between two Au atoms, two Cd atoms, and an Au and a Cd atom, respectively. In *Guthikonda and Elliott (2011)*, the temperature dependence of the material properties is captured by making the free parameters \hat{r} , B , and A temperature dependent. The temperature dependence of the free parameters was selected such that they reproduce the experimental trends. The exact process for obtaining the temperature dependence of the free parameters will not be discussed here. Instead, it is sufficient for the purpose of this work to provide the final form of the Morse potential, which is given in the following equation:

$$\begin{aligned} \phi(r, \theta) &= A(\theta) \left\{ \exp \left[-2B(\theta) \left(\frac{r}{\hat{r}(\theta)} - 1 \right) \right] - 2 \exp \left[-B(\theta) \left(\frac{r}{\hat{r}(\theta)} - 1 \right) \right] \right\}, \quad (3.7) \\ \hat{r}(\theta) &= r_1 + r_2(\exp(r_3(\theta - 1.0)) - 1.0), \\ B(\theta) &= B_1 + B_2(\theta^{B_3} - 1.0), \\ A(\theta) &= A_1 + A_2(\theta^{A_3} - 1.0), \end{aligned}$$

where $r_1, r_2, r_3, B_1, B_2, B_3, A_1, A_2,$ and A_3 are constant parameters. The specific values for the constant parameters in the three different atomic interaction models (Au-Au, Cd-Cd, Au-Cd) can be found in Tbl. 5 in *Guthikonda and Elliott (2011)* and are reproduced here in Tbl. 3.1. The final form of the Morse potential will be used as the atomic interaction potential in this work.

3.2 BFB result of a AuCd model

The results from the BFB study of the AuCd model will be presented in this section. For a more thorough presentation of some of these results, refer to *Guthikonda and Elliott (2011)*.

Potential parameter	Final EIP parameters Value
Au–Au	
A_1^{aa}	0.482917351030788 eV
A_2^{aa}	−0.000855192503400 eV
A_3^{aa}	3.625177992551010
B_1^{aa}	4.708238552507823
B_2^{aa}	−0.083925503055104
B_3^{aa}	2.005750046077795
r_1^{aa}	3.035902348167637 Å
r_2^{aa}	0.040386678684356 Å
r_3^{aa}	0.951807321991185
Cd–Cd	
A_1^{bb}	0.151331888205594 eV
A_2^{bb}	0.002988464666570 eV
A_3^{bb}	2.080803415890498
B_1^{bb}	4.631077953243956
B_2^{bb}	−0.048032181489227
B_3^{bb}	2.117843852765962
r_1^{bb}	3.144811091489227 Å
r_2^{bb}	0.048803776186845 Å
r_3^{bb}	0.784063749845215
Au–Cd	
A_1^{ab}	0.198800688293818 eV
A_2^{ab}	0.002241110733955 eV
A_3^{ab}	2.000000020369559
B_1^{ab}	5.452742482206239
B_2^{ab}	−0.007514062001064
B_3^{ab}	3.298115567103574
r_1^{ab}	3.081961402980163 Å
r_2^{ab}	0.028018026615333 Å
r_3^{ab}	0.842955022337628

Table 3.1: Fitted parameters for the Au-47.5 atomic %Cd Morse EIP model reproduced here from *Guthikonda and Elliott (2011)*.

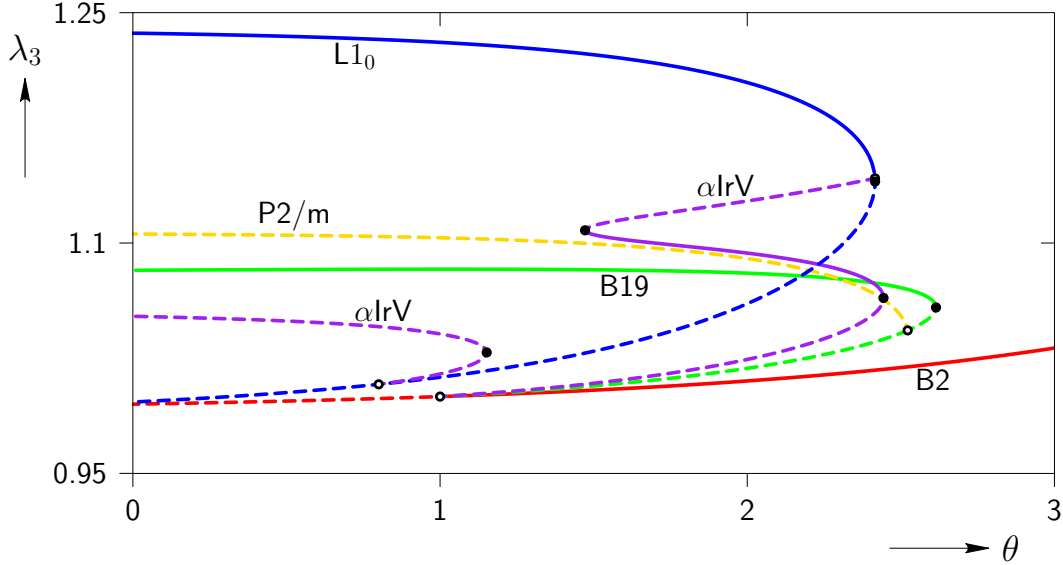


Figure 3.1: Stress-free maximum principal stretch λ_3 versus normalized temperature θ from the BFB result of AuCd using the model of *Guthikonda and Elliott (2011)*.

Fig. 3.1 is obtained from *Guthikonda and Elliott (2011)* and is a plot of one of the results of the BFB study of the AuCd model. Points on any part of the curves in the figure indicate equilibrium points on the energy landscape as projected onto the degree of freedom λ_3 (the largest eigenvalue of \mathbf{U}). Points on solid parts of the curves indicate a stable configuration (phonon and Cauchy-Born stable) whereas dashed segments indicate a configuration that is phonon unstable and/or Cauchy-Born unstable. Curves of different colors correspond to distinct structures of different symmetry. The unit cells of the various structures found are shown in Fig. 3.2, drawn in their respective colors. In Fig. 3.2 the red and blue dots are meant to indicate Au and Cd atoms, respectively.

From Fig. 3.1 we see that at high temperatures ($\theta \approx 3.0$) there is only one stable phase (shown in red) corresponding to a structure of B2 cubic symmetry. This is in agreement with experimental results for AuCd. As the temperature is decreased the B2 structure remains stable and undergoes thermal contraction. The onset of the first instability occurs at $\theta = 1.0$ at which point the B2 structure ceases to be stable. Bifurcating away from the B2 path at $\theta = 1.0$ are two paths corresponding to structures of B19 and α IrV symmetry. Both paths are initially unstable but become stable after their respective turning points. Additionally, while B19 has a wide stable segment (θ approximately 0.0 to 2.5) α IrV becomes unstable after a second turning point and then proceeds to connect near the turning point of the $L1_0$ path. Following the unstable segment of $L1_0$ we see that it connects back to the principal path (B2) around $\theta \approx 0.0$.

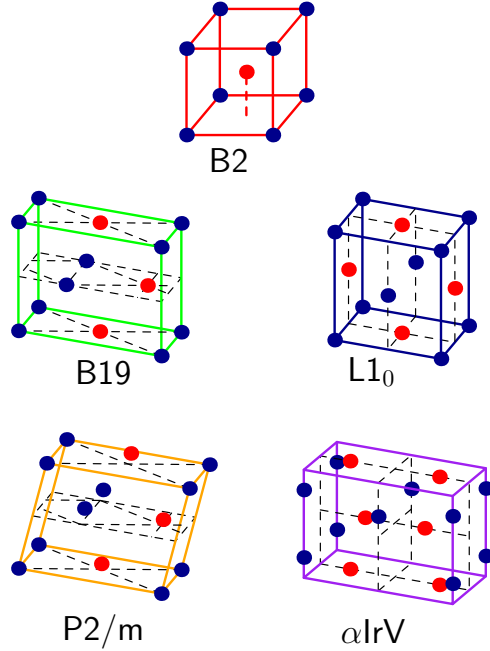


Figure 3.2: The unit cells of the various structures found in the BFB study of AuCd using the model of *Guthikonda and Elliott* (2011). The unit cells are drawn with colors corresponding to curves of the same color in Fig. 3.1.

Figure 3.3 shows a plot of the energy of the different *stable* structures found in *Guthikonda and Elliott* (2011). We see that at low temperatures ($\theta < 2.5$) the principal path is of higher energy as compared to the other structures found by *Guthikonda and Elliott* (2011). Moreover, the B19 path is found to be the lowest in energy until $\theta \approx 2.5$ whereby the B2 path is observed to be the lowest in energy for higher temperatures. Of special interest to this work are the structures B19 and $L1_0$. This is especially true in view of the energy values of these structures where we see that B19 and $L1_0$ are the two lowest energy structures found. Furthermore, since B19 is found to be the minimum energy value (between $\theta \in [0, 2.5)$), it is a likely candidate as the ground state martensite (as interpreted in the viewpoint of modulated martensites).

3.2.1 Some conclusions from *Guthikonda and Elliott* (2011)

The EIP model for AuCd developed by *Guthikonda and Elliott* (2011) was found to have characteristics that compare well with experimental observations of the behavior of AuCd SMAs. For example, the model was found to exhibit a B2 to B19 MPT which has been experimentally observed in AuCd SMAs (see for example, *Olander* (1932), or *Bystrom and Almin* (1947)). The model is capable of “matching lattice parameter, bulk modulus, cohe-

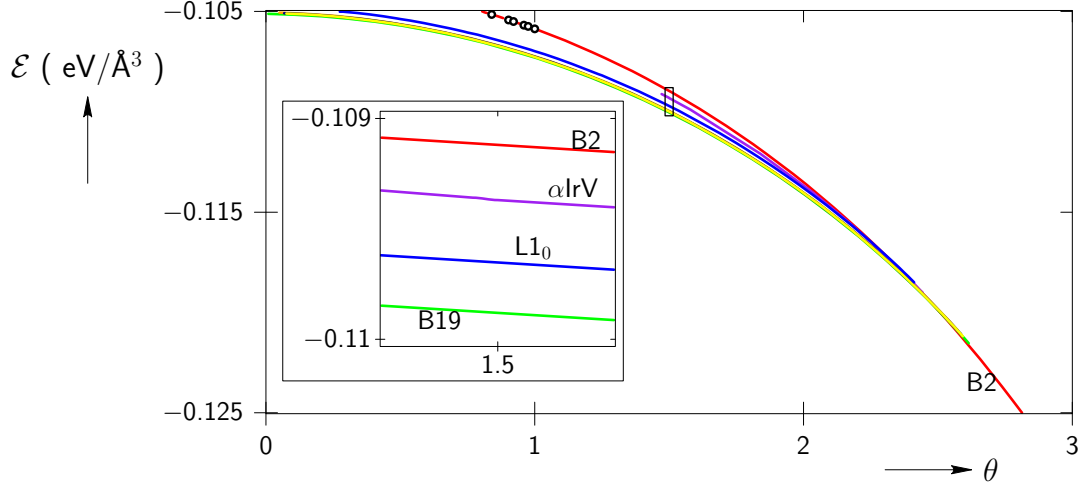


Figure 3.3: A plot of energy versus normalized temperature from the BFB result of AuCd using the model of *Guthikonda and Elliott (2011)*.

sive energy, linear thermal expansion coefficient, and heat capacity at the reference temperature and lattice parameter at temperature of 550 K ” (*Guthikonda and Elliott (2011)*). Although the model was tailored to match the behavior of AuCd SMAs, the Morse EIP was found to be capable of capturing the generic qualitative behavior of SMAs that makes it ideal for the investigation of modulated martensites in this work. For instance, the B2 to B19 MPT captured by the model is consistent with the observation of MPTs in β -phase alloys that commonly exhibit modulated martensites.

It is helpful to interpret the stability of the principal path by considering the phonon dispersion relation of the B2 structure as the temperature θ is varied. First, notice that B19 and αIrV both contain twice the number of atoms in their essential unit cells as there are in the B2 unit cell. Second, the configuration of B19 and αIrV can be described by shuffling of the family of $(110)_{\text{cubic}}$ planes. Hence, the B19 and αIrV bifurcation point is a type of *period-doubling* bifurcation point and corresponds to the onset of instability in the $(110)_{\text{cubic}}$ direction as both B19 and αIrV contain deformation modes corresponding to alternating shifts in $(110)_{\text{cubic}}$ planes (*Guthikonda and Elliott (2011)*). Therefore, from a dispersion relation perspective, this onset of instability can be interpreted as the phonon frequency ω (or more correctly, the square of ω) associated with the wavevector $\mathbf{k} = (110)_{\text{cubic}}$ becoming zero at the zone boundary shown in Fig. 3.4. Fig. 3.5 shows a schematic of a portion of the B2 crystal. The first instability is a period-doubling instability in the $(110)_{\text{cubic}}$ direction as suggested by the wave shown in Fig. 3.5 and indicated as alternating interplanar shift of $(110)_{\text{cubic}}$ planes. The grey region in Fig. 3.5 is to indicate the primitive unit cell necessary to describe this deformation mode. It contains double the number of atoms of the B2 unit

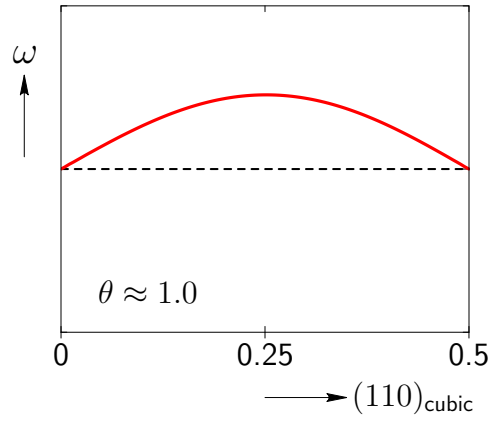


Figure 3.4: The dispersion relation showing the onset of the first instability. The horizontal dashed black line is meant to indicate $\omega = 0.0$.

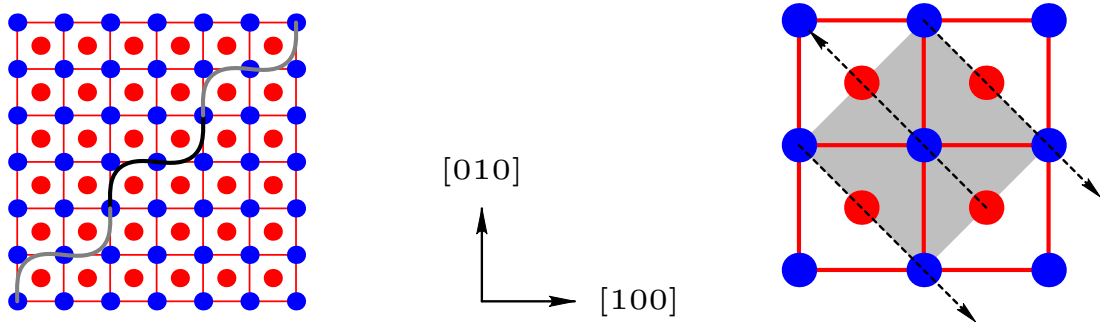


Figure 3.5: A schematic of a B2 crystal subject to period-doubling disturbances corresponding to alternating interplanar shifts of $(110)_{\text{cubic}}$ planes. The grey region indicates the primitive unit cell necessary to describe this deformation mode.

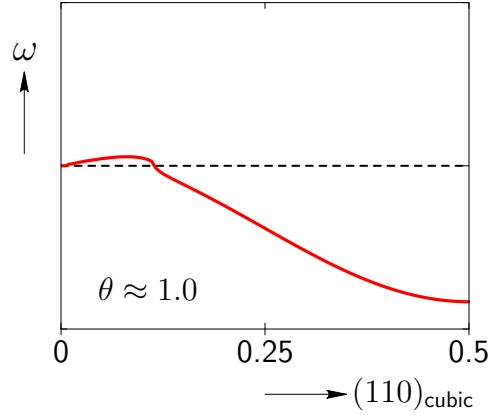


Figure 3.6: The dispersion relation showing the approach to the long wavelength instability corresponding to the Cauchy-Born instability of the B2 structure at $\theta \approx 0.0$.

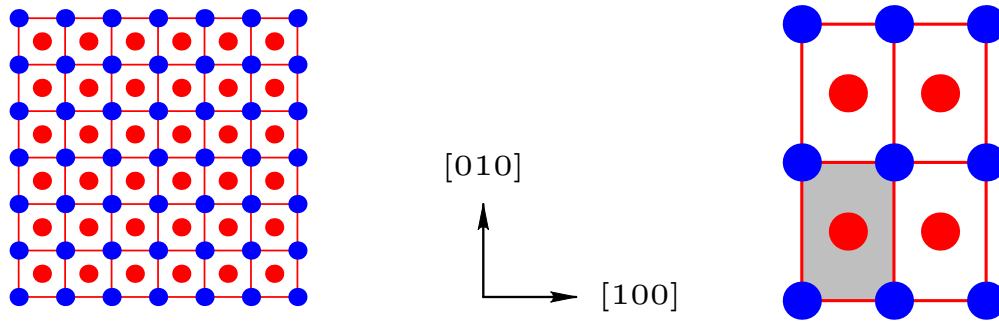


Figure 3.7: A schematic of a B2 crystal at the $L1_0$ period-preserving bifurcation point and the deformation of the primitive unit cell corresponding to the $L1_0$ structure (shown in grey).

cell and hence is a period-doubling bifurcation. On the other end of the spectrum is the bifurcation point for the $L1_0$ path which is a *period-preserving* bifurcation corresponding to a long (infinite) wavelength (Cauchy-Born) instability. Fig. 3.6 shows the dispersion relation near this instability, with Fig. 3.7 showing one example of the deformation mode and primitive unit cell necessary to capture this deformation.

3.3 Modulated martensites in a generic model of SMAs

The use of the AuCd material model of *Guthikonda and Elliott (2011)* is applicable for this work for several reasons. Firstly, while the material model was tailored to display properties of AuCd, it is the ability of the model to show the generic behavior of SMAs that is of interest to this work. Secondly, the model has been shown to capture the phase transformation from the B2 austenite to a B19 martensite which (1) has an austenite struc-

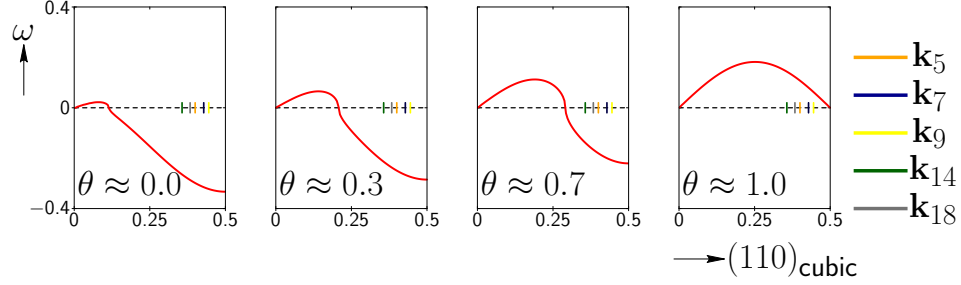


Figure 3.8: The dispersion relation of the B2 configuration as the temperature is varied from $\theta = 0.0$ to $\theta = 1.0$. As the temperature is increased the B2 structure becomes stable to perturbations of different wavelengths. Associated with the change in stability is a period-extending bifurcation point along the principal path where the bifurcating paths correspond to periodic deformations with periodicity as described by the associated wavevector.

ture common to materials that exhibit modulated martensites and (2) is related to the loss of dynamic stability associated with the wavevector $\mathbf{k} = \frac{1}{2}(110)_{\text{cubic}}$. Lastly, since B19 is an orthorhombic structure the existence of modulated martensites in this model cannot be explained by the adaptive martensite hypothesis. Therefore, this material model provides a good opportunity for the development of a new perspective for modulated martensites.

The results from the BFB study obtained by *Guthikonda and Elliott* (2011) suggest the existence of modulated martensites in this model. This is because one can view the bifurcation of the B19 and L1₀ structures from the principal path as lying on opposite extremes of the phonon frequency spectrum for the wavevectors along the $(110)_{\text{cubic}}$ direction ($\mathbf{k} \in \mathcal{K}$). That is, as the temperature is raised from $\theta = 0.0$ to $\theta = 1.0$ a progression of bifurcation points corresponding to instabilities of different wavelengths will be observed. At each point where there is a change in stability ($\omega = 0.0$) a period-extending bifurcation point occurs. The bifurcating paths from these points correspond to periodic deformations with periodicity described by the associated wavevector (the one with $\omega = 0.0$). This idea is shown in Fig. 3.8 showing several snapshots of the dispersion relation computed for the B2 structure as the temperature is increased. The different snapshots in Fig. 3.8 show the results of the dispersion relations calculations (for $\mathbf{k} \in \mathcal{K}$) at different temperatures, starting with $\theta = 0.0$ at the left-most picture in the figure and ending with $\theta = 1.0$ at the right-most picture in the figure.

In order to explore a portion of the infinite number of bifurcating equilibrium paths identified by the previous arguments, we have limited our search to wavevectors that correspond to wavelengths commonly observed in materials that exhibit modulated martensites. Additionally, the wavevectors were chosen so as to be “close” to the $\mathbf{k} = \frac{1}{2}(110)_{\text{cubic}}$ instability.

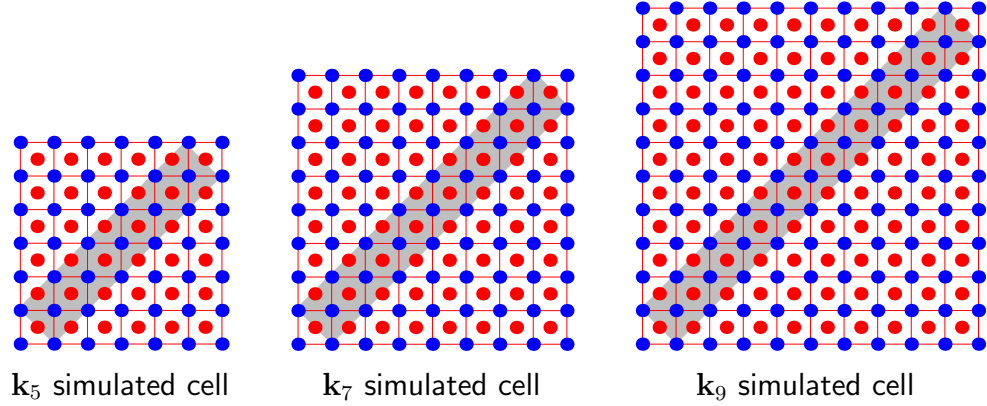


Figure 3.9: Examples of simulated cells necessary to capture period-extending deformations in β -phase alloys.

These wavevectors are shown as the tick marks in Fig. 3.8 and are given by $\mathbf{k}_5 = \frac{2}{5}(110)_{\text{cubic}}$ (shown in orange), $\mathbf{k}_7 = \frac{3}{7}(110)_{\text{cubic}}$ (shown in dark blue), $\mathbf{k}_9 = \frac{4}{9}(110)_{\text{cubic}}$ (shown in yellow), $\mathbf{k}_{14} = \frac{5}{14}(110)_{\text{cubic}}$ (shown in dark green), and $\mathbf{k}_{18} = \frac{7}{18}(110)_{\text{cubic}}$ (shown in gray). The wavevectors corresponds to deformations of five, seven, nine, fourteen, and eighteen B2 primitive unit cells in the $[110]_{\text{cubic}}$ direction (or alternatively 10, 14, 18, 28, and 36 $(110)_{\text{cubic}}$ planes). With the set of relevant wavevectors defined the principal path was re-computed with the additional task of determining each point where the phonon frequency has vanished. These points correspond to period-extending bifurcation points and the paths bifurcating from these points are modulated martensites. At these points the CCB technique can be used to set up the unit cell so as to capture deformations associated with modulated martensites. Examples of the simulated cells defined by the CCB technique for the wavevectors \mathbf{k}_5 , \mathbf{k}_7 , and \mathbf{k}_9 are shown in gray in Fig. 3.9. Using the new unit cell the eigenvectors associated with the zero eigenvalues of the stiffness matrix were computed².

Once the eigenvectors have been identified the bifurcating tangent can be obtained using sophisticated techniques such as the LSK asymptotic analysis and/or symmetry methods that were detailed in Chapter 2 and Appendix A, respectively. However, the high number of DOFs in the system made the use of the LSK asymptotic analysis highly computationally expensive³. Additionally symmetry methods were disregarded as the identification of the symmetry group associated with each respective non-essential unit cell proved to be

²Note that when using the larger unit cell the phonon instability manifests itself as a CB instability. This is because the larger unit cell is able to capture the deformations associated with the unstable wavelength.

³Recall from Ch. 2 that the asymptotic bifurcation analysis provides a parameterization for the bifurcating path in terms of the bifurcation amplitude parameter. The technique then involves solving systems of equations that can involve high order derivatives of the free energy. These derivatives tend to be highly computationally expensive to compute.

difficult. As such, the bifurcating path was obtained in an ad-hoc manner by taking one of the eigenvectors with zero eigenvalue (note that all bifurcation points detected in this work have multiplicities equal to two when the unit cells indicated in Fig. 3.9 are used) and using it as the tangent in a predictor-corrector scheme (specifically, the *Pseudo-Arclength Tangent Approach* as detailed in *Jusuf* (2010)). The scheme allows for an imprecise tangent to be used in the predictor step as an equilibrium path would serve as an attractor for the corrector routine. This approach may be highly sensitive to parameters in the corrector routine and is especially troublesome if the angle between the true bifurcating tangent and the zero-eigenvalue eigenvector is large. However, once a point on the bifurcating path has been obtained the predictor-corrector scheme would proceed to compute the remainder of the path. One disadvantage of this method is that there is no guarantee that the predictor-corrector scheme will not “switch” onto a path of different symmetry. The switching would typically manifest as a jump in the curve or a drastic change in the number of negative eigenvalues of the stiffness matrix. When situations such as these were encountered the stepsize of the predictor-corrector scheme was altered to “skip over” the point of switching. Another disadvantage of this method is that there is no way to identify when all paths bifurcating from a bifurcation point have been computed. However, this proved to be an acceptable compromise considering that this work aims to demonstrate the existence of at least one bifurcating path associated with phase transformation from the B2 austenite to a modulated martensite. As such, the identification of one such path is sufficient to satisfy the objectives of this work.

Figure 3.10 shows a plot of λ_3 (maximum principal stretch) as a function of θ for the structures found in the BFB study performed in this work. As expected, the principal path was found to be of B2 symmetry and is stable at high temperatures. The period-doubling bifurcation point at $\theta = 1.0$ was also observed with bifurcating paths corresponding to α IrV and B19 symmetry. Additionally, the structures L1₀ and P2/m were observed to behave exactly in the manner detailed in *Guthikonda and Elliott* (2011). Also shown in the figure are period-extending bifurcation points indicating the onset of stability (with increasing temperature) for the wavevectors \mathbf{k}_5 (orange), \mathbf{k}_7 (dark blue), \mathbf{k}_9 (yellow), \mathbf{k}_{14} (dark green), and \mathbf{k}_{18} (gray). The bifurcation points are found to occur near the period-doubling instability. Moreover, wavevectors with values closer to $\frac{1}{2}(110)_{\text{cubic}}$ were found to bifurcate closer to the period-doubling bifurcation point (i.e., closer to $\theta = 1.0$). Tbl. 3.2 lists the temperature at which each of the period-extending bifurcation points was found. The table also details the stability of the paths bifurcating from the B2 path. Wavevectors labeled as stable imply that at least one path bifurcating from the associated bifurcation

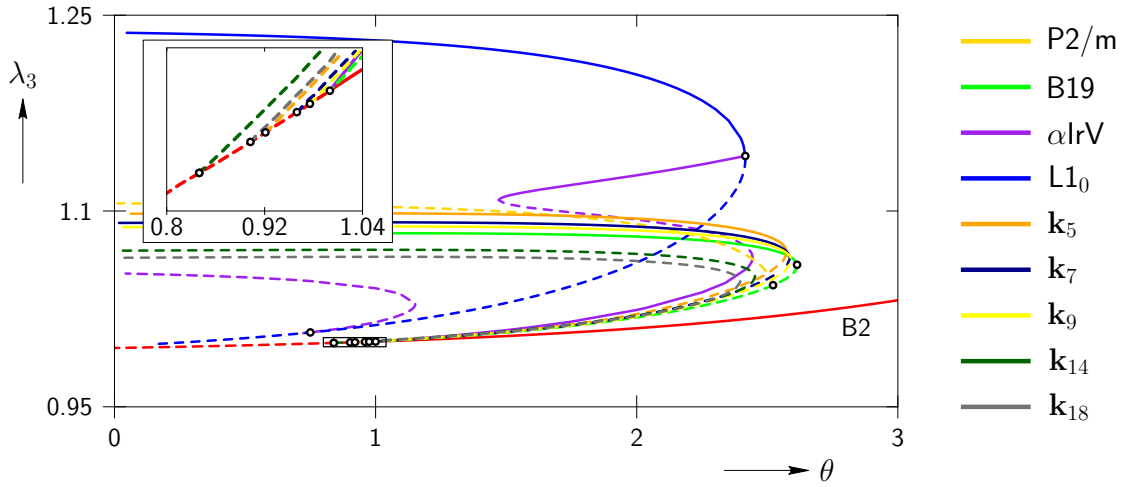


Figure 3.10: BFB study of AuCd showing period-extending bifurcation points. The paths are color coordinated as follows: \mathbf{k}_5 (orange), \mathbf{k}_7 (dark blue), \mathbf{k}_9 (yellow), \mathbf{k}_{14} (dark green), and \mathbf{k}_{18} (gray).

	Wavevector	Bifurcation Temperature	Stability
\mathbf{k}_5	$\frac{2}{5}(110)_{\text{cubic}}$	0.92098	Stable
\mathbf{k}_7	$\frac{3}{7}(110)_{\text{cubic}}$	0.95955	Stable
\mathbf{k}_9	$\frac{4}{9}(110)_{\text{cubic}}$	0.97550	Stable
\mathbf{k}_{14}	$\frac{5}{14}(110)_{\text{cubic}}$	0.83987	Unstable
\mathbf{k}_{18}	$\frac{7}{18}(110)_{\text{cubic}}$	0.90260	Unstable

Table 3.2: Bifurcation temperatures of modulated structures found in the BFB study of AuCd using the model developed by *Guthikonda and Elliott (2011)*. Structures marked as “stable” imply the path contains a stable segment.

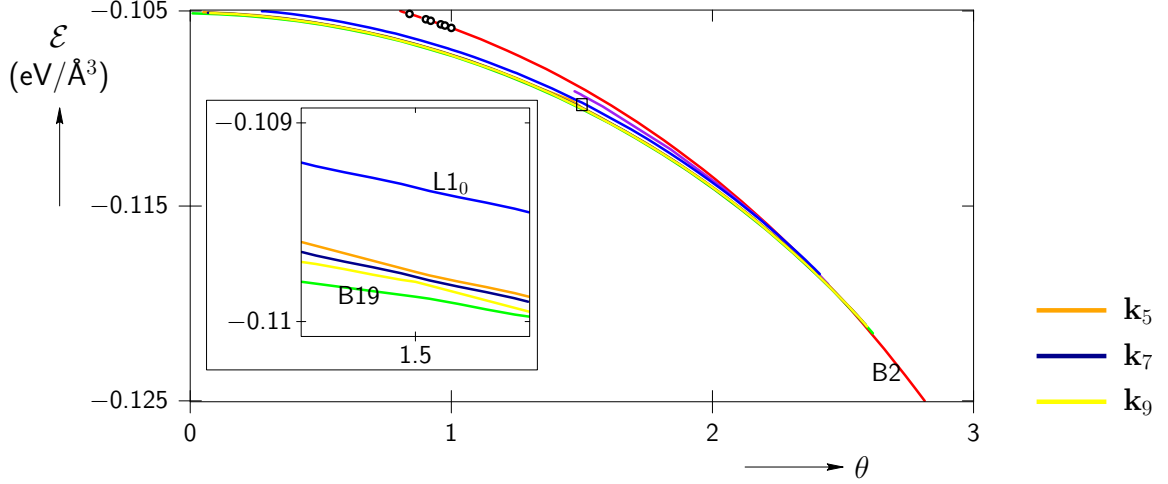


Figure 3.11: Energy of the stable structures found in the BFB study of AuCd.

point contains a stable segment. Qualitatively, the period-extending bifurcating paths were found to follow the same general behavior as the B19 path. That is, all paths bifurcating away from the principal path are initially unstable and contain one turning point. Out of the five wavevectors that were chosen only k_5 , k_7 , and k_9 contained stable segments. These modulations correspond to commonly observed modulated martensites. Moreover, the wavevectors k_{14} and k_{18} were not found to contain any stable segments.

Figure 3.11 shows a plot of the energies of the stable phases that have been found. Two immediate observations can be made from Fig. 3.11. First, the B19 structure is the global minimum in energy for this material over a wide range of temperature (especially for temperatures where the B2 structure is unstable). From the viewpoint of modulated martensites this designates B19 as the ground state martensite for this material. Second, modulated martensites are low energy phases (the energy differences between the modulated phases and B19 are small compared to the energy difference between the B19 and B2 structures). This is especially evident from the close-up of the energy at $\theta = 1.5$ shown in the inset of Fig. 3.11. The modulated structures k_5 , k_7 , and k_9 were found to be very close in energy to each other and much closer to the value of the ground state martensite B19 than any of the other structures found in *Guthikonda and Elliott (2011)*. Tbl. 3.3 lists the energy differences between the different structures at $\theta = 1.5$. The values are shown as differences relative to the energy of the ground state martensite B19. Note that the modulated martensites are of similar values in energy with k_9 being the minimum. Also note that the differences in energy between the modulated martensites and B19 are much smaller than any of the structures found in *Guthikonda and Elliott (2011)*.

	\mathcal{E} (relative (eV/Å ³))
B2	1.06761×10^{-3}
B19	0.0
α IrV	7.14629×10^{-4}
L1 ₀	3.16381×10^{-4}
\mathbf{k}_5	6.64208×10^{-5}
\mathbf{k}_7	4.75274×10^{-5}
\mathbf{k}_9	3.69999×10^{-5}

Table 3.3: Relative energy values (at $\theta = 1.5$) of the various structures found in the BFB of AuCd. The values are shown relative to the energy of the ground state martensite B19.

3.3.1 Discussion

It is important to note that the existence of modulated structures was not considered in the design of the model by *Guthikonda and Elliott (2011)*. This suggests that *the presence of modulated martensites is a general feature of the energy landscape of the model*. This information augments our understanding of the energy landscape with details not provided by the AMH or *Balandraud and Zanzotto (2007)*. By performing a BFB study of this material model, it was found that modulated martensites are in fact natural features of the energy landscape of the material and they exist as a possible configuration which is accessible to the material during a phase transformation.

The BFB study of the model also found that \mathbf{k}_{14} and \mathbf{k}_{18} do not contain any stable segments. This is because these paths contain deformation modes that have not yet become stable. Specifically, the \mathbf{k}_{14} and \mathbf{k}_{18} paths bifurcate away from the principal path at temperatures where wavevectors of lower wavelengths have not yet become phonon stable. To be more precise, in the case of \mathbf{k}_{14} and \mathbf{k}_{18} , these lower wavelength wavevectors are ones with wavelengths that are multiples⁴ of 14 and 18. These include the wavevectors \mathbf{k}_7 , \mathbf{k}_9 , and the $\mathbf{k} = \frac{1}{2}(110)_{\text{cubic}}$ which have all not yet become phonon stable. Additionally, since it was found (see Fig. 3.8) that the period-doubling instability is the last instability prior to the principal path becoming stable, one can expect that wavevectors corresponding to modulations with wavelengths equal to two will be unstable. For this reason, *modulated martensites with odd wavelength modulations are more likely to be stable*. This is summarized in the following conjecture. Should the conjecture be proven true, then this observation is one possible explanation as to why five, seven, and nine are commonly observed modulated martensite periodicities.

⁴The reason why only multiples of the wavelengths are relevant is that perturbations with wavelengths that are not multiples of the unit cell cannot be simulated using the same unit cell.

Conjecture 1. *If N is an odd wavelength modulation, then $\exists a \in \mathbb{Z}$ such that $\frac{a-1}{N} < \frac{a}{N} < \frac{1}{2}$ and the modulated martensite associated with the wavevector $\mathbf{k} = \frac{a}{N}(110)_{cubic}$ is stable.*

3.4 The effect of a kinematic compatibility constraint on modulated martensites

In the above section it was mentioned that the BFB study found that modulated martensites are natural features of the energy landscape. Furthermore, the result from the BFB study suggests that the ground state martensite is B19 and as such would be the structure that is experimentally observed. For this reason, it still is an open question as to why modulated martensites have been observed in several β -phase alloys. One idea put forth (in works such as the AMH viewpoint and *Balandraud and Zanzotto (2007)*) is that the kinematic compatibility constraint at the austenite-martensite interface would substantially raise the energy of B19. This is especially important when considering the nucleation in a martensitic phase transformation where a more compatible modulated phase may be thermodynamically more favorable. The existence of modulated martensites in experiments was then justified as being stabilized by the kinematic compatibility to the austenite.

Assuming that the austenite structure has a deformation gradient equal to the identity \mathbf{I} , the compatibility of a particular structure to the austenite is measured by how close the middle eigenvalue of the deformation gradient of a structure (denoted here as λ_2) is to 1.0 (*Bhattacharya (2003)*). The closer that λ_2 is to 1.0, the more compatible the structure. Fig. 3.12 provides a plot of the middle eigenvalue of \mathbf{U} for the B19, \mathbf{k}_5 , \mathbf{k}_7 , and \mathbf{k}_9 structures. Also shown is the plot of the middle eigenvalue of \mathbf{U} for the B2 structure. A perfectly compatible structure would have λ_2 values that match exactly with that of the B2 structure. Note that only the stable segments of B19 and the modulated structures are shown. From the figure we see that the modulated structures \mathbf{k}_5 , \mathbf{k}_7 , and \mathbf{k}_9 are more kinematically compatible than B19. This suggests that while B19 is lower in energy, the high mismatch in kinematic compatibility could indeed substantially raise its energy when forming an austenite-martensite interface. Therefore, the viewpoints of AMH and *Balandraud and Zanzotto (2007)* seem to be valid since it was found that the modulated structures are low energy phases (i.e., the differences in energy between B19 and the modulated martensites are a small fraction of the difference in energy between B19 and B2) that exhibit better kinematic compatibility. To verify that viewpoint it would be helpful to determine how kinematic compatibility constraints affect the energy of the various modulated structures found as compared to the constrained B19. A key benefit to the BFB method applied to an atomistic material model is that it is able to probe the effect of such a constraint on the energy landscape of the

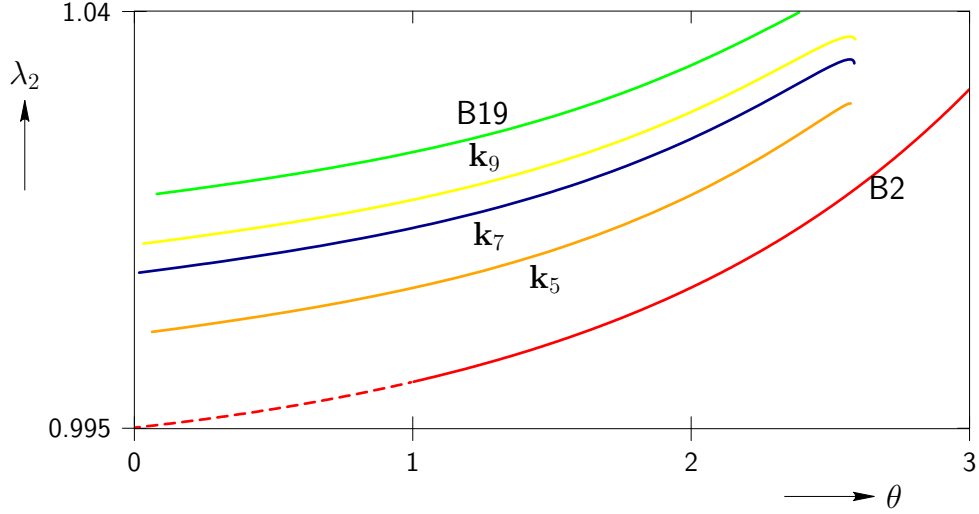


Figure 3.12: The middle eigenvalue of \mathbf{U} of the modulated structures as compared to the ground state martensite. A perfectly compatible structure would have λ_2 values that match exactly with that of the B2 structure. The plot shows λ_2 for the B2, B19, k_5 , k_7 , and k_9 structures.

material.

For the set of constrained simulations the B19, k_5 , k_7 , and k_9 structures were simulated subject to the constraint that the middle eigenvalue λ_2 of \mathbf{U} was equal to that found for B2 at $\theta = 1.5$. While a more appropriate constraint would be such that the constraint is a function of θ , the BFB code used in this work does not support this type of constraint. As such, it is important to note that conclusions derived from the results of this simulation are only strictly valid for $\theta = 1.5$. However, the relatively small change, with temperature, in the value of λ_2 for the B2 structure render the effect of this limitation minimal. For a more detailed explanation of the application of the kinematic compatibility constraint used in this work, refer to Appendix D. Figure 3.13 shows a plot of the result of the constrained calculation. The inset in Fig. 3.13 shows a close-up of the results at $\theta = 1.5$. The dashed lines indicate the value of the energy of the *unconstrained* structures. Solid lines indicate the constrained structures. Tbl. 3.4, lists the unconstrained relative energy values of B19, k_5 , k_7 , and k_9 . $\Delta\mathcal{E}$ in the table gives the increase in energy between the constrained and unconstrained structures. It is evident from both Fig. 3.13 and Tbl. 3.4 that B19 shows the largest increase in energy when constrained. The modulated structures are much better at accommodating a kinematic compatibility constraint giving smaller increases in energy as compared to B19. Moreover, k_5 was found to show the least increase in energy that is an order of magnitude smaller than that for B19. For this reason, it is reasonable to assume that k_5 would be the experimentally observed modulated structure.

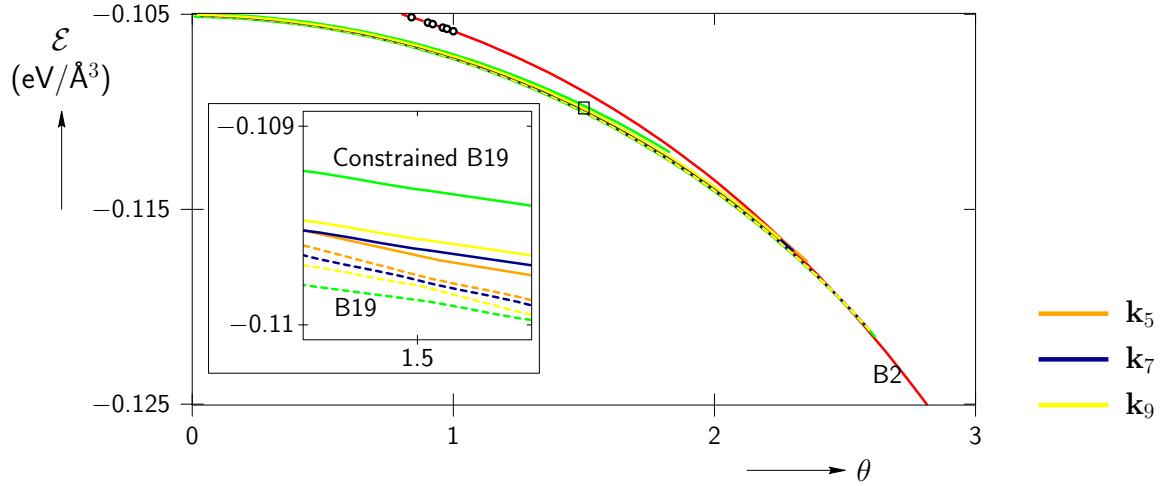


Figure 3.13: The effect of a kinematic compatibility constraint on the energy of B19, k_5 (orange), k_7 (dark blue), and k_9 (yellow). The dashed and solid lines indicate the energy of the unconstrained and constrained structures, respectively.

	\mathcal{E} (unconstrained (eV/Å ³))	$\Delta\mathcal{E}$ (eV/Å ³)	\mathcal{E} (constrained (eV/Å ³))
B19	0.0	3.43073×10^{-4}	3.43073×10^{-4}
k_5	6.64208×10^{-5}	3.65389×10^{-5}	1.02960×10^{-4}
k_7	4.75274×10^{-5}	1.04718×10^{-4}	1.52245×10^{-4}
k_9	3.69999×10^{-5}	1.55275×10^{-4}	1.92275×10^{-4}

Table 3.4: Constrained energy values of B19, k_5 , k_7 , and k_9 . $\Delta\mathcal{E}$ gives the increase in energy between the constrained and unconstrained energy values.

3.5 Summary and conclusion of the BFB study of AuCd for period-extending bifurcations

This chapter detailed the results for the atomistic simulations of modulated martensites using a material model for shape memory alloys. The model was developed by *Guthikonda and Elliott* (2011) to exhibit the generic behavior of the shape memory alloy AuCd that undergoes a martensitic phase transformation from a B2 austenite to a B19 martensite. It is important to note that this model was not developed to exhibit modulated martensites. However, the nature of the phonon dispersion relations along the principal path suggested the existence of an infinite number of modulations some of which may be associated to commonly observed modulated martensites. Five wavevectors were chosen in order to demonstrate the existence of modulated martensites in SMAs. These wavevectors are close to the period-doubling instability and correspond to modulations equal to 10, 14, 18, 28, and 36 $(110)_{\text{cubic}}$ planes. Using a method known as the Cascading Cauchy-Born technique (see *Dobson et al.* (2008) and *Sorkin et al.* (submitted 2012)), bifurcation points associated with the desired wavevectors were identified. The technique also has the added feature of detailing the construction of an appropriately sized simulated unit cell that can be used to capture deformations associated with the bifurcating path. In this manner, we extended the understanding of the SMA modeled by *Guthikonda and Elliott* (2011) by explicitly showing the existence of equilibria corresponding to stable modulated martensites. The results of this work yielded one of the first insights into the energies of modulated martensites. To date, these have not been well understood and have only been speculated as in the AMH and *Balandraud and Zanzotto* (2007). This BFB study is unique in that it gives us a clear insight into the energy and therefore relative stability of various modulated structures and the ground state martensite. It was found that modulated martensites are low energy phases in that the differences in energy between the modulated structures and B19 is small compared to the energy difference between B19 and B2.

The results from a set of constrained simulations were also detailed in this chapter. Here, paths corresponding to the structures B19, k_5 , k_7 , and k_9 were computed subject to the constraint that the middle eigenvalue of the structure was equal to that found for B2 at $\theta = 1.5$. The results show that B19 has a drastic increase in energy due to the constraint whereas the modulated structures are much better at accommodating the kinematic compatibility constraint. Additionally, the low overall energy of the constrained modulated martensite k_5 suggests that a modulated martensite with a modulation of ten $(110)_{\text{cubic}}$ planes would be experimentally observed.

Chapter 4

The modulated martensite mixture model

The BFB study detailed in Ch. 3 found that modulated martensites are natural features of the energy landscape of the material. Additionally, the dispersion relations suggest that modulated martensites of arbitrary periodicities can be found as equilibria on the energy landscape. Although the BFB method can in principle be utilized to find paths corresponding to modulated structures of all periodicities, this is not feasible in practice due to (1) the infinite number of periodicities that the modulated martensites can attain and (2) the computational power required to simulate ever larger systems. As such, a tool that can predict the properties of arbitrary modulated martensites would be useful. As part of this work we developed a predictive tool named the “Modulated Martensite Mixture Model” (M^4). This model is capable of accurately predicting the energy, lattice constants, and structural details of an arbitrary modulated martensite phase. This is demonstrated by comparing the M^4 predictions to the computational results obtained in Ch. 3. Finally, this chapter also details how the capabilities of M^4 are used to predict the existence of a structure that is highly kinematically compatible to B2. This prediction is then verified via an atomistic simulation of the predicted structure.

The main idea behind the M^4 is that modulated martensites are a mixture of two short-period *base martensites*. This view is similar to that of the ‘fundamental unit’ of *Balandraud and Zanzotto* (2007). However unlike their work, the M^4 uses *two* base martensite structures that correspond to equilibrium points on the energy landscape. The properties of these structures are available (from the BFB study detailed in Ch. 3) for use as input data. Predicting the features of an arbitrary modulated martensite then involves using the

input data along with the mixture composition to determine the final properties of the modulated structure. One benefit of the M^4 is that the calculations to obtain the properties of the base martensites are “easy” in that they only involve a small number of DOFs¹. Additionally, the base martensites belong to a symmetry group that can be easily identified. This information can be used to enhance the BFB performance due to the consequences of symmetry discussed in Appendix A. In contrast, modulated martensites have a high number of DOFs making explicit atomistic calculations computationally expensive. Furthermore, it is a non-trivial process to determine the symmetry group G of a particular modulated martensite. This makes it difficult to ensure that the path-following is restricted to the G -symmetric path (i.e., it does not ‘skip’ onto a path of different symmetry). Thus, an accurate predictive tool such as the M^4 will be valuable in determining the properties of a modulated martensite (by possibly bypassing path-following altogether).

4.1 The M^4 in detail

Despite the particular model used here being a long-range model, it is found that the primary factor that determines the overall properties of a modulated martensite is that of nearest-neighbor plane location. As such, one assumption in the M^4 is that the overall properties of a modulated martensite are determined by nearest neighbor planar interaction. The interactions are defined using input data from two short-period structures we have called base martensites. In order to get a better sense of what structures may serve as base martensites, recall that period-extending bifurcation points were found along the principal path corresponding to wavevectors from the family \mathcal{K} (defined in Eq. (3.1)). This phenomenon is captured in the dispersion relation which is reproduced from Ch. 3 in Fig. 4.1. To emphasize, notice from the dispersion relation that the bifurcation of paths corresponding to modulated martensites is flanked between two extremes corresponding to points where the $L1_0$ and $B19$ paths, respectively, bifurcate. As such, it is reasonable to expect that the *base martensites are $L1_0$ and $B19$* . Fig. 4.2 shows the unit cells of the base martensites.

In the M^4 the two base martensites (and their variants) are used to define four ‘planar unit cells’. A modulated martensite is then defined by first identifying its periodicity (number of $(110)_{\text{cubic}}$ planes in each modulated martensite unit cell) and then assigning one of four planar unit cell types to each of the $(110)_{\text{cubic}}$ planes. Therefore, the M^4 identifies the primitive unit cell of a modulated martensite through the sequence of plane types assigned to each $(110)_{\text{cubic}}$ plane in the modulated martensite unit cell. This sequence can then be

¹In fact, the input data for the model used in this work was already computed by *Guthikonda and Elliott* (2011).

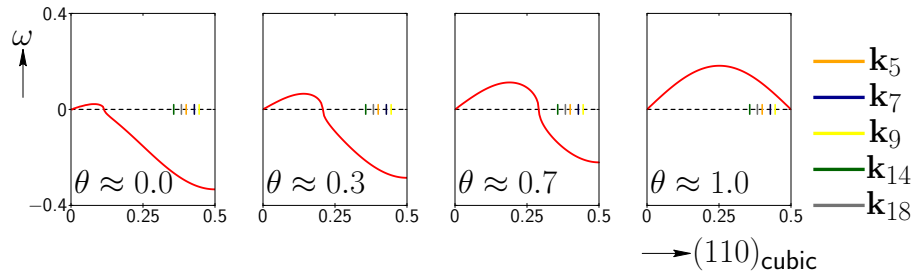


Figure 4.1: The dispersion relation of the B2 configuration as the temperature is varied from $\theta = 0.0$ to $\theta = 1.0$. As the temperature is increased the B2 structure becomes stable to perturbations of different wavelengths. Associated with each change in stability is a period-extending bifurcation point along the principal path where the bifurcating path (possibly paths) corresponds to periodic deformations with periodicity as described by the wavevector.

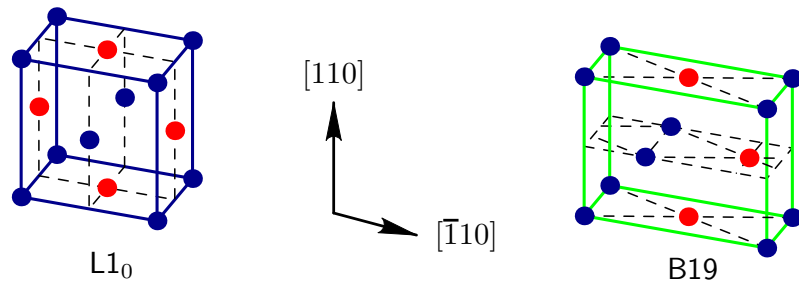


Figure 4.2: The two base martensites— $L1_0$ and B19.

used to compute the properties of the modulated martensite such as the lattice parameters and energy. The M^4 ‘constructs’ a modulated martensite from the sequence by ‘joining’ the planar unit cells sequentially. As such, the M^4 requires that the deformation of the planar unit cells satisfies the following conditions, henceforth referred to as condition (\star).

1. The deformations must preserve the orientation of $(110)_{\text{cubic}}$ planes. This condition is to make sure that the definition of modulated martensites according to the M^4 is consistent with the literature (i.e., as shuffling of $(110)_{\text{cubic}}$ planes in the $[\bar{1}10]_{\text{cubic}}$ direction).
2. Any two deformations must satisfy the Hadamard jump condition:

$$\mathbf{F}_{p_1} - \mathbf{F}_{p_2} = \mathbf{a} \otimes \mathbf{n}, \quad (4.1)$$

where $\mathbf{n} = \frac{1}{\sqrt{2}}[1, 1, 0]^\top$ in the standard basis² \mathcal{B} and \mathbf{F}_{p_1} and \mathbf{F}_{p_2} are deformation gradients associated with any two planar unit cells.

The above conditions guarantee that modulated martensites are described by a continuous deformation of the B2 structure (that preserves the orientation of $(110)_{\text{cubic}}$ planes) while allowing for jumps in the deformation gradient.

To understand the nature of the four planar unit cells consider the crystal of B2 symmetry schematically shown in Fig. 4.3. The M^4 assumes a B2 reference configuration with lattice vectors of the form given in the standard basis \mathcal{B} by:

$$\mathbf{e}_1 = N \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \quad (4.2)$$

Thus, the M^4 uses a non-primitive description of a B2 structure which contains $2N$ $(110)_{\text{cubic}}$ planes. Each of the four planar unit cells starts as an undeformed B2 planar unit cell shown in gray. The final configuration of the planar unit cell will be determined by the plane type that is assigned to it. Additionally, the use of Cauchy-Born kinematics in the BFB study makes it easier to understand the deformation of the planar unit cell as a two step process.

²The standard basis is defined as $\mathcal{B} = (\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3)$, where $\mathbf{g}_1 = [1, 0, 0]^\top$, $\mathbf{g}_2 = [0, 1, 0]^\top$, and $\mathbf{g}_3 = [0, 0, 1]^\top$ are vectors aligned with the cubic axes of the B2 crystal.

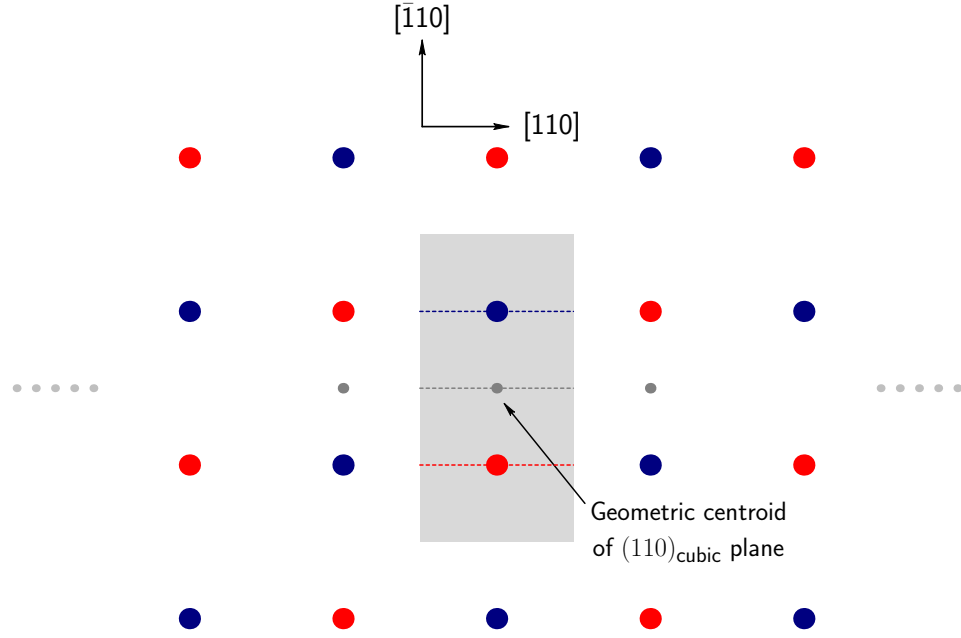


Figure 4.3: A portion of a crystal with B2 symmetry. The gray area indicates an undeformed (B2) planar unit cell.

4.1.1 Two step decomposition of the deformation of a planar unit cell

Recall that Cauchy-Born kinematics relates the final position of atoms within a crystal to the uniform deformation gradient. Specifically, this is done by allowing atoms within a unit cell to move away from their initial positions and then subjecting the unit cell to a uniform deformation. In the M^4 the two step decomposition of a planar unit cell deformation is part of the input data and hence follows the Cauchy-Born kinematics.

Step 1. Interatomic shifts of atoms within the planar unit cell

The interatomic shifts of atoms within the planar unit cell are expressed as fractional vectors and indicate motion along the lattice vector directions. The actual fractional values for B19 and $L1_0$ are obtained from the BFB of AuCd and are detailed in *Guthikonda and Elliott (2011)*. It turns out that for B19 the motion of the atoms results in alternating zig-zag shifts of the geometric centroid in the $[\bar{1}10]_{\text{cubic}}$ direction (see Fig. 4.4). This identifies two types of B19 planar unit cells, denoted here as B^+ and B^- , that are associated with whether the geometric centroid of that plane has moved up or down (in the $[\bar{1}10]_{\text{cubic}}$ direction) from its original position. $L1_0$ is distinct from B19 in that there is no fractional motion of the atoms within the unit cell and hence no in-plane motion of the geometric centroid. For this reason the two $L1_0$ plane types are treated identically at this point. It is only after the uniform

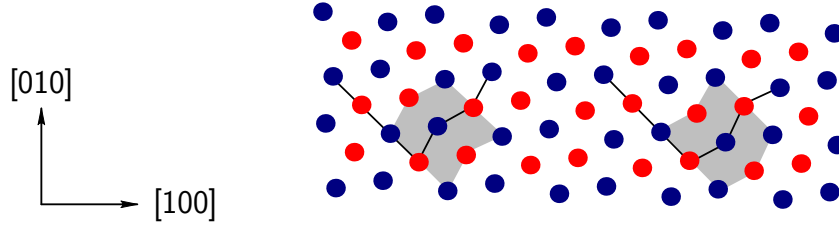


Figure 4.4: A projection of two variants of a B19 crystal. For B19 the motion of the atoms results in alternating zig-zag shifts of the geometric centroid in the $[\bar{1}10]_{\text{cubic}}$ direction.

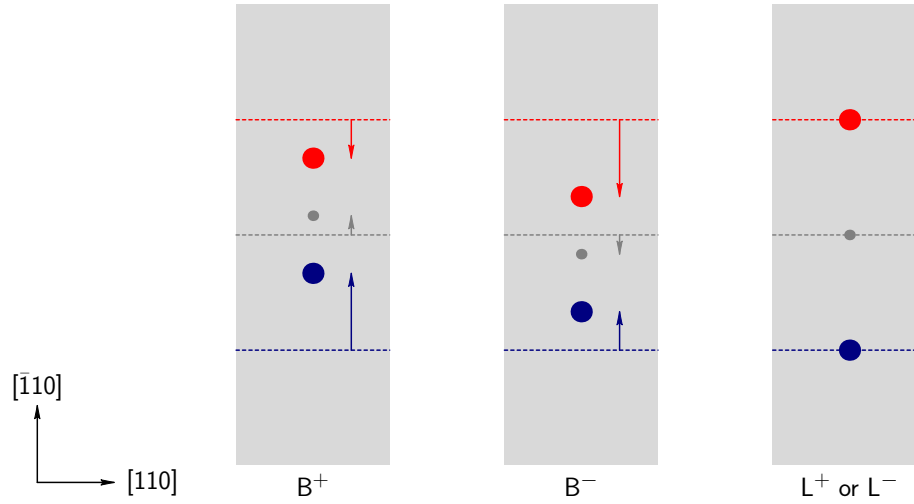


Figure 4.5: The four plane types in the M^4 after an in-plane atomic motion.

deformation of the unit cell that the two L_0 plane types— L^+ and L^- —are identified. In the standard basis \mathcal{B} the geometric centroid shift vectors are given by

$$\mathbf{S}_{B^+} = [-s_{B^+}, s_{B^+}, 0]^\top \quad (4.3)$$

$$\mathbf{S}_{B^-} = [-s_{B^-}, s_{B^-}, 0]^\top, \quad (4.4)$$

$$\mathbf{S}_{L^+} = \mathbf{S}_{L^-} = \mathbf{0}, \quad (4.5)$$

where $s_{B^+} = -s_{B^-}$ is obtained from the BFB study of B19. Fig. 4.5 shows a schematic of the planar unit cells for B^+ , B^- , L^+ , and L^- after the in-plane atomic motion.

Step 2. Uniform deformation of the planar unit cell

From the BFB study of the base martensites detailed in Ch. 3 the deformation of the B19 structure is described through the deformation gradient (given in the standard

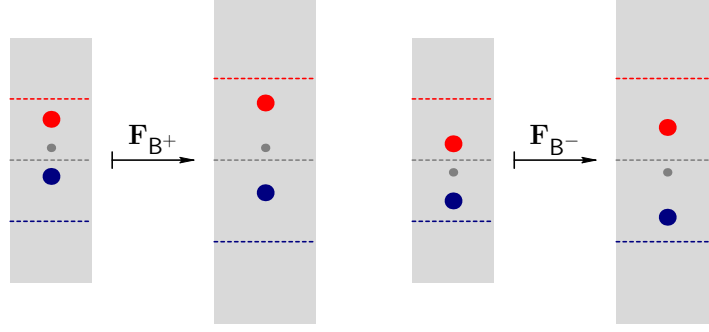


Figure 4.6: The plane types B^+ and B^- in the M^4 after the application of uniform deformations \mathbf{F}_{B^+} and \mathbf{F}_{B^-} . The deformation is given by a stretch of $a + \gamma$, $a - \gamma$, and b in the $[110]_{\text{cubic}}$, $[\bar{1}10]_{\text{cubic}}$, and $[001]_{\text{cubic}}$ directions, respectively.

basis \mathcal{B}) as

$$\mathbf{U}_{B^{19}} = \begin{bmatrix} a & \gamma & 0 \\ \gamma & a & 0 \\ 0 & 0 & b \end{bmatrix}, \quad (4.6)$$

where $a > 1$, $b < 1$, and $\gamma < 0$. The deformation gradients \mathbf{F}_{B^+} and \mathbf{F}_{B^-} that deform the planes B^+ and B^- are both given by $\mathbf{U}_{B^{19}}$. It is useful to view the deformation gradients \mathbf{F}_{B^+} and \mathbf{F}_{B^-} in the basis $\mathcal{B}_{M^4} = (\tilde{\mathbf{g}}_1, \tilde{\mathbf{g}}_2, \tilde{\mathbf{g}}_3)$, where

$$\tilde{\mathbf{g}}_1 = \frac{1}{\sqrt{2}}[1, 1, 0]^\top, \quad \tilde{\mathbf{g}}_2 = \frac{1}{\sqrt{2}}[-1, 1, 0]^\top, \quad \text{and} \quad \tilde{\mathbf{g}}_3 = [0, 0, 1]^\top. \quad (4.7)$$

This basis is a useful setting to describe modulated martensites since they are characterized by shuffles of $(110)_{\text{cubic}}$ planes in the $[\bar{1}10]_{\text{cubic}}$ direction (i.e., planes with normal equal to $\tilde{\mathbf{g}}_1$ move in the $\tilde{\mathbf{g}}_2$ direction). In this basis the deformation gradients \mathbf{F}_{B^+} and \mathbf{F}_{B^-} are described as follows:

$$\mathbf{F}_{B^+} = \mathbf{F}_{B^-} = \begin{bmatrix} a + \gamma & 0 & 0 \\ 0 & a - \gamma & 0 \\ 0 & 0 & b \end{bmatrix}. \quad (4.8)$$

Thus, the deformation of the planar unit cell under \mathbf{F}_{B^+} and \mathbf{F}_{B^-} is a stretch by $a + \gamma$, $a - \gamma$, and b in the $[110]_{\text{cubic}}$, $[\bar{1}10]_{\text{cubic}}$, and $[001]_{\text{cubic}}$ directions, respectively. This deformation is schematically shown in Fig. 4.6. Note that these values for \mathbf{F}_{B^+} and \mathbf{F}_{B^-} satisfy condition (\star) mentioned above. Indeed, (1) is satisfied since $[\bar{1}10]_{\text{cubic}}$ and $[001]_{\text{cubic}}$ are principal directions and (2) is satisfied since $\mathbf{F}_{B^+} - \mathbf{F}_{B^-} = \mathbf{0}$.

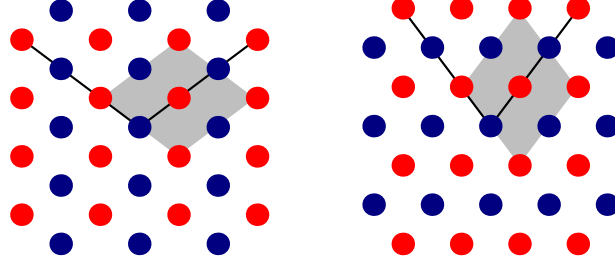


Figure 4.7: A projection of two variants of $L1_0$ crystals.

To obtain the deformation gradients for L^+ and L^- we start with the deformation gradients for the variants of $L1_0$ obtained from the BFB study. In the standard basis \mathcal{B} they are given by

$$\mathbf{U}_{L1_0^+} = \begin{bmatrix} c & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & d \end{bmatrix}, \quad \mathbf{U}_{L1_0^-} = \begin{bmatrix} d & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & d \end{bmatrix}, \quad (4.9)$$

where $c > 1$ and $d < 1$. These deformations produce the deformed crystals schematically shown in Fig. 4.7.

Unlike for the case of B^+ and B^- the deformation gradients for L^+ and L^- cannot simply be taken as those given by $\mathbf{U}_{L1_0^+}$ and $\mathbf{U}_{L1_0^-}$. This is because these two deformation gradients do not satisfy condition (\star) . In particular, the deformation gradients $\mathbf{U}_{L1_0^+}$ and $\mathbf{U}_{L1_0^-}$ do not preserve the orientation of $(110)_{\text{cubic}}$ planes and hence they cannot be used to describe modulated martensites. To illustrate this, consider the deformation gradients $\mathbf{U}_{L1_0^+}$ and $\mathbf{U}_{L1_0^-}$ in the \mathcal{B}_{M^4} basis given by

$$\mathbf{U}_{L1_0^+} = \frac{1}{2} \begin{bmatrix} \ell_1 & \ell_2 & 0 \\ \ell_2 & \ell_1 & 0 \\ 0 & 0 & 2d \end{bmatrix}, \quad \mathbf{U}_{L1_0^-} = \frac{1}{2} \begin{bmatrix} \ell_1 & -\ell_2 & 0 \\ -\ell_2 & \ell_1 & 0 \\ 0 & 0 & 2d \end{bmatrix}, \quad (4.10)$$

where $\ell_1 = c + d$ and $\ell_2 = c - d$. To satisfy (1) of condition (\star) we seek rotation matrices \mathbf{R}^+ and \mathbf{R}^- such that the $\tilde{\mathbf{g}}_1 \otimes \tilde{\mathbf{g}}_2$ and $\tilde{\mathbf{g}}_3 \otimes \tilde{\mathbf{g}}_2$ components of $\mathbf{R}^\pm \mathbf{F}_{L1_0^\pm}$ are equal to zero. It can be shown that the rotation matrices are given by

$$\mathbf{R}^\pm = \begin{bmatrix} \cos(\pm\theta) & -\sin(\pm\theta) & 0 \\ \sin(\pm\theta) & \cos(\pm\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (4.11)$$

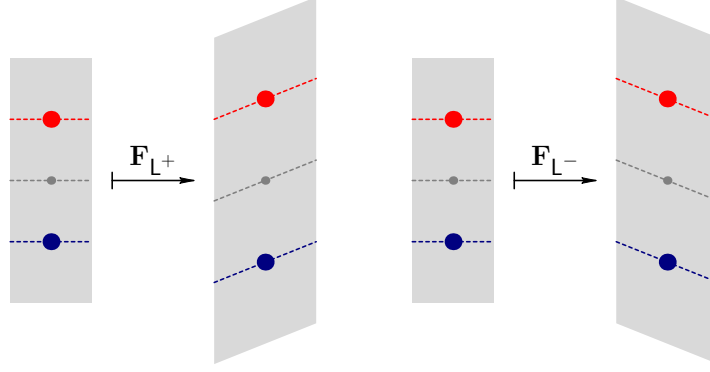


Figure 4.8: The plane types L^+ and L^- in the M^4 after the application of uniform deformations \mathbf{F}_{L^+} and \mathbf{F}_{L^-} .

where $\tan(\theta) = \frac{\ell_1}{\ell_2}$. In the \mathcal{B}_{M^4} basis the final deformation gradients are given by

$$\mathbf{F}_{L^+} = \mathbf{R}^+ \mathbf{U}_{L1_0^+} = \begin{bmatrix} L_1 & 0 & 0 \\ \zeta & L_2 & 0 \\ 0 & 0 & d \end{bmatrix}, \quad (4.12)$$

$$\mathbf{F}_{L^-} = \mathbf{R}^- \mathbf{U}_{L1_0^-} = \begin{bmatrix} L_1 & 0 & 0 \\ -\zeta & L_2 & 0 \\ 0 & 0 & d \end{bmatrix}. \quad (4.13)$$

The above definition of \mathbf{F}_{L^+} and \mathbf{F}_{L^-} forms a twin system with normal given by $\mathbf{n} = [1, 0, 0]^\top$ (in the \mathcal{B}_{M^4} basis). That is, $\mathbf{F}_{L^+} - \mathbf{F}_{L^-} = [0, 2\zeta, 0]^\top \otimes [1, 0, 0]^\top$, which implies that part two of condition (\star) is satisfied. The effect of these deformation gradients on the planar unit cell is schematically given in Fig. 4.8.

We would like to emphasize that *the actual deformation gradients of each planar unit cell need to be further altered in order to satisfy both parts of condition (\star) for all pairs of plane types*. For example, it can be shown that $\mathbf{F}_{B^-} - \mathbf{F}_{L^-} \neq \mathbf{a} \otimes \mathbf{n}$ for any $\mathbf{a}, \mathbf{n} \in \mathbb{R}^3$. In order to reconcile this problem the M^4 adjusts the deformation of each planar unit cell in a given modulated martensite such that the in-plane transformations of all $(110)_{\text{cubic}}$ planes are described by the same stretches. This planar stretch is given by the deformation gradient obtained from the weighted average of plane types in a given M^4 sequence (and hence varies on a case to case basis). The details of this last step of the deformation of the planar unit cells is demonstrated in a later section via an explicit example. Finally, the four deformed plane types are shown in Fig. 4.9. It is worth noting that these four plane type definitions allow the M^4 to exactly describe the two base martensites. In particular, B19 is

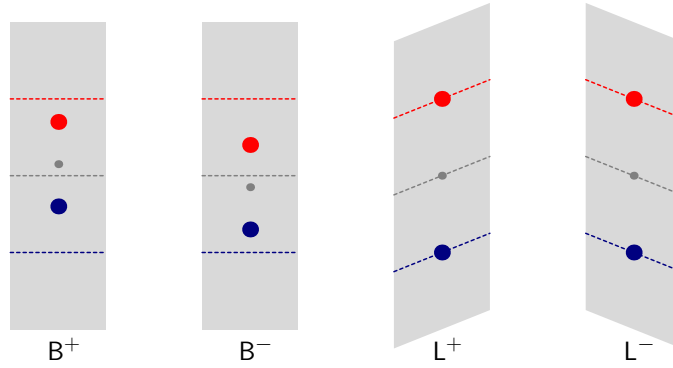


Figure 4.9: The four plane types in the M^4 .

Plane type	Followed by	Shown in Fig. 4.10
B^+	L^- or B^-	(a) (b)
B^-	L^+ or B^+	(c) (d)
L^+	L^+ or B^+	(e) (f)
L^-	L^- or B^-	(g) (h)

Table 4.1: The M^4 mixture rules.

described as a sequence of alternating B^+ and B^- planes. Similarly, $L1_0$ is described as a constant sequence of either L^+ or L^- planes.

4.1.2 The M^4 mixture rules

The M^4 also provides a set of rules defining the types of planes that can precede and follow the planar unit cells of each of the four plane types in a modulated martensite sequence. The rules ensure that each modulated structure is a mixture of the four plane types as defined by their nearest-neighbor planes. Moreover, it is found that the mixture rules increase the probability that the predicted modulated martensite will be dynamically stable. Tbl. 4.1 and Fig. 4.10 summarize the M^4 mixture rules. Note that while the table provides the types of planes that can follow a given plane type, it also implicitly prescribes the rules for what can precede a given plane. For example, combining the results from the second and third lines with the first line we find that B^+ can be followed by L^- or B^- and can be preceded by B^- or L^+ .

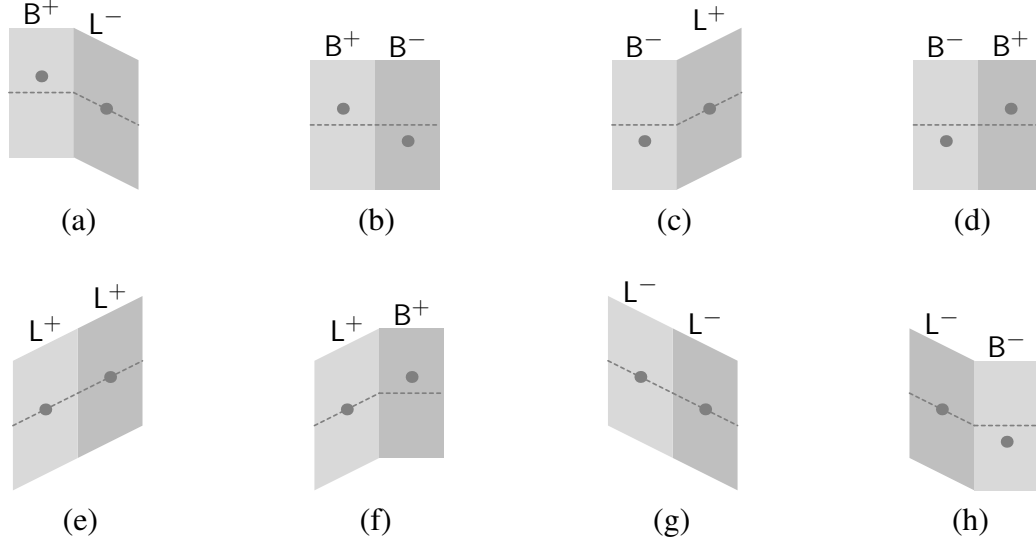


Figure 4.10: The M^4 mixture rules.

4.1.3 Recipe to construct a modulated martensite according to the M^4

With the M^4 rules defined the procedure to construct/interpret an arbitrary modulated martensite may be defined. The recipe presented here is meant to provide a general understanding of the M^4 . Refer to Appendix E for a more detailed presentation of the technique including an algorithm used to construct a modulated martensite configuration according to the M^4 . Note that this procedure defines a modulated martensite at a given temperature. In the M^4 the temperature dependence of the modulated martensite is captured through the temperature dependence of the input data. The M^4 procedure is as follows:

1. Compute the properties of B2, B19, and L1₀ and identify the following quantities to be used as the M^4 input data:

- Energy of the four plane types \mathcal{E}_{B^+} , \mathcal{E}_{B^-} , \mathcal{E}_{L^+} , and \mathcal{E}_{L^-} given by

$$\mathcal{E}_{B^+} = \mathcal{E}_{B^-} = \mathcal{E}_{B19}, \quad (4.14)$$

$$\mathcal{E}_{L^+} = \mathcal{E}_{L^-} = \mathcal{E}_{L1_0}, \quad (4.15)$$

where \mathcal{E}_{B19} and \mathcal{E}_{L1_0} are the energies obtained from the BFB study for the B19 and L1₀ structures, respectively.

- In-plane shifts of the geometric centroid of the $(110)_{\text{cubic}}$ planes for all plane types: \mathbf{S}_{B^+} , \mathbf{S}_{B^-} , \mathbf{S}_{L^+} , and \mathbf{S}_{L^-} according to Eqs. (4.3)–(4.5).
- Uniform deformation gradients \mathbf{F}_{B^+} , \mathbf{F}_{B^-} , \mathbf{F}_{L^+} , and \mathbf{F}_{L^-} from Eqs. (4.6), (4.12),

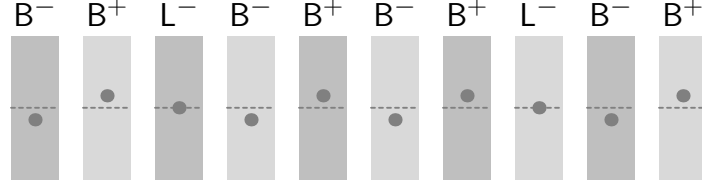


Figure 4.11: A schematic of the intermediate planar unit cells for the modulated martensite described by $(B^-, B^+, L^-, B^-, B^+, B^-, B^+, L^-, B^-, B^+)$.

and (4.13).

2. Determine the sequence of plane types in the unit cell of the modulated martensite along the $[110]_{\text{cubic}}$ direction. For example, suppose the unit cell of the modulated martensite can be described by the following sequence of plane types: $\mathcal{P} = (B^-, B^+, L^-, B^-, B^+, B^-, B^+, L^-, B^-, B^+)$ (as shown in the next section, this is actually the modulated martensite k_5).
3. For each plane in the M^4 sequence apply the appropriate in-plane atomic shifts (step 1 in the two step decomposition of the planar unit cell deformation) according to the plane type. Fig. 4.11 provides a schematic of the intermediate planar unit cells (i.e., after the application of step 1) in the current example. Note that while only the geometric centroid is shown within each planar unit cell there is also an associated atomic motion (that is part of the input data obtained from the BFB result for B19 and L1₀).
4. To complete the transformation of the planar unit cells in-plane deformations must first be applied to ensure that each plane is commensurate to the others. This is to ensure that ‘attaching’ the planar unit cells would not result in a discontinuous deformation where the planes meet. For example, a straightforward application of the uniform deformations \mathbf{F}_{B^+} , \mathbf{F}_{B^-} , \mathbf{F}_{L^+} , and \mathbf{F}_{L^-} in the current example would result in the deformed planar unit cells shown in Fig. 4.12. Note the mismatch of the in-plane height between the B19 and L1₀ planar unit cells. Therefore, the planar unit cells must be deformed to form coherent $(110)_{\text{cubic}}$ planes satisfying both parts of condition (\star) . In order to reconcile this problem first note that every \mathbf{F}_p can be written in the following form (in the \mathcal{B}_{M^4} basis)

$$\mathbf{F}_p = \begin{bmatrix} a & 0 & 0 \\ \eta & b & 0 \\ 0 & 0 & c \end{bmatrix}, \quad (4.16)$$

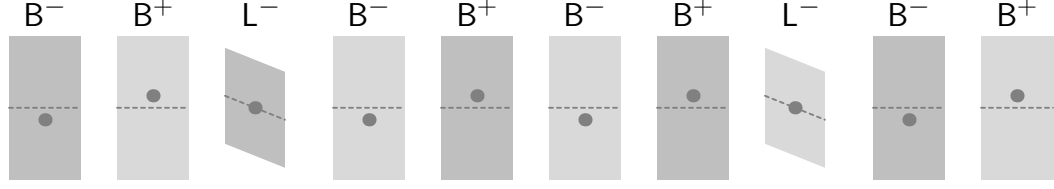


Figure 4.12: The transformed planar unit cells for the modulated martensite described by $(B^-, B^+, L^-, B^-, B^+, B^-, B^+, L^-, B^-, B^+)$. Note that there is a discrepancy in the in-plane width and height between the $B19$ and $L1_0$ planar unit cells. ‘Attaching’ the planar unit cells at this point would result in a discontinuity where the planes meet.

where $\eta = 0$ for \mathbf{F}_{B^+} and \mathbf{F}_{B^-} , $\eta > 0$ for \mathbf{F}_{L^+} , and $\eta < 0$ for \mathbf{F}_{L^-} . Hence, for the deformation gradients of the form given in Eq. (4.16) it is ensured that $\tilde{\mathbf{g}}_2$ and $\tilde{\mathbf{g}}_3$ are principal directions for every \mathbf{F}_p . This implies that we can decompose the deformation gradients as a sum of in-plane and out-of-plane deformation gradients as follows:

$$\mathbf{F}_p = \mathbf{F}_p^{\text{in-plane}} + \mathbf{F}_p^{\text{out-of-plane}}, \quad (4.17)$$

where $p \in \mathbb{P} = (B^+, B^-, L^+, L^-)$. The in-plane and out-of-plane deformation gradients are defined as follows:

$$\mathbf{F}_p^{\text{in-plane}} = b (\tilde{\mathbf{g}}_2 \otimes \tilde{\mathbf{g}}_2) + c (\tilde{\mathbf{g}}_3 \otimes \tilde{\mathbf{g}}_3), \quad (4.18)$$

$$\mathbf{F}_p^{\text{out-of-plane}} = a (\tilde{\mathbf{g}}_1 \otimes \tilde{\mathbf{g}}_1) + \eta (\tilde{\mathbf{g}}_2 \otimes \tilde{\mathbf{g}}_1), \quad (4.19)$$

where b and c are components of \mathbf{F}_p as in Eq. (4.16). This decomposition defines $\mathbf{F}_p^{\text{in-plane}}$ such that it transforms $(110)_{\text{cubic}}$ planes purely as an in-plane stretch and thus implies that

$$\mathbf{F}_p^{\text{in-plane}} \tilde{\mathbf{g}}_1 = \mathbf{0}, \quad (4.20)$$

$$\mathbf{F}_p^{\text{in-plane}} \tilde{\mathbf{g}}_2 = b \tilde{\mathbf{g}}_2, \quad (4.21)$$

$$\mathbf{F}_p^{\text{in-plane}} \tilde{\mathbf{g}}_3 = c \tilde{\mathbf{g}}_3. \quad (4.22)$$

We define the in-plane deformation gradient of the planar unit cell as

$$\mathbf{F}_{M^4}^{\text{in-plane}} = \frac{1}{2N} \sum_{p \in \mathcal{P}} \mathbf{F}_p^{\text{in-plane}}, \quad (4.23)$$

where $\mathcal{P} = (p_1, \dots, p_{2N})$ is the M^4 sequence of plane types for a particular modu-

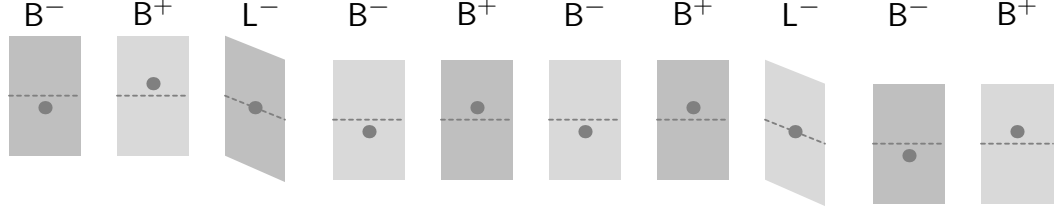


Figure 4.13: The planar unit cells for $(B^-, B^+, L^-, B^-, B^+, B^-, B^+, L^-, B^-, B^+)$ after they have been fully deformed.

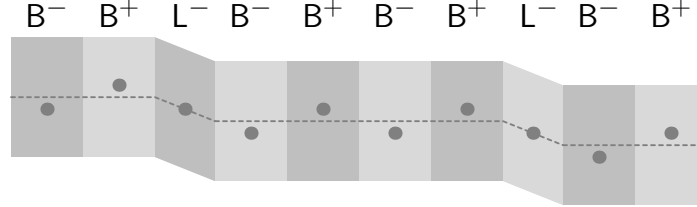


Figure 4.14: The modulated structure $(B^-, B^+, L^-, B^-, B^+, B^-, B^+, L^-, B^-, B^+)$.

lated martensite of interest (such as the one in the ongoing example defined above). The application of this in-plane deformation results in a stretching of the in-plane geometric centroid shift and the relative position vector of the atoms from the geometric centroid by an amount given by the principal value of $\mathbf{F}_{M^4}^{\text{in-plane}}$ in the $\tilde{\mathbf{g}}_2$ direction. The final form for the deformation gradient of the p -th planar unit cell in the M^4 sequence is given by

$$\mathbf{F}_p^{\text{planar-cell}} = \mathbf{F}_{M^4}^{\text{in-plane}} + \mathbf{F}_p^{\text{out-of-plane}}, \quad (4.24)$$

where³ $p \in \mathbb{P}$. The final effect of this for the above example is shown in Fig. 4.13. Note that there is no longer a mismatch between the planes.

5. Finally, the modulated martensite is described by ‘stacking’ the deformed planar unit cells sequentially. Fig. 4.14 shows the modulated martensite for the current example. One consequence of this definition of the unit cell of a modulated martensite is that the M^4 out-of-plane deformation gradient is given by

$$\mathbf{F}_{M^4}^{\text{out-of-plane}} = \frac{1}{2N} \sum_{p \in \mathcal{P}} \mathbf{F}_p^{\text{out-of-plane}}, \quad (4.25)$$

³Note that the final form of the deformation gradient cannot be obtained prior to the identification of the M^4 sequence of plane types that define a particular modulated martensite.

since the reference lattice vector⁴ \mathbf{e}_1 is transformed in the following manner:

$$\begin{aligned}
\mathbf{F}_{p_1}^{\text{planar-cell}} \frac{\mathbf{e}_1}{2N} + \dots + \mathbf{F}_{p_{2N}}^{\text{planar-cell}} \frac{\mathbf{e}_1}{2N} &= \sum_{p \in \mathcal{P}} \left(\mathbf{F}_{M^4}^{\text{in-plane}} + \mathbf{F}_p^{\text{out-of-plane}} \right) \frac{\mathbf{e}_1}{2N} \\
&= \sum_{p \in \mathcal{P}} \mathbf{F}_{M^4}^{\text{in-plane}} \frac{\mathbf{e}_1}{2N} + \sum_{p \in \mathcal{P}} \mathbf{F}_p^{\text{out-of-plane}} \frac{\mathbf{e}_1}{2N} \\
&= \sum_{p \in \mathcal{P}} \mathbf{0} + \sum_{p \in \mathcal{P}} \mathbf{F}_p^{\text{out-of-plane}} \frac{\mathbf{e}_1}{2N} \\
&= \left(\frac{1}{2N} \sum_{p \in \mathcal{P}} \mathbf{F}_p^{\text{out-of-plane}} \right) \mathbf{e}_1 \\
&= \mathbf{F}_{M^4}^{\text{out-of-plane}} \mathbf{e}_1.
\end{aligned} \tag{4.26}$$

It is important to note that in the nano-scale the deformation of the modulated martensite is not smooth but is instead described by piecewise uniform deformations corresponding to deformations of each planar unit cell in the modulated martensite. However, in the micro-scale (a scale where the discreteness of the unit cells is no longer apparent) the M^4 defines the deformation as the weighted average of deformations of the plane types in the modulated martensite. Thus, we have

$$\mathbf{F}_{M^4} = \mathbf{F}_{M^4}^{\text{in-plane}} + \mathbf{F}_{M^4}^{\text{out-of-plane}}. \tag{4.27}$$

Equivalent forms of \mathbf{F}_{M^4} are given by

$$\mathbf{F}_{M^4} = \frac{1}{2N} \sum_{p \in \mathcal{P}} \mathbf{F}_p \tag{4.28}$$

$$= \frac{1}{2N} \sum_{p \in \mathbb{P}} n_p \mathbf{F}_p, \tag{4.29}$$

where n_p is the number of planes of type p in the M^4 sequence \mathcal{P} . Similarly, motivated by the micro-scale description of the uniform deformation, the M^4 defines the energy \mathcal{E}_{M^4} of the modulated martensite as follows:

$$\mathcal{E}_{M^4} = \frac{1}{2N} \sum_{p \in \mathcal{P}} \mathcal{E}_p. \tag{4.30}$$

As a final note, although the M^4 is able to predict the properties of a modulated martensite to a high degree of accuracy, it *will in general exhibit a small but non-zero error*. That is, the

⁴Note that by definition, $\mathbf{e}_1 = \sum_{n=1}^{2N} \frac{\mathbf{e}_1}{2N}$.

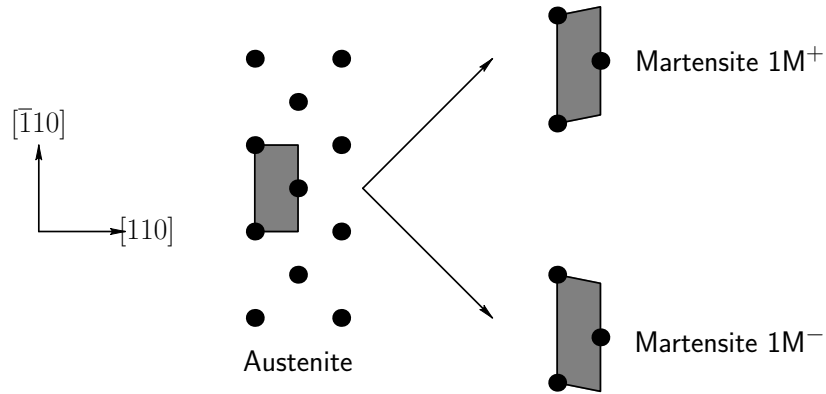


Figure 4.15: The two types of unit cells identified in *Balandraud and Zanzotto (2007)* corresponding to shears in the positive and negative $[\bar{1}10]_{\text{cubic}}$ directions (labeled $1M^+$ and $1M^-$), respectively.

predicted structure will not be an exact equilibrium structure but is a “short distance” away from the actual equilibrium structure. A discussion of the relative errors in the predictions of the M^4 will be presented in a later section.

4.1.4 Two interpretations of the *Balandraud and Zanzotto (2007)* model

Balandraud and Zanzotto (2007) view modulated martensites as periodic stacking of identical close-packed planes. The viewpoint identifies a fundamental planar unit cell composed of adjacent $(110)_{\text{cubic}}$ planes that “slide” relative to one another. Furthermore, *Balandraud and Zanzotto (2007)* describe this sliding as a shear of the $(110)_{\text{cubic}}$ planes in the $[\bar{1}10]_{\text{cubic}}$ direction. Hence, *Balandraud and Zanzotto (2007)* identify two types of unit cells corresponding to shears in the positive and negative $[\bar{1}10]_{\text{cubic}}$ directions. This implies that the only DOFs allowed by the *Balandraud and Zanzotto (2007)* interpretation are those of uniform deformation (i.e., no interatomic shifts). The resulting unit cells are shown in Fig. 4.15 and labeled as $1M^+$ and $1M^-$. The modulated martensite is then described by the sequence of $1M^+$ and $1M^-$ cells. The reason we discuss (below) two interpretations of the *Balandraud and Zanzotto (2007)* model is that we have identified two possible martensites from which to construct the *Balandraud and Zanzotto (2007)* planar unit cell. As such, it is unclear which of the two martensites should be used to define the fundamental planar unit cells.

1. $B\&Z_1$

The unit cells in the $B\&Z_1$ model are related to the equilibrium B19 structure. Notice that the structure described by an alternate stacking of $1M^+$ and $1M^-$ planes is

similar to the nearly close-packed B19 configuration. In this manner, the definition of $1M^+$ and $1M^-$ in $B\&Z_1$ is similar to the two B19 plane types identified as B^+ and B^- by the M^4 . However unlike in the M^4 , the $1M^+$ and $1M^-$ unit cells must only be described through a uniform deformation of the B2 lattice mimicking the B19 configuration and *must not include any interatomic shifts*. Therefore, in the \mathcal{B}_{M^4} basis the DOFs of the planar unit cells in $B\&Z_1$ are given by⁵

$$\mathbf{F}_{1M^+} = \begin{bmatrix} a & 0 & 0 \\ \gamma & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}, \quad \mathbf{F}_{1M^-} = \begin{bmatrix} a & 0 & 0 \\ -\gamma & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}, \quad (4.31)$$

where λ_2 and λ_3 are the principal values of \mathbf{U}_{B19} in the $[\bar{1}10]_{\text{cubic}}$ and $[001]_{\text{cubic}}$ directions, respectively. a and γ are obtained such that the deformation of the B2 lattice (consisting of alternating $1M^+$ and $1M^-$ planes) mimics the B19 configuration. We emphasize that the structure described in this manner is not an exact B19 configuration (since it does not contain the necessary interatomic shift DOFs) and hence, is not in equilibrium.

2. $B\&Z_2$

The unit cells in the $B\&Z_2$ model are related to the equilibrium $L1_0$ structure. As such, the DOFs for the $B\&Z_2$ model are taken to be the following:

$$\mathbf{F}_{1M^+} = \mathbf{F}_{L^+}, \quad \mathbf{F}_{1M^-} = \mathbf{F}_{L^-}, \quad (4.32)$$

where \mathbf{F}_{L^+} and \mathbf{F}_{L^-} are as defined in the M^4 .

4.2 The modulated martensites k_5 , k_7 , and k_9 as interpreted by the M^4 and *Balandraud and Zanzotto (2007)*

This section presents how the M^4 , $B\&Z_1$, and $B\&Z_2$ models can be used to predict the properties of the modulated martensites found in Ch. 3. The primary step in the interpretation of a known modulated martensite according to the M^4 and *Balandraud and Zanzotto (2007)* is the determination of the sequence of plane types that make up the stacking arrangement of the modulated martensite. This is done by analyzing the essential unit cell of the modulated martensite in order to determine the shuffling order of $(110)_{\text{cubic}}$ planes. For the M^4 each plane type in the sequence is identified by its nearest neighbor planes and

⁵It is easy to verify that these deformation gradients satisfy condition (\star) .

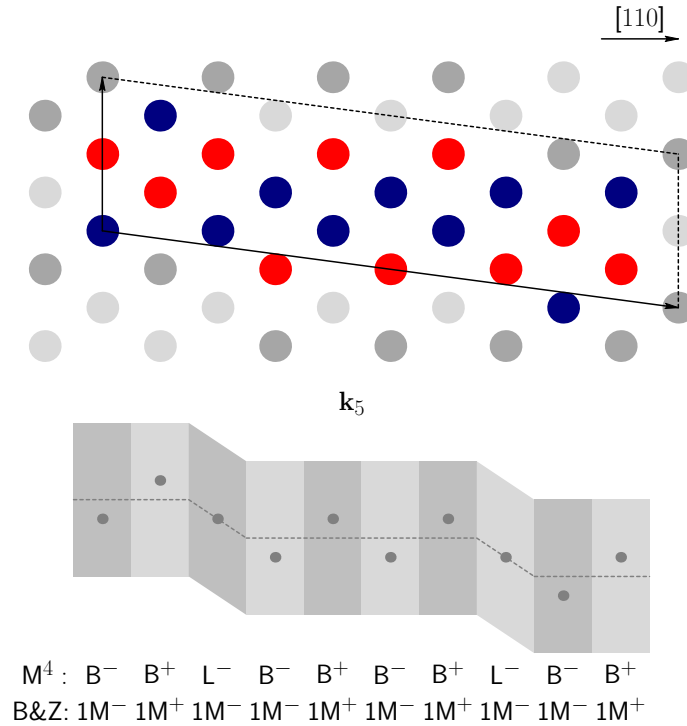


Figure 4.16: The modulated martensite k_5 as interpreted by the M^4 and *Balandraud and Zanzotto* (2007).

hence the B^+ or B^- type planes are characterized by atoms in adjacent planes undergoing alternating (zig-zag) interplanar shifts. Similarly, the planes of type L^+ or L^- are characterized by their adjacent planes undergoing unidirectional shifts associated with the uniform shear of the planar unit cells. For the *Balandraud and Zanzotto* (2007) models $1M^+$ and $1M^-$ are characterized by a shear of the fundamental unit cells in the positive and negative $[\bar{1}10]_{\text{cubic}}$ directions, respectively.

Here, we have used the modulated martensites k_5 , k_7 , and k_9 to demonstrate the interpretation of a modulated martensite according to the M^4 , $B\&Z_1$, and $B\&Z_2$ models. Schematics of the crystals are shown in Figs. 4.16–4.18 along with their corresponding plane type sequence as interpreted by the M^4 . Also in the picture is the interpretation of the modulated martensite according *Balandraud and Zanzotto* (2007) which is labelled B&Z. Note that the B&Z sequence in Figs. 4.16–4.18 applies for both $B\&Z_1$ and $B\&Z_2$. Shown in the figures are the lattice vectors (depicted as arrows) that define the unit cell of the modulated martensite. Additionally, the colored circles indicate atoms in the $(110)_{\text{cubic}}$ planar unit cells used to define each plane type according to its nearest neighbors.

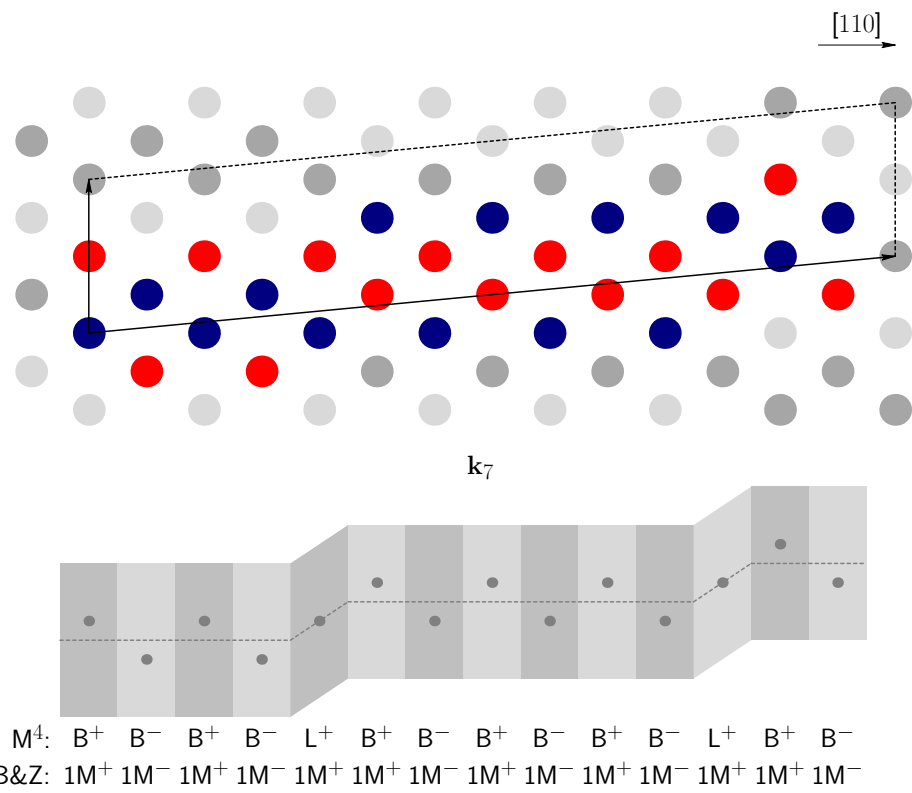


Figure 4.17: The modulated martensite k_7 as interpreted by the M^4 and *Balandraud and Zanzotto* (2007).

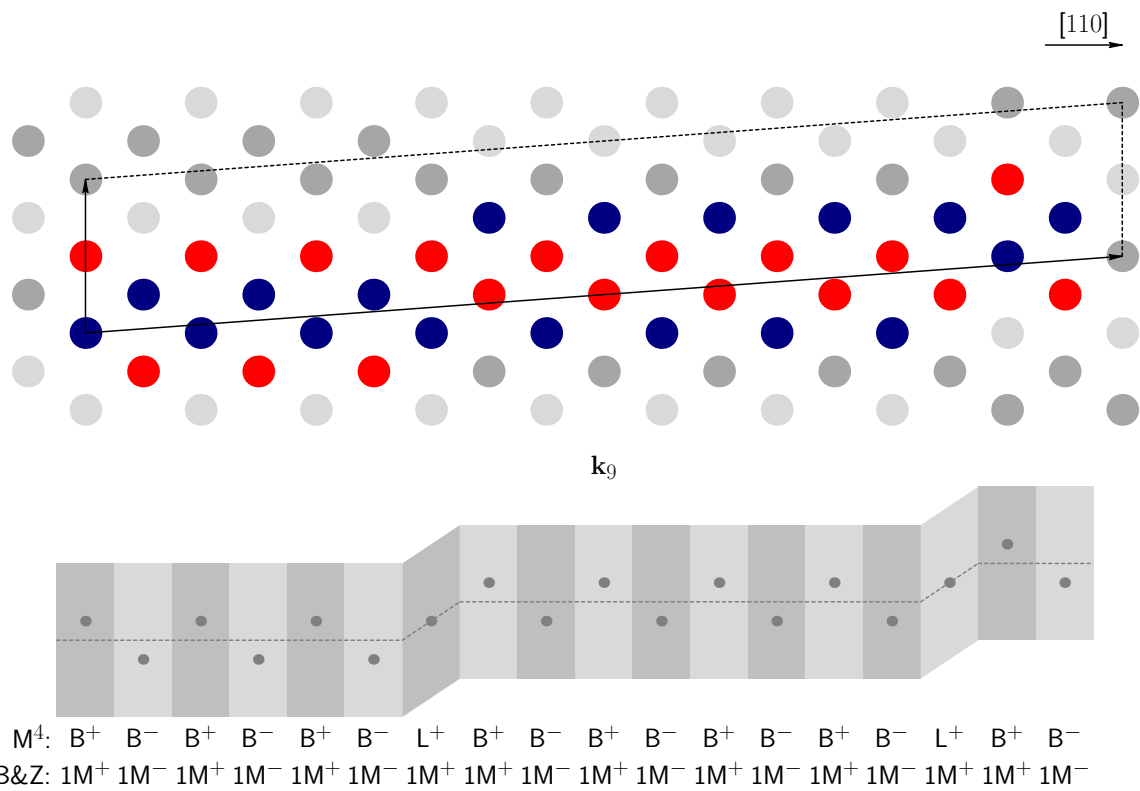


Figure 4.18: The modulated martensite k_9 as interpreted by the M^4 and *Balandraud and Zanzotto* (2007).

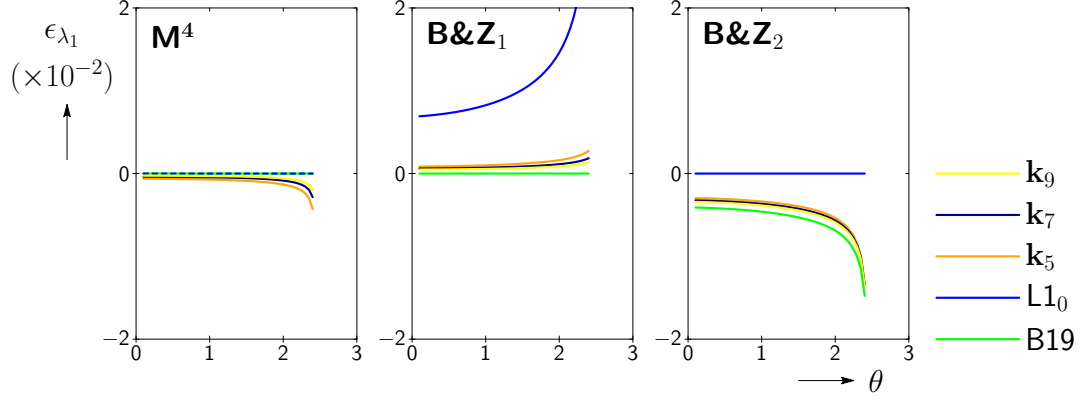


Figure 4.19: The relative error for λ_1 (the smallest eigenvalue of \mathbf{U}) for the structures B19, L1₀, \mathbf{k}_5 , \mathbf{k}_7 , and \mathbf{k}_9 using the M^4 , B&Z₁, and B&Z₂ models.

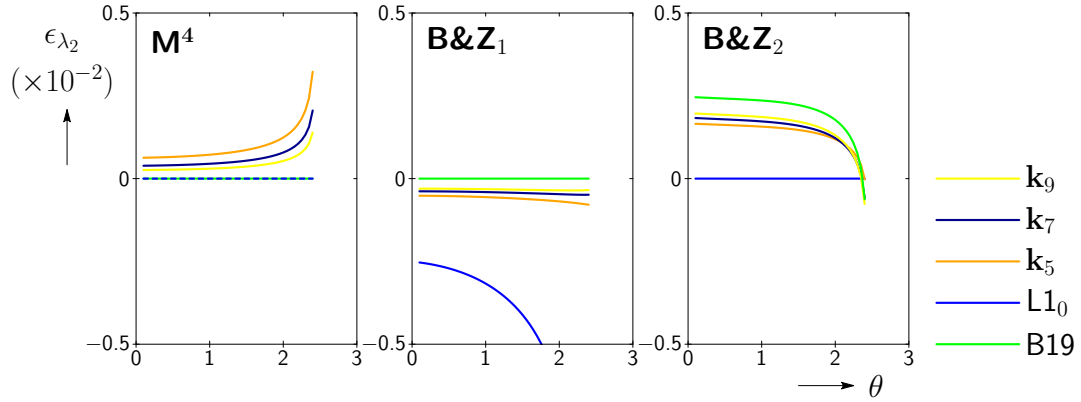


Figure 4.20: The relative error for λ_2 (the middle eigenvalue of \mathbf{U}) for the structures B19, L1₀, \mathbf{k}_5 , \mathbf{k}_7 , and \mathbf{k}_9 using the M^4 , B&Z₁, and B&Z₂ models.

4.3 Relative errors of the M^4 predictions

The predictive performance of the M^4 , B&Z₁, and B&Z₂ models is evaluated by computing the difference between the predicted and actual properties of the modulated martensites \mathbf{k}_5 , \mathbf{k}_7 , and \mathbf{k}_9 . The actual properties for the modulated martensites are obtained via the BFB study detailed in Ch. 3. The models were evaluated by computing the difference in the predicted and actual values for the principal stretches of the deformation gradient ($\lambda_1, \lambda_2, \lambda_3$) and the interatomic fractional shift DOFs ($\mathbf{S} = [S_0^1, S_0^2, S_0^3, \dots, S_{N-1}^1, S_{N-1}^2, S_{N-1}^3]$, where $S_i^j \mathbf{e}_j$, $j = 1, 2, 3$, is the shift vector of the i -th atom). The structures used to evaluate the predicted relative errors of the models are B19, L1₀, \mathbf{k}_5 , \mathbf{k}_7 , and \mathbf{k}_9 . Figs. 4.19–4.23 show the results of the relative errors for the different models (where the colors green, blue, orange, dark blue, and yellow correspond to B19, L1₀, \mathbf{k}_5 , \mathbf{k}_7 , and \mathbf{k}_9 , respectively). Here the

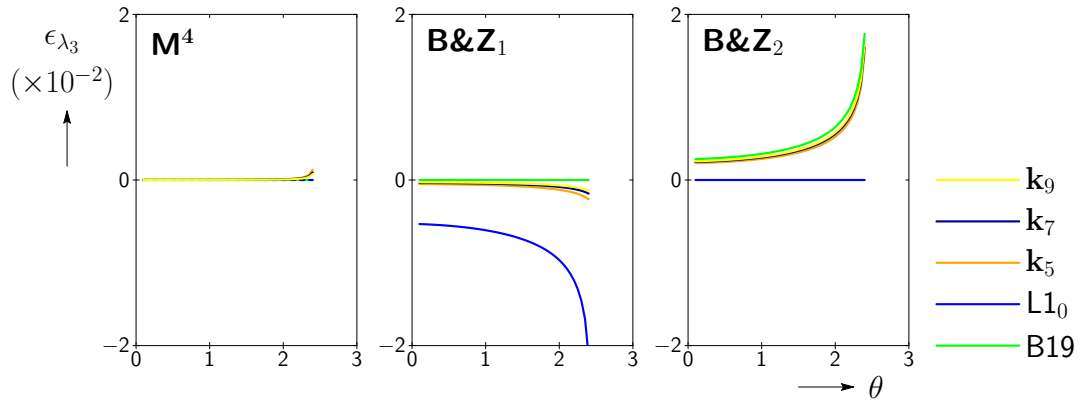


Figure 4.21: The relative error for λ_3 (the largest eigenvalue of U) for the structures $B19$, $L1_0$, k_5 , k_7 , and k_9 using the M^4 , $B\&Z_1$, and $B\&Z_2$ models.

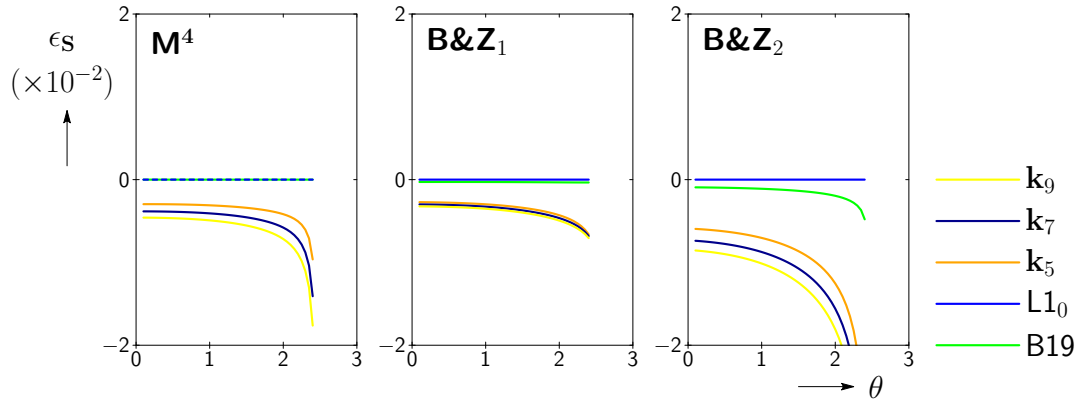


Figure 4.22: The relative error for S (the interatomic shifts DOFs) for the structures $B19$, $L1_0$, k_5 , k_7 , and k_9 using the M^4 , $B\&Z_1$, and $B\&Z_2$ models.

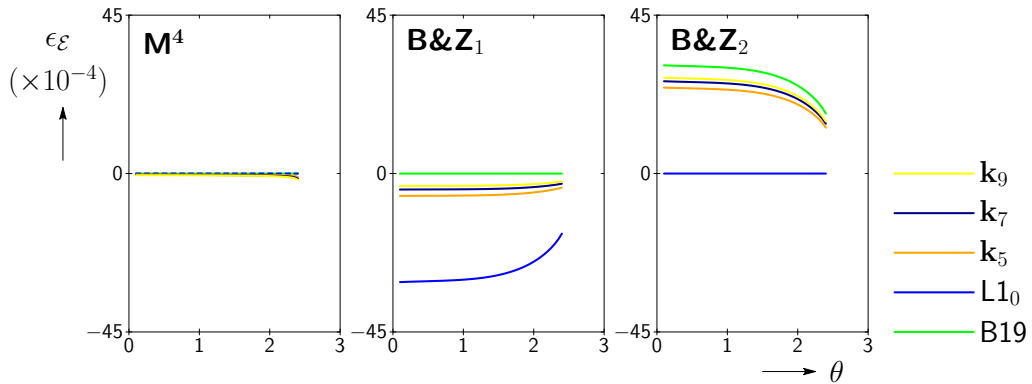


Figure 4.23: The relative error for \mathcal{E} (the energy) for the structures $B19$, $L1_0$, k_5 , k_7 , and k_9 using the M^4 , $B\&Z_1$, and $B\&Z_2$ models.

relative errors are computed as follows:

$$\epsilon_{\lambda_i} = \Delta\lambda_i/\lambda_i^{\text{actual}} = \left(\lambda_i^{\text{predicted}} - \lambda_i^{\text{actual}} \right) / \lambda_i^{\text{actual}}, \quad i = 1, 2, 3, \quad (4.33)$$

$$\epsilon_{\mathbf{S}} = \|\mathbf{S}^{\text{predicted}} - \mathbf{S}^{\text{actual}}\| / \|\mathbf{S}^{\text{actual}}\|, \quad (4.34)$$

$$\epsilon_{\mathcal{E}} = \left(\mathcal{E}^{\text{predicted}} - \mathcal{E}^{\text{actual}} \right) / \mathcal{E}^{\text{actual}}. \quad (4.35)$$

Only stable equilibria were used to compute the prediction relative errors. This is because the M^4 was not developed to predict unstable configurations. The performance of the M^4 model in predicting energy was also compared to the performance of the *Balandraud and Zanzotto* (2007) models $B\&Z_1$ and $B\&Z_2$. It is important to note that *Balandraud and Zanzotto* (2007) do not comment on what the energies of a modulated martensite are. Therefore, to compare the performance of the M^4 model we extended the interpretation made by *Balandraud and Zanzotto* (2007) and defined the energies of a modulated martensite as the energy of the respective structures in the definition of the $B\&Z_1$ and $B\&Z_2$ models. Hence, the definition of the energies $\mathcal{E}_{B\&Z_1}$ and $\mathcal{E}_{B\&Z_2}$ are

$$\mathcal{E}_{B\&Z_1} = \mathcal{E}_{B19}, \quad (4.36)$$

$$\mathcal{E}_{B\&Z_2} = \mathcal{E}_{L1_0}. \quad (4.37)$$

Note that the equilibrium energy of B19 was used to define the energy of the modulated martensite according to⁶ $B\&Z_1$. Fig. 4.23 shows a plot of the relative error in energy for the three models.

4.3.1 Discussion on the M^4 , $B\&Z_1$, and $B\&Z_2$ prediction relative errors

In general, all the models exhibit configuration relative errors that are within 1%. Moreover, it was found that the $B\&Z_2$ model yields predictions with the largest relative error in all cases. Also notice that the relative errors increase near the turning point (at high temperature) of B19 and $L1_0$ regardless of the model. This suggests that the models should not be expected to perform well near a loss of dynamical stability of the base martensite(s). Furthermore, *unlike $B\&Z_1$ and $B\&Z_2$, the M^4 perfectly predicts the properties of B19 and $L1_0$* . This is because B19 and $L1_0$ are the base martensites in the M^4 .

⁶Strictly speaking, the energy of the fundamental unit cell $B\&Z_1$ would be that of the energy of the sheared configuration used in the definition of the $B\&Z_1$ model. Instead, we use the definition in Eq. (4.37) because the view in this work is that all predictions from the theory are built from information related to just one of the base martensite phases (B19 or $L1_0$), and hence, information regarding the non-equilibrium sheared configuration was not used.

In the prediction of the configuration for \mathbf{k}_5 , \mathbf{k}_7 , and \mathbf{k}_9 we see from Figs. 4.19–4.21 that the M^4 performs best in predicting the principal stretches λ_1 and λ_3 . Moreover, it was found that the M^4 performs comparably to B&Z₁ in predicting λ_2 . Also notice that unlike the B&Z₂ model, the M^4 and B&Z₁ models show decreasing relative error with increasing modulation. This may be due to long range effects that come into play for shorter modulations. At this point it is unclear why the B&Z₂ model does not display this trend. For the prediction of the interatomic shift DOFs of \mathbf{k}_5 , \mathbf{k}_7 , and \mathbf{k}_9 it was found that the B&Z₁ model performs best, as can be seen in Fig. 4.22. However, as was the case with the relative error in predicting λ_2 , the M^4 performs comparably to the B&Z₁ model.

Finally, we emphasize that the primary strength of the M^4 is its ability to accurately predict the energy of an arbitrary stable modulated martensite. This is demonstrated in Fig. 4.23 where it is seen that the *Balandraud and Zanzotto (2007)* models have energy relative errors that are *one order of magnitude* larger than those of the M^4 .

4.4 Using the M^4 to search for the most kinematically compatible modulated martensite

The M^4 can be used to predict the properties of an arbitrary modulated martensite. One interesting way to test this capability is to use the M^4 to find a highly kinematically compatible modulated structure. This was done by searching within a set of possible phase fraction combinations $\left\{ \left(\frac{n_{B^+}}{2N}, \frac{n_{B^-}}{2N}, \frac{n_{L^+}}{2N}, \frac{n_{L^-}}{2N} \right) \mid \sum_{p \in \mathbb{P}} n_p = 2N \right\}$, where n_p is the number of planes of type p , $\mathbb{P} = (B^+, B^-, L^+, L^-)$, and $2N$ gives the total number of planes. At each phase fraction combination the middle eigenvalue of the deformation gradient is computed through Eq. (4.29) and compared to the middle eigenvalue of the B2 structure⁷. In this work, the search was performed for modulations of 10 to 20 (110)_{cubic} planes at $\theta = 1.5$.

The search produced only one modulated structure (referred to as $\mathbf{k}_7^{(2)}$) more kinematically compatible than any of the structures found in the BFB study detailed in Ch. 3. The identified MM has a modulation of 14 (110)_{cubic} planes. The existence of the predicted modulated martensite was verified by computing the corresponding equilibrium configuration with the BFB atomistic code. Tbl. 4.2 provides a list of middle eigenvalue ratios⁸ for B19, \mathbf{k}_5 , \mathbf{k}_7 , \mathbf{k}_9 , and $\mathbf{k}_7^{(2)}$. Note that the value for $\mathbf{k}_7^{(2)}$ in Tbl. 4.2 is the actual middle eigenvalue of the structure obtained from atomistic calculation using the M^4 predicted structure as an input

⁷The closer that the middle eigenvalue of the modulated structure is to that of the B2 structure, the more compatible it is.

⁸Note, in general one cannot simply use the ratios of middle eigenvalues to determine the compatibility between two phases. However, since B2 has $\mathbf{U}_{B2} = a(\theta)\mathbf{I}$, the approach gives the correct value in this case.

Equilibrium structures	λ_2/λ_2^{B2}
B19	1.02435
\mathbf{k}_5	1.00996
\mathbf{k}_7	1.01628
\mathbf{k}_9	1.01926
$\mathbf{k}_7^{(2)}$	0.99953

Table 4.2: The ratio of middle eigenvalues λ_2/λ_2^{B2} for various structures at $\theta = 1.5$. The structure $\mathbf{k}_7^{(2)}$ was found using the M^4 and is the most compatible structure to B2.

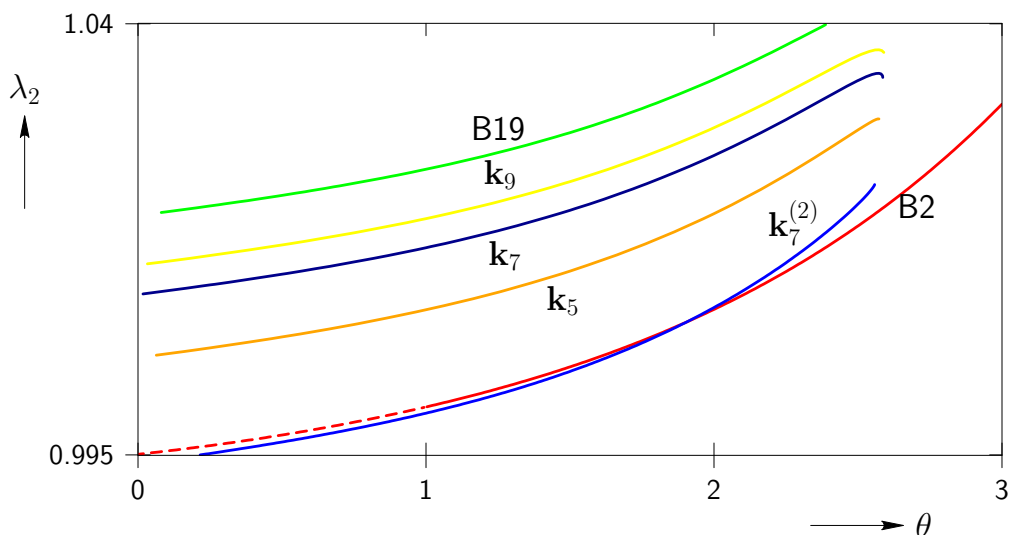


Figure 4.24: Middle eigenvalue of U for $\mathbf{k}_7^{(2)}$ compared to B19 and other modulated structures found.

configuration⁹. We can see from Tbl. 4.2 that the new structure is found to be much more compatible at $\theta = 1.5$ than any of the structures found previously. Fig. 4.24 shows a plot of the middle eigenvalues for $0.0 \leq \theta \leq 3.0$ of B19, \mathbf{k}_5 , \mathbf{k}_7 , \mathbf{k}_9 , and $\mathbf{k}_7^{(2)}$ (shown in light blue). Note that this figure is identical to Fig. 3.12 except for the addition of the $\mathbf{k}_7^{(2)}$ path which was computed with the BFB atomistic code. The red line is again the principal path and so a compatible structure would match that curve perfectly. Observe from Fig. 4.24 that $\mathbf{k}_7^{(2)}$ is much more compatible than B19. Moreover, $\mathbf{k}_7^{(2)}$ remains compatible over a wide range of temperature. It is unclear why $\mathbf{k}_7^{(2)}$ is the only structure found in the search that is more compatible than the other structures obtained through the BFB study.

The unit cell of the modulated structure $\mathbf{k}_7^{(2)}$ consists of the combination of plane types

⁹The M^4 predicted eigenvalue was 1.00045.

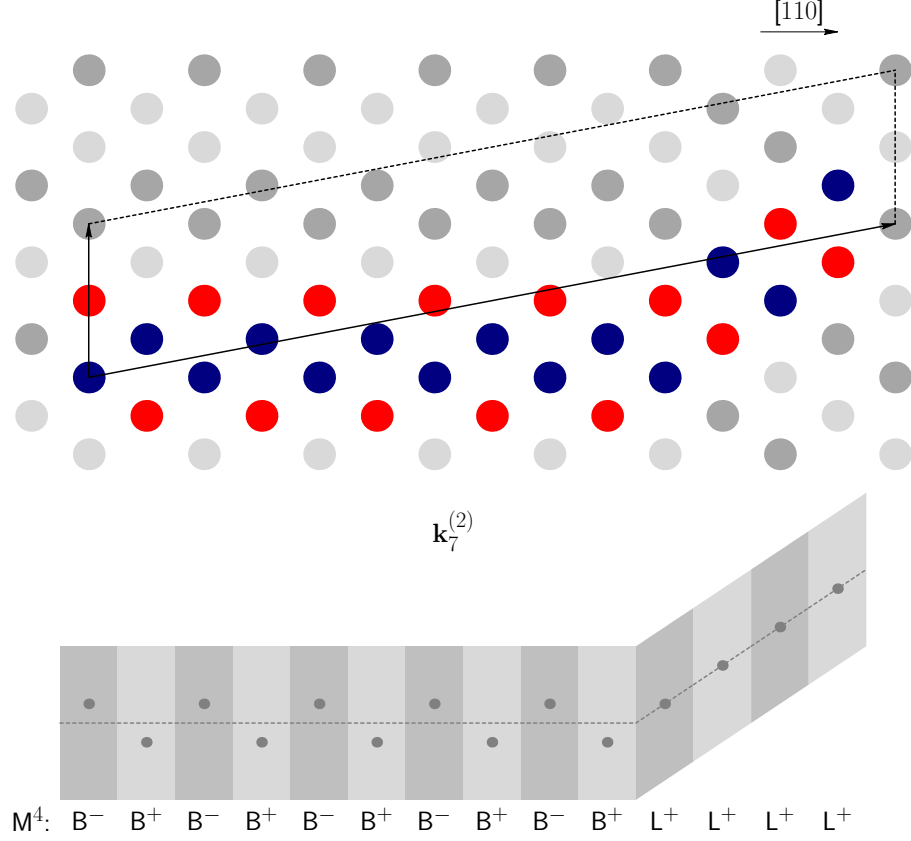


Figure 4.25: The modulated martensite $\mathbf{k}_7^{(2)}$.

given by $(\frac{n_{B^+}}{N}, \frac{n_{B^-}}{N}, \frac{n_{L^+}}{N}, \frac{n_{L^-}}{N}) = (\frac{5}{14}, \frac{5}{14}, \frac{4}{14}, 0)$. Fig. 4.25 provides a schematic of a portion of the $\mathbf{k}_7^{(2)}$ crystal. The configuration shown in Fig. 4.25 is the one that was simulated in this work and the one used to obtain the data in Fig. 4.24. In the figure the simulated cell is outlined by the lattice vectors (shown as arrows) and dashed lines. The M^4 sequence and the planar unit cells are also shown in the figure. It is important to note that the configuration of the unit cell shown below is just one from the possible set of configurations that satisfies the above phase-fraction combination. For instance, the M^4 sequence of $\mathbf{k}_7^{(3)} = (B^-, B^+, L^+, B^-, B^+, L^+, B^-, B^+, L^+, B^-, B^+, L^+, B^-, B^+)$ would also satisfy the phase-fraction condition and be a valid modulated martensite according to the M^4 . Although this configuration is distinct from that of $\mathbf{k}_7^{(2)}$, it was found to exhibit similar deformation and energy¹⁰ suggesting that the phase-fraction combination is the dominant degree of freedom in determining the properties of a modulated martensite. This agrees with the M^4 formulation where the properties of the modulated martensite are obtained through the rule of mixtures.

¹⁰For example, at $\theta = 1.5$, $\lambda_2^{\mathbf{k}_7^{(3)}} / \lambda_2^{B^2} = 0.999172$ and $\Delta \mathcal{E}^{\mathbf{k}_7^{(3)}} = 9.45023 \times 10^{-5}$ eV/Å compared to $\lambda_2^{\mathbf{k}_7^{(2)}} / \lambda_2^{B^2} = 0.999529$ and $\Delta \mathcal{E}^{\mathbf{k}_7^{(2)}} = 9.28681 \times 10^{-5}$ eV/Å.

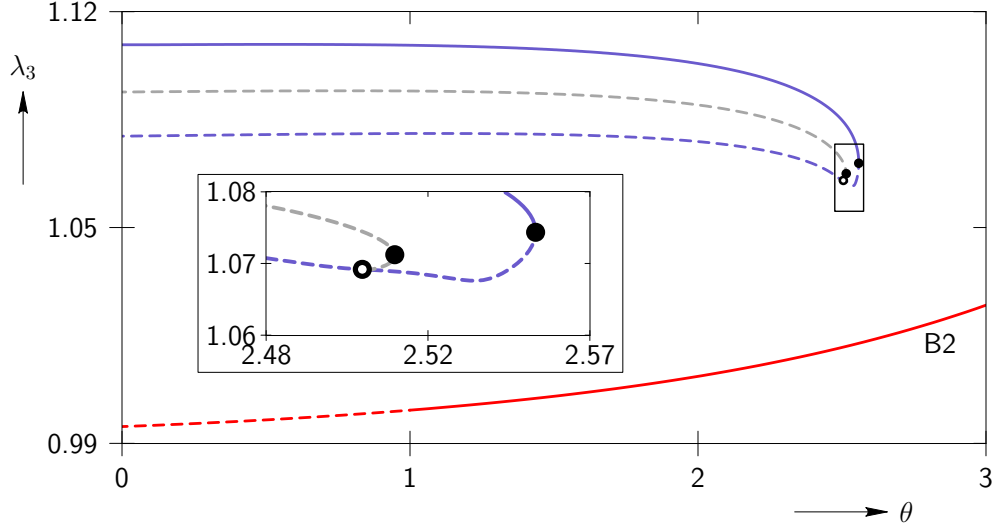


Figure 4.26: $\mathbf{k}_7^{(2)}$ path is shown to be an isolated path that does not connect to the principal path via the branching of paths.

It is important to emphasize that the $\mathbf{k}_7^{(2)}$ path was not found through the branching of paths via a BFB study but was discovered using the M^4 . As such, there is no guarantee that this path will connect to the principal path directly or through any of its primary (or secondary, tertiary, and so on) bifurcating paths. In fact, $\mathbf{k}_7^{(2)}$ was found to be an isolated path as can be seen in Fig. 4.26. The $\mathbf{k}_7^{(2)}$ path contains a stable segment for $\theta \in (0, 2.558)$ and a turning point at $\theta = 2.558$. The path then becomes unstable with one unstable mode associated with the turning point for $\theta \in (2.558, 2.507)$. A bifurcation point at $\theta = 2.507$ is found along the path and then remains unstable with two negative eigenvalues for the remainder of the path¹¹. The bifurcating path from $\mathbf{k}_7^{(2)}$ at $\theta = 2.507$ was also computed and was observed to have no stable segment and no other bifurcation points. Therefore, $\mathbf{k}_7^{(2)}$ and its bifurcating path are isolated from the principal path. This raises the question of whether a reversible transition between B2 and $\mathbf{k}_7^{(2)}$ can exist. Thus, we searched for a direct transition pathway between $\mathbf{k}_7^{(2)}$ and any of the previously found structures. Steepest descent calculations were performed starting from the saddle point at $\theta = 2.52$ and $\theta = 2.55$. These temperatures were chosen because they belong to the segment of the path ($\theta \in (2.507, 2.558)$) that contains only one unstable mode allowing for a straightforward steepest descent calculation (the remainder of the path is either stable or contains more than one unstable mode). Additionally, two temperatures were chosen in order to ensure that the results obtained are not specific to one temperature but are generic features of the transition pathways from that segment of $\mathbf{k}_7^{(2)}$.

¹¹No phonon instabilities were found for this path.

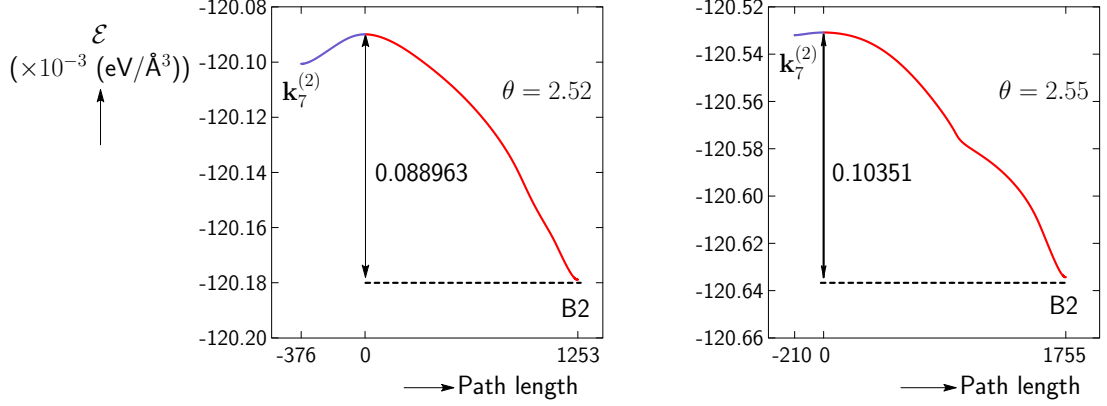


Figure 4.27: Steepest descent calculation showing $B2 \leftrightarrow \mathbf{k}_7^{(2)}$ transition at $\theta = 2.52$ and $\theta = 2.55$.

Figure 4.27 shows the results of the steepest descent calculations performed (two directions at two temperatures). The axis labeled “Path length” is meant to denote a measure of the distance¹² between the two minima on either side of the saddle point. The calculations showed the existence of a direct transition pathway between $B2 \leftrightarrow \mathbf{k}_7^{(2)}$ at both $\theta = 2.52$ and $\theta = 2.55$. The existence of the transition pathway implies that $\mathbf{k}_7^{(2)}$ is in the B2 basin of attraction and indicates that a reversible martensitic phase transformation between these two phases is possible.

4.4.1 Constrained simulation result

A constrained simulation was performed to determine the increase in energy of $\mathbf{k}_7^{(2)}$ due to the austenite kinematic compatibility constraint. The applied constraint was again chosen to be so that the middle eigenvalue of \mathbf{U} for $\mathbf{k}_7^{(2)}$ is equal to that of the B2 at¹³ $\theta = 1.5$. The results are summarized in Fig. 4.28 and Tbl. 4.3. Fig. 4.28 shows a plot of the unconstrained and constrained values of energy for the various modulated structures. From the unconstrained plot we see that $\mathbf{k}_7^{(2)}$ is a low energy phase and is comparable to the other modulated structures found. However, $\mathbf{k}_7^{(2)}$ is energetically more favorable than the remaining structures upon the application of the kinematic compatibility constraint. The values given in Tbl. 4.3 confirm this observation. For this reason and also the fact that $\mathbf{k}_7^{(2)}$ remains compatible for a wide range of temperature, it is reasonable to expect $\mathbf{k}_7^{(2)}$ to be

¹²Specifically, the path length is the sum of all the steepest descent steps, du_i , taken in the DOF space from the saddle point to either minima. The quantity du_i is defined as, $du_i = \|\delta \mathbf{U}, \delta \mathbf{S}\|_i^T - \|\delta \mathbf{U}, \delta \mathbf{S}\|_{i-1}^T$, where $du_0 = 0$ and \mathbf{U} and \mathbf{S} are vectors comprising of components of the uniform deformation gradient and interatomic shifts DOFs, respectively.

¹³The application of the kinematic compatibility is performed in the same way as described in Ch. 3 and detailed in Appendix D.

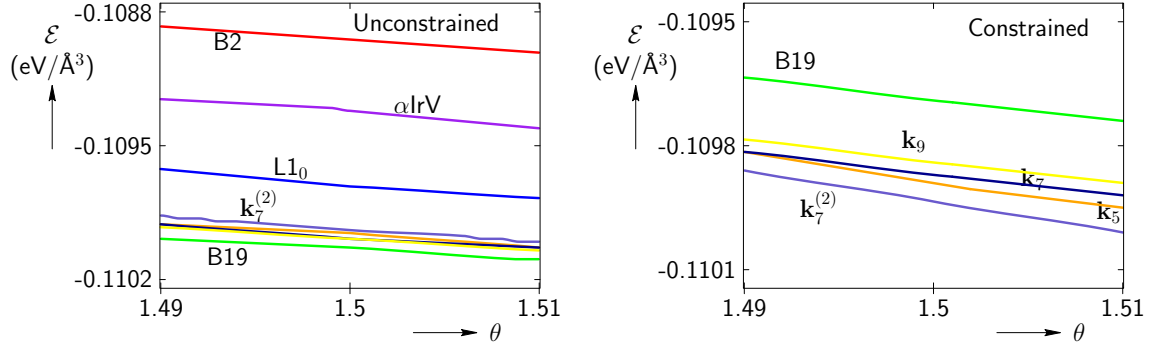


Figure 4.28: Energy for constrained and unconstrained modulated structures. From the unconstrained plot we see that $\mathbf{k}_7^{(2)}$ is a low energy phase and is comparable to the other modulated structures found. However, $\mathbf{k}_7^{(2)}$ is energetically more favorable than the remaining structures upon the application of the kinematic compatibility constraint.

Structure	Unconstrained $\Delta\mathcal{E}$ (eV/Å ³)	Constrained $\Delta\mathcal{E}$ (eV/Å ³)
B19	0.0	3.43073×10^{-4}
\mathbf{k}_5	6.64208×10^{-5}	1.02960×10^{-4}
\mathbf{k}_7	4.75274×10^{-5}	1.52245×10^{-4}
\mathbf{k}_9	3.69999×10^{-5}	1.92275×10^{-4}
$\mathbf{k}_7^{(2)}$	9.28681×10^{-5}	9.29432×10^{-5}

Table 4.3: The energy difference (relative to unconstrained B19) for constrained and unconstrained structures at $\theta = 1.5$. $\mathbf{k}_7^{(2)}$ is energetically more favorable upon the application of a kinematic compatibility constraint.

the experimentally observed modulated martensite.

4.5 The M^4 : Summary and conclusions

The M^4 was developed in view of the dispersion relations and structural analysis of the modulated martensites found in the BFB study detailed in Ch. 3. By interpreting modulated martensites as a mixture of two base martensites the M^4 provides a technique to predict the properties (such as lattice parameters and energy) of arbitrary stable modulated phases. By comparing the relative error of the M^4 to the *Balandraud and Zanzotto (2007)* models, it was found that the M^4 performs comparably to *Balandraud and Zanzotto (2007)* in predicting the configuration of modulated martensites. Additionally, the M^4 was found to have an accuracy that is at least one order of magnitude better than the *Balandraud and Zanzotto (2007)* model in predicting the energy of an arbitrary modulated martensite. Unlike the *Balandraud and Zanzotto (2007)* model the M^4 is able to perfectly predict the B19 properties.

The M^4 also provides the capability to search for new modulated martensites with desired properties. This was demonstrated by performing a brute force search for the most kinematically compatible modulated structure with modulations between 10 to 20 $(110)_{\text{cubic}}$ planes. Only one additional structure was found in the search that was more compatible than the modulated martensites found through the BFB study. This structure consists of fourteen $(110)_{\text{cubic}}$ planes and was found to be compatible for a wide range of temperature. The equilibrium path corresponding to this structure was found to be isolated from the principal path. However, steepest descent computations were performed to confirm that a direct transition pathway between the modulated martensite and the B2 austenite exists. This implies that the discovered structure belongs to the B2 basin of attraction and thus, a reversible transformation between B2 and the discovered modulated martensite is possible. Additionally, this structure was also found to be the lowest energy structure upon the application of a kinematic compatibility constraint. This suggests that the discovered structure may be a viable candidate for the observed modulated martensite at a phase transition. It is important to note that the search explicitly demonstrated that the M^4 was able to predict the properties of a modulated martensite without a-priori knowledge of the existence of that structure. Through the use of the BFB method and the M^4 it is therefore, in principle, possible to completely obtain the properties of all modulated martensites in this model.

One important question that can be asked here is why are MMs not experimentally observed in AuCd. For that matter, why are MMs not observed in all SMAs? This work suggests

one possible answer—the energies of MMs depend on the energy of the $L1_0$ phase. As such, some materials may have a high enough energy for $L1_0$ so that a nucleation of a MM in this material would substantially increase the overall energy (despite a lower increase in elastic energy compared to the increase in elastic energy due to the nucleation of the ground state martensite). In these cases, it is the ground state martensite that is observed. Another answer to the aforementioned question may be that the energy barrier involved in the transition between a B2 structure and to a MM may be too high. However, the computation of such an energy barrier will not be presented in this work.

In conclusion, the M^4 provides a tool that can be used to predict the existence and properties of MMs in a SMA. This can be achieved with only a small amount of input data which can be obtained from empirical atomistic simulations or DFT calculations. One place where the M^4 has already proven to be useful is shown in the discovery of the MM $k_7^{(2)}$ —a path not connected to the principal path and hence unlikely to be found by the BFB method. As such, it is our belief that the M^4 provides a valuable tool in the investigation of the role of MMs in β -phase SMAs.

Chapter 5

Summary and conclusions

This thesis details the results from an atomistic study of SMAs. In particular, the objective of this work was to examine the existence and features of modulated martensites using a generic material model for SMAs. In Ch. 1 a review of the current understanding of SMAs was presented. The chapter highlights how the shape memory effect and pseudo-elasticity is heavily tied to martensitic phase transformations and the formation of microstructure. Ch. 1 also presented a special type of martensitic phase transformation—one that takes the austenite to a structure known as modulated martensite. These type of MPTs are commonly observed in β -phase alloys where the austenite structure is described by a body centered cubic crystal structure (ignoring atomic species) and the modulated martensite is comprised of a periodic stacking of $(110)_{\text{cubic}}$ planes. Despite common belief modulated martensites are not thermodynamically stable structures but are in fact meta-stable states of higher energy. In fact, DFT calculations performed by *Zayak et al.* (2003) have determined that in these types of materials there exists a martensite structure of lower energy. With this in mind, one question that can be asked is what is the mechanism that stabilizes the formation of modulated martensites in certain SMAs? Two existing theories were presented in Ch. 1 to explain the experimental observation of modulated martensites:

1. *The adaptive martensite hypothesis (Khachaturyan et al. (1991)).*

The AMH views modulated martensites as a nano-scale twinned structure built from a tetragonal ground state martensite. The viewpoint assumes a high crystal-lattice mismatch between the ground state martensite and austenite crystal structures. *Khachaturyan et al.* (1991) argue that the ground state martensite phase cannot homogeneously nucleate because the elastic energy generated by a coherent homogenous nucleus is too high. Therefore, instead of the ground state martensite a homogenous nucleation

of a kinematically compatible modulated martensite is observed. In other words, the modulated martensite is stabilized by an austenite kinematic compatibility condition at the austenite-martensite interface.

2. *The viewpoint of Balandraud and Zanzotto (2007)*

Like the AMH, *Balandraud and Zanzotto (2007)* claim that the observed modulated martensite is stabilized by the austenite kinematic compatibility condition at the austenite-martensite interface. In contrast to the AMH, *Balandraud and Zanzotto (2007)* do not view the modulated structure as a nano-twinned tetragonal martensite but instead view modulated martensites as periodic stacking of close-packed planes. The viewpoint identifies a fundamental planar unit cell whereby the final stacking sequence defines the modulated martensite.

While both theories have found successes in explaining the features of modulated martensites they fail to provide any insight regarding the dynamic stability (local minima) or the thermodynamic stability (global minima) of modulated martensites. This work aimed to address some of the shortcomings of the AMH and the viewpoint of *Balandraud and Zanzotto (2007)*.

Some other questions that arise when considering materials that exhibit modulated martensites are:

- Do stable martensite phases of all modulations exist in the same material?
- What factors prohibit some modulations from being observed in experiments?
- Is the ability of certain materials to undergo a MPT into martensites of multiple modulations inherently special to those materials, or is that ability a general characteristic of all materials that undergo a MPT?

To address the above questions a tool known as the branch-following and bifurcation (BFB) method was used to perform empirical atomistic simulations to probe the energy characteristics of a typical SMA. The BFB study yielded further understanding that could be useful in interpreting what modulated martensites are, why they form, and what role they play in SME and PE. To provide the reader with some background, a brief discussion on the BFB method was presented in Ch. 2.

Chapter 3 details the results of the BFB study of the material model used in this work. The model was developed by *Guthikonda and Elliott (2011)* and was demonstrated to cap-

ture the general behavior of β -phase SMAs. Additionally, the model was able to capture dynamic instabilities associated with wavevectors in the $(110)_{\text{cubic}}$ direction, which is the primary mode in the phase transition into modulated martensites. As such, a BFB study of this model was performed to investigate the existence of modulated martensites. The BFB method was augmented with the Cascading Cauchy Born technique to detect period-extending bifurcation points associated with the wavevectors $\mathbf{k}_5 = \frac{2}{5}(110)_{\text{cubic}}$, $\mathbf{k}_7 = \frac{3}{7}(110)_{\text{cubic}}$, $\mathbf{k}_9 = \frac{4}{9}(110)_{\text{cubic}}$, $\mathbf{k}_{14} = \frac{5}{14}(110)_{\text{cubic}}$, and $\mathbf{k}_{18} = \frac{7}{18}(110)_{\text{cubic}}$. These wavevectors correspond to periodicities of commonly observed modulated martensites. The existence of these paths were verified. Note that this result *explicitly demonstrates the existence of modulated martensites in this material model and hence addresses one of the questions outlined in the beginning of this chapter*. This suggests that the presence of modulated martensites is a general feature of the energy landscape of the model. Additionally, the BFB study of the model also found that only paths corresponding to wavevectors \mathbf{k}_5 , \mathbf{k}_7 , \mathbf{k}_9 contain stable segments. This is because the \mathbf{k}_{14} and \mathbf{k}_{18} paths bifurcate away from the principal path at points where wavevectors of lower wavelengths have not yet become phonon stable. Motivated by this reasoning, we put forth the conjecture that modulated martensites with odd wavelengths are more likely to be stable. Ch. 3 also details the effect of an austenite kinematic compatibility constraint on modulated martensites. Here, it is found that the ground state B19 structure showed a dramatic increase in energy when subjected to an austenite compatibility constraint. This result lends credence to the viewpoints of the AMH and *Balandraud and Zanzotto (2007)* who claim that the high mismatch in kinematic compatibility could substantially increase the elastic energy when forming an austenite-martensite interface. In contrast to the B19 structure the modulated martensites are better able to accommodate the compatibility constraint. It was found that \mathbf{k}_5 showed the smallest increase in energy when constraint in this manner. For this reason, it is reasonable to assume that \mathbf{k}_5 would be the experimentally observed modulated structure.

In Ch. 4 a predictive tool called the *Modulated Martensite Mixture Model* (M^4) was presented. The M^4 can accurately predict the energy, lattice constants, and structural details of an arbitrary modulated martensite. The main idea behind the M^4 is that the model interprets modulated martensites as a mixture of two short period base martensites—B19 and $L1_0$. The predictive performance of the M^4 was evaluated by comparing the relative error between the predicted and actual properties of \mathbf{k}_5 , \mathbf{k}_7 , and \mathbf{k}_9 . The M^4 was found to be comparable to the *Balandraud and Zanzotto (2007)* model in predicting configurational properties such as principal values of the deformation gradient and fractional interatomic shifts of atoms in the deformed unit cell. However, the M^4 provides an additional capability

in that it is able to accurately predict the energies and stability of an arbitrary modulated martensite. The capabilities of the M^4 were also demonstrated by using it to predict the existence of a highly compatible modulated martensite. The existence of the compatible modulated martensite was verified by explicit atomistic computation where it was also found to belong to the B2 basin of attraction. This implies that a reversible MPT between B2 and the discovered modulated martensite is possible. Finally, this structure was also found to be the lowest energy structure upon the application of a kinematic compatibility constraint suggesting that the structure may be a viable candidate for the observed modulated martensite at a phase transition.

In conclusion, the BFB study of this model suggested that modulated martensites are natural features of the energy landscape of SMAs. In contrast to the AMH and *Balandraud and Zanzotto (2007)* viewpoint the results from the BFB study provided information regarding the energies of the modulated martensites present in the model. It was found that the modulated martensites are low energy phases (in that the difference in energy between a modulated martensite and the ground state martensite is much smaller than the difference in energy between the austenite and ground state martensite). Furthermore, it was demonstrated that the modulated martensites are better able to accommodate an austenite compatibility constraint providing a possible explanation for the experimental observation of modulated martensites. Finally, the BFB study allowed for the development of the M^4 . The M^4 provides a tool that can be used to predict the existence and properties of modulated martensites in a SMA. This can be achieved with only a small amount of input data which can be obtained from empirical atomistic simulations or DFT calculations.

5.1 Some open questions

While this work was successful at explicitly demonstrating the existence of modulated martensites as well as providing information regarding the relative stability of different modulated martensites, there remain some open questions that can be addressed as future work on this topic. This section will list a few of the immediate tasks that can be done to advance the understanding of the role of modulated martensites in SMAs.

1. In this work we argued that modulated martensites are experimentally observed because they accommodate an austenite compatibility constraint better than the ground state martensite. While this explanation is applicable at the habit plane, it does not explain the observation of modulated martensites in the bulk, and far away from the austenite-martensite interface. One explanation for this can be found in *Kaufmann*

et al. (2010) where they suggest that once the compatible modulated martensite is generated at the austenite-martensite interface, a high energy barrier between the modulated and base martensite “locks” the configuration in place so as to be observed in the bulk. As such, investigation into the energy barriers between modulated martensites and its base martensite will yield information that may support the conjecture stated in *Kaufmann et al.* (2010).

2. In some materials a microstructure often referred to as “tweed” (*Kartha et al.* (1991)) is observed during martensitic phase transformation (*Sethna et al.* (1992)). These are fabric-like cross-hatched patterns with stripes along families of $\langle 110 \rangle_{\text{cubic}}$ directions with widths that are a few atomic spacing (*Kartha et al.* (1995)). As it was natural to view modulated martensites as the onset of instability associated with wavevectors in the $(110)_{\text{cubic}}$ direction, the tweed pattern naturally lends itself to an extension of the idea as the onset instability along two (or three) directions from the family of $\{110\}_{\text{cubic}}$ planes. As such, it will be useful to demonstrate the existence of tweed patterns using the *Guthikonda and Elliott* (2011) model.
3. To assess the validity of the *Guthikonda and Elliott* (2011) model, it would be useful to compare the configuration of \mathbf{k}_5 , \mathbf{k}_7 , and \mathbf{k}_9 to experimentally observed modulated martensites with modulations of 10, 14, and 18 $(110)_{\text{cubic}}$ planes, respectively.
4. Finally, the M^4 model was developed to be independent of the particular material model. That is, the M^4 is designed to be able to predict the properties of modulated martensites in any material model as long as the model is able to capture the B19 and $L1_0$ configurations. As such, it would be useful to explore the sensitivity of the predictive accuracy of the M^4 on the model parameters in order to validate the transferability of the M^4 .

Bibliography

- E. L. Allgower. *Numerical continuation methods*. Springer-Verlag, 1980.
- M. Arndt, V. Sorkin, and E. B. Tadmor. Efficient algorithms for discrete lattice calculations. *Journal of Computational Physics*, 228(13):4858–4880, 2009.
- N. Ashcroft and N. Mermin. *Solid state physics*. Saunders College, 1976.
- P. Bak and J. von Boehm. Ising model with solitons, phasons, and “the devil’s staircase”. *Physical Review B*, 21(11):5297–5308, 1980.
- X. Balandraud and G. Zanzotto. Stressed microstructures in thermally induced M9R-M18R martensites. *Journal of the Mechanics and Physics of Solids*, 55(1):194–224, 2007.
- J. M. Ball and R. D. James. Fine phase mixtures as minimizers of energy. *Archive for Rational Mechanics and Analysis*, 100(1):13–52, 1987.
- K. Bhattacharya. Wedge-like microstructure in martensites. *Acta Metallurgica et Materialia*, 39(10):2431–2444, 1991.
- K. Bhattacharya. *Microstructure of martensite: Why it forms and how it gives rise to the shape memory effect*. Oxford University Press, 2003.
- K. Bhattacharya, S. Conti, G. Zanzotto, and J. Zimmer. Crystal symmetry and the reversibility of martensitic transformations. *Nature*, 428:55–59, 2004.
- H. Bohm. Modulated structures at phase transitions. *American Mineralogist*, 68:11–17, 1983.
- M. Born and K. Huang. *Dynamical theory of crystal lattices*. Oxford University Press, 1954.

- J. S. Bowles and C. M. Wayman. The bain strain, lattice correspondences, and deformations related to martensitic transformations. *Metallurgical and materials transactions B*, 3:1113–1121, 1972.
- A. L. Browne, A. C. Keefe, P. W. Alexander, N. Mankame, P. Usoro, N. L. Johnson, J. Aase, P. Sarosi, G. P. McKnight, G. Herrera, C. Churchill, J. Shaw, and J. Brown. Lightweight thermal energy recovery system based on shape memory alloys: a DOE ARPA-E initiative. *SPIE Conference Series*, 8343, 2012.
- B. Budiansky. Theory of buckling and post-buckling behavior of elastic structures. *Advances in applied mechanics*, 14:1–65, 1974.
- W. J. Buehler and F. E. Wang. A summary of recent research on the nitinol alloys and their potential application in ocean engineering. *Ocean Engineering*, 1(1):105–120, 1968.
- A. Bystrom and K. E. Almin. X-ray investigation of AuCd alloys rich in Au. *Acta Chemica Scandinavica*, 1(1):76–89, 1947.
- E. Cancès, M. Defranceschi, W. Kutzelnigg, C. L. Bris, and Y. Maday. Computational quantum chemistry: A primer. *Handbook of Numerical Analysis*, 10, 2003.
- Y. Chen. Local stress and heat flux in atomistic systems involving three-body forces. *The Journal of Chemical Physics*, 124(5):054113, 2006.
- C. B. Churchill and J. Shaw. Shakedown response of conditioned shape memory alloy wire. *Proceedings of the SPIE 15th annual international symposium on smart structures and materials*, page 69291F, 2008.
- R. A. Cowley. Structural phase transitions I. Landau theory. *Advances in Physics*, 29(1):1–110, 1980.
- P. J. Davis. *Circulant matrices*. Chelsea Publishing, 1994.
- A. F. Devonshire. Theory of barium-titanate—Part I. *Philosophical Magazine*, 40:1040–1063, 1949.
- A. F. Devonshire. Theory of barium-titanate—Part II. *Philosophical Magazine*, 42:1065–1079, 1951.
- M. Dobson, R. S. Elliott, and E. B. Tadmor. A multilattice quasicontinuum for phase transforming materials: Cascading Cauchy Born kinematics. *Journal of Computer-Aided Materials Design*, 14(S1):219–237, 2008.

- M. T. Dove. *Introduction to Lattice Dynamics*. Cambridge University Press, 1993.
- T. Duerig, A. Pelton, and C. Trepanier. Nitinol. *SMST e-Elastic newsletter*, 2011.
- B. Eck, R. Dronskowski, M. Takahashi, and S. Kikkawa. Theoretical calculations on the structures, electronic and magnetic properties of binary 3D transition metal nitrides. *Journal of Materials Chemistry*, 9(7):1527–1537, 1999.
- R. S. Elliott. Multiscale bifurcation and stability of multilattices. *Journal of Computer-Aided Materials Design*, 14:143–157, 2007.
- R. S. Elliott, N. Triantafyllidis, and J. A. Shaw. Stability of thermally-induced martensitic transformations in bi-atomic crystals. *Journal of the Mechanics and Physics of Solids*, 50(11):2463–2493, 2002.
- R. S. Elliott, N. Triantafyllidis, and J. A. Shaw. Stability of crystalline solids—I: Continuum and atomic lattice considerations. *Journal of the Mechanics and Physics of Solids*, 54(1):161–192, 2006a.
- R. S. Elliott, N. Triantafyllidis, and J. A. Shaw. Stability of crystalline solids—II: Application to temperature-induced martensitic phase transformations in a bi-atomic crystal. *Journal of the Mechanics and Physics of Solids*, 54(1):193–232, 2006b.
- J. L. Ericksen. *Phase transformations and material instabilities in solids*. Academic Press, 1984.
- P. F. Gostin, S. Oswald, L. Schultz, and A. Gebert. Acid corrosion process of Fe-based bulk metallic glass. *Corrosion Science*, 62(C):112–121, 2012.
- V. S. Guthikonda and R. S. Elliott. An effective interaction potential model for the shape memory alloy AuCd. *Continuum Mechanics and Thermodynamics*, 21(4):269–295, 2009.
- V. S. Guthikonda and R. S. Elliott. Erratum to: An effective interaction potential model for the shape memory alloy AuCd. *Continuum Mechanics and Thermodynamics*, 23(2):177–183, 2011.
- R. Hamilton, H. Sehitoglu, C. Efstathiou, and H. Maier. Inter-martensitic transitions in Ni-Fe-Ga single crystals. *Acta Materialia*, 55(14):4867–4876, 2007.
- K. F. Hane and T. W. Shield. Microstructure in the cubic to monoclinic transition in titanium-nickel shape memory alloys. *Acta Materialia*, 47(9):2603–2617, 1999.

- T. J. Healey. A group-theoretic approach to computational bifurcation problems with symmetry. *Computer Methods in Applied Mechanics and Engineering*, 67(3):257–295, 1988.
- T. J. Healey and U. Miller. Two-phase equilibria in the anti-plane shear of an elastic solid with interfacial effects via global bifurcation. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 463(2080):1117–1134, 2007.
- D. Helm and P. Haupt. Shape memory behaviour: modelling within continuum thermomechanics. *International Journal of Solids and Structures*, 40(4):827–849, 2003.
- G. A. Holzapfel. *Nonlinear solid mechanics*. John Wiley & Sons, LTD., 2000.
- R. D. James. Displacive phase transformations in solids. *Journal of the Mechanics and Physics of Solids*, 34(4):359–394, 1986.
- R. D. James and D. Kinderlehrer. Theory of diffusionless phase transitions. *PDEs and continuum models of phase transitions*, 344:51–84, 1989.
- V. Jusuf. *Algorithms for branch-following and critical point identification in the presence of symmetry*. Master’s thesis, The University of Minnesota, 2010.
- V. Jusuf. *A new framework for the interpretation of modulated martensites in shape memory alloys*. Ph.D. thesis, The University of Minnesota, 2013.
- S. Kartha, T. Castán, J. A. Krumhansl, and J. P. Sethna. Spin-glass nature of tweed precursors in martensitic transformations. *Physical Review Letters*, 67(25):3630–3633, 1991.
- S. Kartha, J. A. Krumhansl, J. P. Sethna, and L. K. Wickham. Disorder-driven pretransitional tweed pattern in martensitic transformations. *Physical Review B*, 52(2):803, 1995.
- G. B. Kauffman and I. Mayo. The story of Nitinol: The serendipitous discovery of the memory metal and its applications. *The Chemical Educator*, 2(2):1–21, 1997.
- S. Kaufmann, R. Niemann, T. Thersleff, U. K. Röbber, O. Heczko, J. Buschbeck, B. Holzapfel, L. Schultz, and S. Fähler. Modulated martensite: why it forms and why it deforms easily. *New Journal of Physics*, 13(5):053029, 2011.
- S. Kaufmann, U. K. Röbber, O. Heczko, M. Wuttig, J. Buschbeck, L. Schultz, and S. Fähler. Adaptive modulations of martensites. *Physical Review Letters*, 104(14):145702, 2010.
- A. G. Khachaturyan, S. M. Shapiro, and S. Semenovskaya. Adaptive phase formation in martensitic transformation. *Physical Review B*, 43(13):10832–10843, 1991.

- C. Kittel. *Introduction to solid state physics*. John Wiley & Sons, Inc., 1986.
- I. Kiyohiro, M. Kazuo, and F. Hiroshi. Bifurcation hierarchy of symmetric structures. *International Journal of Solids and Structures*, 27(12):1551–1573, 1991.
- M. Krystian and W. Pichl. In situ optical microscopy of the martensitic phase transformation of lithium. *Physical Review B*, 62(21):13956–13962, 2000.
- F. Lovey. The fault density in 9R type martensites: a comparison between experimental and calculated results. *Acta Metallurgica*, 35(5):1103–1108, 1987.
- L. E. Malvern. *Introduction to the mechanics of a continuous medium*. Prentice-Hall, inc., 1969.
- P. J. McCluskey, C. Zhao, O. Kfir, and J. J. Vlassak. Precipitation and thermal fatigue in Ni-Ti-Zr shape memory alloy thin films by combinatorial nanocalorimetry. *Acta Materialia*, 59(13):5116–5124, 2011.
- R. McWeeny. *Symmetry: An introduction to group theory and its applications*. Dover Publications, 2002.
- H. D. Megaw. Temperature changes in the crystal structure of barium titanium oxide. *Proceedings of the Royal Society of London A.*, 189(1017):261–283, 1947.
- K. N. Melton, J. L. Proft, and T. W. Duerig. Wide hysteresis shape memory alloys based on the Ni-Ti-Nb system. *Proceedings of the MRS International Meeting on Advanced Materials*, 9:165–170, 1988.
- W. Miller. *Symmetry groups and their applications*. Academic Press, 1972.
- V. J. Minkiewicz, G. Shirane, and R. Nathans. Phonon dispersion relation for iron. *Physical Review*, 162(3):528–531, 1967.
- T. Natsuki and M. Endo. Stress simulation of carbon nanotubes in tension and compression. *Carbon*, 42(11):2147–2151, 2004.
- Z. Nishiyama and S. Kajiwara. Electron microscope study of the crystal structure of the martensite in a copper-aluminium alloy. *Japanese journal of Applied Physics*, 2(8):478–486, 1963.
- A. Olander. The crystal structure of AuCd. *Zeitschrift Fur Kristallographie*, 83(1/2):145–148, 1932.

- K. Otsuka and C. M. Wayman. *Shape memory materials*. Cambridge University Press, 1999.
- A. Pelton, D. Stockel, and T. Duerig. Medical uses of Nitinol. *Materials Science Forum*, 327-328:63–70, 2000.
- J. Pons, V. A. Chernenko, R. Santamarta, and E. Cesari. Crystal structure of martensitic phases in Ni-Mn-Ga shape memory alloys. *Acta Materialia*, 48(12):3027–3038, 2000.
- P. L. Potapov, S. E. Kulkoval, A. V. Shelyakov, K. Okutsu, S. Miyazaki, and D. Schryvers. Crystal structure of orthorhombic martensite in TiNi-Cu and TiNi-Pd. *Intermetallics*, 112:727–730, 2003.
- D. E. Sands. *Introduction to Crystallography*. Dover, 1993.
- M. A. Savi, A. Paiva, and P. M. C. L. Pacheco. Phenomenological modeling of shape memory alloy thermomechanical behavior. *Proceedings of the International Symposium on Solid Mechanics*, pages 497–511, 2007.
- S. Semenovskaya and A. G. Khachatryan. Structural transformations in nonstoichiometric $\text{YBa}_2\text{Cu}_3\text{O}_{6+\delta}$. *Physical Review B*, 46(10):6511–6534, 1992.
- J. P. Sethna, S. Kartha, T. Cast'an, and J. A. Krumhansl. Tweed in martensites: A potential new spin glass. *Physica Scripta*, T42:214–219, 1992.
- J. A. Shaw. *Material instabilities in a nickel-titanium shape memory alloy*. Ph.D. thesis, The University of Texas at Austin, 1997.
- V. Sorkin, R. S. Elliott, and E. B. Tadmor. A local quasicontinuum for 3D multilattice crystalline materials: application to shape-memory alloys. *Physical Review B*, pages 1–17, submitted 2012.
- M. O. Steinhauser. *Computational multiscale modeling of fluids and solids: Theory and applications*. Springer Publishing Company, 2007.
- Q. P. Sun, T. T. Xu, and X. Zhang. On deformation of A-M interface in single crystal shape memory alloys and some related issues. *Journal of Engineering Materials and Technology*, 121(1):38–43, 1999.
- E. B. Tadmor, R. E. Miller, and R. S. Elliott. *Continuum mechanics and thermodynamics: From fundamental concepts to governing equations*. Cambridge University Press, 2012.

- S. Tanisaki. Microdomain structure in paraelectric phase of NaNO_2 . *Journal of the Physical Society of Japan*, 16(3):579–579, 1961.
- S. Tanisaki. X-ray study on the ferroelectric phase transition of NaNO_2 . *Journal of the Physical Society of Japan*, 18(8):1181–1191, 1963.
- R. J. D. Tilley. *Crystals and crystal structures*. Wiley, 2006.
- J.-C. Toledano and P. Toledano. *The Landau theory of phase transitions*. World Scientific Publishing, 1987.
- C. Torres, A. Condó, F. Lovey, and J. Guimpel. Microstructure in Cu-Al-Ni shape memory alloy thin films. *Acta Microscopica*, 16(1):153–154, 2007.
- N. Triantafyllidis and R. Peek. On stability and the worst imperfection shape in solids with nearly simultaneous eigenmodes. *International Journal of Solids and Structures*, 29(18):2281–2299, 1992.
- K. Tsuzaki, Y. Natsume, and T. Maki. Transformation reversibility in Fe-Mn-Si shape memory alloy. *Le Journal de Physique IV*, 05(C8):409–414, 1995.
- H. Unoki and T. Sakudo. Electron spin resonance of Fe^{3+} in SrTiO_3 with special reference to the 110° K phase transition. *Journal of the Physical Society of Japan*, 23(3):546–552, 1967.
- H. Warlimont and L. Delaey. *Martensitic Transformations in Copper-, Silver- and Gold-based alloys (Progress in Materials Science Vol 18)*. Pergamon Press, 1974.
- M. Wechsler. *On the theory of the formation of martensite*. Ph.D. thesis, Columbia University, 1954.
- Y. Yamada, I. Shibuya, and S. Hoshino. Phase transition in NaNO_2 . *Journal of the Physical Society of Japan*, 18(11):1594–1603, 1963.
- A. Zayak, P. Entel, J. Enkovaara, A. Ayuela, and R. Nieminen. First-principles investigations of homogeneous lattice-distortive strain and shuffles in Ni_2MnGa . *Journal of Physics: Condensed Matter*, 15:159, 2003.
- Z. Zhang, R. D. James, and S. Müller. Energy barriers and hysteresis in martensitic phase transformations. *Acta Materialia*, 57(15):4332–4352, 2009.

Appendix A

Symmetry

Since symmetry and phase transitions are deeply tied, the consequence of symmetry will be discussed in this appendix. We will show that the presence of symmetry will imply certain restrictions regarding the nature of the equilibrium paths. These restrictions will aid in the computation of quantities such as $E_{,\xi\xi\lambda}$, $E_{,\xi\xi\xi}$ or $E_{,\xi\xi\xi\xi}$ that are needed in performing an LSK asymptotic analysis at a bifurcation point. These quantities are computationally expensive to compute and as such, any knowledge of the form of these quantities will imply an increase in computational efficiency. One benefit in problems with symmetry is that the presence of symmetry will dictate the form of any quantities obeying the symmetry of the problem.

A.1 Symmetry

While the computation of equilibrium paths through the use of the BFB method mentioned in the preceding sections is applicable to problems that have no symmetry, the presence of symmetry implies several desirable consequences that may be used to characterize and compute equilibrium paths.

In this section, the theory behind symmetry techniques will be briefly discussed. The purpose of this discussion is to provide some context behind the mathematical tools that will be used in various parts of this work. It is not meant to be an in-depth presentation and as such will not cover various proofs of the results presented. For a more detailed presentation of the various topics, refer to *McWeeny* (2002) or *Miller* (1972).

A.1.1 Group Representation Theory

Here, we will follow the general theory of linear group representations as presented in *McWeeny* (2002) and *Miller* (1972). Let \mathbb{R}^N be a vector space with the usual inner product $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x} \cdot \mathbf{y} = \sum_i^N x_i y_i$ and let G be an abstract group.

In this work, G will be a finite abstract symmetry group whose order will be denoted as $|G|$ (where $|G| < \infty$). In general, the matrix representation of G is non-unique. In fact, every choice of a basis for \mathbb{R}^N yields a new matrix representation of G . However, any two matrix representations \mathbf{T} and $\widehat{\mathbf{T}}$ can be equivalent in the sense defined as follows.

Definition. Two matrix representation \mathbf{T} and $\widehat{\mathbf{T}}$ are *equivalent* (denoted $\mathbf{T} \cong \widehat{\mathbf{T}}$) if there exists $\mathbf{S} \in \mathbb{R}^{N \times N}$ such that

$$\mathbf{T}(g) = \mathbf{S} \widehat{\mathbf{T}}(g) \mathbf{S}^{-1}, \quad \forall g \in G. \quad (\text{A.1})$$

One consequence of the finite order group G is that any matrix representation \mathbf{T} of G on the inner product space \mathbb{R}^N , is equivalent to an *orthogonal* matrix representation.

A.1.1.1 Invariant Subspaces and Irreducible Representations

Definition. $W \subset \mathbb{R}^N$ is an *invariant subspace* under a matrix representation \mathbf{T} if

$$\mathbf{T}(g)\mathbf{w} \in W, \quad \forall \mathbf{w} \in W, \quad \text{and} \quad \forall g \in G. \quad (\text{A.2})$$

Definition. The matrix representation \mathbf{T} is *reducible* if there is a proper subspace W of \mathbb{R}^N invariant under \mathbf{T} . Otherwise, \mathbf{T} is *irreducible*.

Note that by the above definition, a representation is irreducible if the only invariant subspaces of \mathbb{R}^N are the trivial subspaces \mathbb{R}^N and $\{\mathbf{0}\}$.

If \mathbf{T} is reducible and $W \subset \mathbb{R}^N$ is an invariant subspace of dimension k where $k \in [1, N)$, then it can be shown that \mathbf{T} is equivalent to a *block-diagonal matrix representation*. That is, in the appropriate bases, the matrices of $\mathbf{T}(g)$ takes the following form:

$$\begin{bmatrix} \widetilde{\mathbf{T}}(g) & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{T}}(g) \end{bmatrix}, \quad \widetilde{\mathbf{T}}(g) \in \mathbb{R}^{k \times k}, \quad \text{and} \quad \widehat{\mathbf{T}}(g) \in \mathbb{R}^{(N-k) \times (N-k)}, \quad (\text{A.3})$$

where $\widetilde{\mathbf{T}}(g) \in \mathbb{R}^{k \times k}$ is a matrix representation on the invariant subspace W . An important result here is that, It can be shown that all finite dimensional matrix representations of G

can be written as *direct sum of a finite number of irreducible representations*. For many abstract symmetry groups G , these irreducible representations can be found tabulated in books on symmetry such as *McWeeny* (2002). What this means is that in an appropriate basis, any finite dimensional matrix representation $\mathbf{T}(g)$ can be expressed as a block diagonal representation whose blocks are comprised of irreducible representations where each irreducible representation can occur with a certain multiplicity. So, if $\mathbf{T}^\mu(g)$, $\mu = 1, \dots, \alpha$ denotes α distinct irreducible representations and n_μ gives the number of times that the μ -th irreducible representation occurs in $\mathbf{T}(g)$, then

$$\mathbf{T}(g) = \bigoplus_{\mu=1}^{\alpha} \bigoplus_{i=1}^{n_\mu} \mathbf{T}^\mu(g), \quad \forall g \in G. \quad (\text{A.4})$$

With every irreducible representations, there is an associated invariant subspace W^μ that transforms according to the irreducible representation. That is, for each $\mu \in 1, 2, \dots, \alpha$, there exists an invariant subspace $W^\mu \subset \mathbb{R}^N$ such that

$$\mathbf{T}^\mu(g)\mathbf{w} \in W^\mu, \quad \forall \mathbf{w} \in W^\mu, \quad \text{and} \quad \forall g \in G. \quad (\text{A.5})$$

In this manner, \mathbb{R}^N can be decomposed into a direct sum of invariant subspaces as reflected by the following equation:

$$\mathbb{R}^N = \text{span}\{\boldsymbol{\psi}_{i,\beta}^\mu \mid i = 1, 2, \dots, N_\mu; \beta = 1, 2, \dots, n_\mu; \mu = 1, \dots, \alpha\}, \quad (\text{A.6})$$

where $\{\boldsymbol{\psi}_{i,\beta}^\mu \mid i = 1, 2, \dots, N_\mu; \beta = 1, 2, \dots, n_\mu\}$ form a basis for the $(N_\mu n_\mu)$ -dimensional invariant subspaces W^μ and $\mathbf{T}^\mu(g) \in \mathbb{R}^{N_\mu \times N_\mu}$ (i.e., \mathbf{T}^μ is a matrix representation of G on \mathbb{R}^{N_μ}). It can be shown that n_μ is given by

$$n_\mu = \frac{1}{|G|} \sum_{g \in G} \overline{\text{tr}(\mathbf{T}^\mu(g))} \text{tr}(\mathbf{T}(g)), \quad (\text{A.7})$$

where $\text{tr}(\cdot)$ indicates the trace of a matrix and $\overline{(\cdot)}$ indicates complex conjugation. Often, the quantity $\text{tr}(\mathbf{T}^\mu(g))$ is referred to as the *character* for the μ -th irreducible representation of $g \in G$ and can be found in character tables in books such as *McWeeny* (2002). One surprising result is that, if $\mathbf{S} \in \mathbb{R}^{N \times N}$ is such that $\mathbf{S} = [\boldsymbol{\psi}_{i=1,\beta=1}^{\mu=1} \dots \boldsymbol{\psi}_{i=N_\alpha,\beta=n_\alpha}^{\mu=\alpha}]$, then \mathbf{S} simultaneously block-diagonalizes $\mathbf{T}(g)$, $\forall g \in G$.

One method to obtain the bases for a particular subspace is through the use of the *Projection operator*. The projection operator $\mathbf{P}^\mu : \mathbb{R}^N \rightarrow W^\mu$ is a linear operator that maps any vector

$\mathbf{v} \in \mathbb{R}^N$ onto its projection into the invariant subspace W^μ and is constructed as follows:

$$\mathbf{P}^\mu = \frac{\text{tr}(\mathbf{T}^\mu(e))}{|G|} \sum_{g \in G} \overline{\text{tr}(\mathbf{T}^\mu(g))} \mathbf{T}(g), \quad (\text{A.8})$$

where $\mathbf{T}^\mu(e)$ denotes the μ -th irreducible representation associated with the identity element denoted by 'e'. By definition, $\mathbf{P}^\mu \boldsymbol{\psi}^\mu = \boldsymbol{\psi}^\mu, \forall \boldsymbol{\psi}^\mu \in W^\mu$. Therefore, the $(N_\mu n_\mu)$ basis vectors for the subspace W^μ can be found by solving for the $(N_\mu n_\mu)$ eigenvectors of \mathbf{P}^μ with eigenvalue equal to one.

A.1.2 Some Consequences of Symmetry

Now that group representation theory has been briefly presented, the consequences of symmetry will be discussed in this section. In particular, the decomposition of \mathbb{R}^N into invariant subspaces (discussed above) will yield certain desirable results that can be used to categorize equilibrium paths based on their symmetry. Additionally, the block-diagonal structure of the irreducible matrix representation will impose certain restrictions regarding the null space in a symmetry-breaking bifurcation point that will dictate the forms of invariances in the null space. These results will be used in this work to aide in the computation of equilibrium solutions (via a BFB method) to systems with symmetry.

A.1.2.1 Branch-following in the presence of symmetry

In the presence of symmetry, the equilibrium equations can be reduced so that certain solutions can be obtained by solving a lower-dimensional set of equations. The idea behind this approach is that the equilibrium equations can be solved so that only solutions that preserve the symmetry of the original problem are found. In this way, symmetry related degrees of freedom that trivially satisfy the equilibrium equations can be identified, from which a lower-dimensional set of equations can then be obtained.

Recall that in this work, the equilibrium equations are derived from a potential function as in Eq. (2.2). For systems with symmetry (say symmetry group G), then the potential function will remain invariant under the action of the elements of the symmetry group as given by

$$\mathcal{E}(\mathbf{T}(g)\mathbf{u}, \lambda) = \mathcal{E}(\mathbf{u}, \lambda), \quad \forall g \in G. \quad (\text{A.9})$$

A consequence of this invariance is the following condition obtained by taking the gradient with respect to \mathbf{u} of Eq. (A.9):

$$\mathbf{F}(\mathbf{T}(g)\mathbf{u}, \lambda) = \mathbf{T}(g)\mathbf{F}(\mathbf{u}, \lambda), \quad \forall g \in G. \quad (\text{A.10})$$

Equations that satisfy the above condition are said to be *equivariant* under the action of the symmetry group G .

As mentioned in the preceding subsection, the symmetry group G induces a decomposition of \mathbb{R}^N into invariant subspaces. One particular invariant subspace is the *G-symmetric* subspace of \mathbb{R}^N (denoted by S_G) associated with the identity irreducible representation (denoted by E) and is defined as follows:

$$S_G \equiv \{\mathbf{u} \in \mathbb{R}^N \mid \mathbf{T}(g)\mathbf{u} = \mathbf{u}, \forall g \in G\}. \quad (\text{A.11})$$

Note that the identity irreducible representation \mathbf{T}^E is a representation of \mathbb{R} which implies that $N_E = 1$ and therefore $\dim(S_G) = N_E n_E = n_E$. The basis vectors that span S_G can be found by solving for n_E eigenvectors of \mathbf{P}^E with eigenvalue equal to one where n_E is obtained from Eq. (A.7) and is given by

$$n_E = \frac{1}{|G|} \sum_{g \in G} \text{tr}(\mathbf{T}(g)), \quad (\text{A.12})$$

and \mathbf{P}^E is obtained from Eq. (A.8) and is given by

$$\mathbf{P}^E = \frac{1}{|G|} \sum_{g \in G} \mathbf{T}(g). \quad (\text{A.13})$$

Note that Eq. (A.12) and Eq. (A.13) are obtained from Eq. (A.7) and Eq. (A.8) with $\text{tr}(\mathbf{T}^E(g)) = 1, \forall g \in G$.

Proposition 1. $\mathbf{F}(\mathbf{u}, \lambda) \in S_G$ whenever $\mathbf{u} \in S_G$

Proof. By the definition of the G -symmetric subspace, for any $\mathbf{u} \in S_G$, then

$$\mathbf{T}(g)\mathbf{u} = \mathbf{u}, \quad \forall g \in G. \quad (\text{A.14})$$

Therefore, for any $\mathbf{u} \in S_G$, then

$$\mathbf{F}(\mathbf{u}, \lambda) = \mathbf{F}(\mathbf{T}(g)\mathbf{u}, \lambda), \quad \forall g \in G. \quad (\text{A.15})$$

Furthermore, by equivariance (Eq. (A.10)) then $\mathbf{F}(\mathbf{T}(g)\mathbf{u}, \lambda) = \mathbf{T}(g)\mathbf{F}(\mathbf{u}, \lambda)$, $\forall g \in G$ which gives the desired result:

$$\mathbf{F}(\mathbf{u}, \lambda) = \mathbf{T}(g)\mathbf{F}(\mathbf{u}, \lambda), \quad \forall g \in G, \quad (\text{A.16})$$

which satisfies the criteria for a vector to be in S_G , therefore $\mathbf{F}(\mathbf{u}, \lambda) \in S_G$. \square

The above proposition shows that $\mathbf{F}(\mathbf{u}, \lambda) \in S_G$ whenever $\mathbf{u} \in S_G$. Therefore, to solve for $\mathbf{u} \in S_G$ (G -symmetric equilibrium solutions), it suffices to solve the equilibrium equations restricted to S_G since in general, any tensor that is invariant under the action of a representation \mathbf{T} must belong to that invariant subspace associated with the identity irreducible representation.

Equation (A.16) allows for the construction of the symmetry-reduced set of equations $\mathbf{F}_G(\mathbf{v}, \lambda)$ in the following way:

$$\mathbf{F}_G(\mathbf{v}, \lambda) \equiv (\mathbf{\Psi}_G)^\top \mathbf{F}(\mathbf{\Psi}_G \mathbf{v}, \lambda) = \mathbf{0}, \quad (\text{A.17})$$

where $\mathbf{v} \in \mathbb{R}^{n_E}$, $\mathbf{\Psi}_G \equiv [\boldsymbol{\psi}_1 \dots \boldsymbol{\psi}_{n_E}]$ is an $N \times n_E$ matrix that maps \mathbf{v} into S_G and $\{\boldsymbol{\psi}_i | i = 1, 2, \dots, n_E\}$ are eigenvectors of \mathbf{P}^E with eigenvalue equal to one.

An important result (Healey (1988)) is that (\mathbf{v}, λ) satisfies $\mathbf{F}_G(\mathbf{v}, \lambda) = \mathbf{0}$ if and only if $\mathbf{F}(\mathbf{\Psi}_G \mathbf{v}, \lambda) = \mathbf{0}$. Thus, there is a one-to-one relationship between the solution (\mathbf{v}, λ) of the reduced equations and the G -symmetric solutions of the full equations. In this manner, by using \mathbf{F}_G to solve for G -symmetric solutions, one can systematically categorize the equilibrium solution paths based on its symmetry group G .

A.1.2.2 Implications of symmetry on asymptotic bifurcation analysis at a symmetry-breaking bifurcation point

Assume that a system Ω with DOFs in \mathbb{R}^N can be described by a potential energy function \mathcal{E} . Also, assume that Ω is symmetric with respect to a symmetry group G with matrix representation $\mathbf{T}(g)$, $g \in G$, on \mathbb{R}^N . In general, any functions on the system will inherit the symmetry of the system. This will yield desirable consequences in the computation of the coefficients in the expansion of the bifurcation equation when performing an asymptotic bifurcation analysis.

Recall the decomposition of \mathbb{R}^N into invariant subspaces as in Eq. (A.5) where each invariant subspace is associated to a particular irreducible representation of G . It can be shown (Cowley (1980)) that at a second-order phase transition, the *distortion in the structure can*

be described to be contained in a subspace associated with only one irreducible representation (say, μ -th irreducible representation). This means that at a symmetry-breaking bifurcation point, the null-space is then a subspace of W^μ , the invariant subspace associated with the μ -th irreducible representation. Furthermore, the dimension of the null-space is equal to the dimension of the irreducible representation and is here denoted by N_μ . This will yield a desirable consequence in terms of LSK decomposition, namely that the form of the coefficients in the expansion of the bifurcation equation can be constructed through the irreducible representation associated with the null space. First we will establish the following result that states that the bifurcation equation is equivariant under the action of G (i.e., that the bifurcation equations inherit the group symmetry of the full equations).

Proposition 2. *Let the bifurcation equation $E_{,\eta}(\boldsymbol{\eta}, \Delta\lambda)$ be defined as in Eq. (2.14) in Sec. 2.2.1. Let \mathcal{N} be the null space at the G -symmetry-breaking bifurcation point $(\mathring{\mathbf{u}}(\lambda^c), \lambda^c)$ with $\dim(\mathcal{N}) = N_{\mathcal{N}}$. Then,*

$$E_{,\eta}(\mathbf{T}^{\mathcal{N}}(g)\boldsymbol{\eta}, \Delta\lambda) = \mathbf{T}^{\mathcal{N}}(g)E_{,\eta}(\boldsymbol{\eta}, \Delta\lambda), \quad \forall g \in G, \quad (\text{A.18})$$

where $\mathbf{T}^{\mathcal{N}} \in \mathbb{R}^{N_{\mathcal{N}} \times N_{\mathcal{N}}}$ is the irreducible representation of symmetry group G associated with the null space \mathcal{N} .

Proof. First, recall that the LSK decomposition allows the decomposition of the DOF vector as follows: $\mathbf{u} = \mathring{\mathbf{u}} + \boldsymbol{\xi} + \mathbf{v}$. Let $\mathcal{N} = \text{span}\{\mathbf{u}_1^{\mathcal{N}}, \dots, \mathbf{u}_{N_{\mathcal{N}}}^{\mathcal{N}}\}$ and $\mathcal{N}^\perp = \text{span}\{\mathbf{u}_1^{\mathcal{N}^\perp}, \dots, \mathbf{u}_{(N-N_{\mathcal{N}})}^{\mathcal{N}^\perp}\}$. Therefore, $\boldsymbol{\xi} = \sum_{i=1}^{N_{\mathcal{N}}} \eta_i \mathbf{u}_i^{\mathcal{N}} \in \mathcal{N}$ and $\mathbf{v} = \sum_{i=1}^{(N-N_{\mathcal{N}})} w_i \mathbf{u}_i^{\mathcal{N}^\perp}$.

Define

$$\widehat{\mathcal{E}}(\boldsymbol{\eta}, \mathbf{w}, \Delta\lambda) \equiv \mathcal{E}(\mathring{\mathbf{u}} + \mathbf{P}^\top \boldsymbol{\eta} + \mathbf{Q}^\top \mathbf{w}, \lambda^c + \Delta\lambda), \quad (\text{A.19})$$

where $\boldsymbol{\eta} = \{\eta_1, \eta_2, \dots, \eta_{N_{\mathcal{N}}}\}$, $\mathbf{w} = \{w_1, w_2, \dots, w_{(N-N_{\mathcal{N}})}\}$ and $\mathbf{P} \in \mathbb{R}^{N \times N_{\mathcal{N}}}$, $\mathbf{Q} \in \mathbb{R}^{N \times (N-N_{\mathcal{N}})}$ such that $\mathbf{P} : \mathbb{R}^{N_{\mathcal{N}}} \rightarrow \mathcal{N}$ and $\mathbf{Q} : \mathbb{R}^{(N-N_{\mathcal{N}})} \rightarrow \mathcal{N}^\perp$. Note that $\mathbf{P}^\top \boldsymbol{\eta} = \boldsymbol{\xi}$ and $\mathbf{Q}^\top \mathbf{w} = \mathbf{v}$ so Eq. (A.19) becomes

$$\widehat{\mathcal{E}}(\boldsymbol{\eta}, \mathbf{w}, \Delta\lambda) \equiv \mathcal{E}(\mathring{\mathbf{u}} + \boldsymbol{\xi} + \mathbf{v}, \lambda^c + \Delta\lambda) = \mathcal{E}(\mathbf{u}, \lambda). \quad (\text{A.20})$$

Now, since \mathcal{E} is invariant under G , then it must satisfy $\mathcal{E}(\mathbf{u}, \lambda) = \mathcal{E}(\mathbf{T}(g)\mathbf{u}, \lambda)$, $\forall g \in G$. Therefore, the right hand side of Eq. (A.19) must satisfy

$$\mathcal{E}(\mathring{\mathbf{u}} + \mathbf{P}^\top \boldsymbol{\eta} + \mathbf{Q}^\top \mathbf{w}, \lambda^c + \Delta\lambda) = \mathcal{E}(\mathbf{T}(g)[\mathring{\mathbf{u}} + \mathbf{P}^\top \boldsymbol{\eta} + \mathbf{Q}^\top \mathbf{w}], \lambda^c + \Delta\lambda), \quad (\text{A.21})$$

for every $g \in G$. Now, since $\mathring{\mathbf{u}}$ is G -symmetric by assumption, then $\mathring{\mathbf{u}} = \mathbf{T}(g)\mathring{\mathbf{u}}, \forall g \in G$. Additionally, since \mathcal{N} and \mathcal{N}^\perp are mutually orthogonal then it must be that

$$\mathbf{P}^\top \boldsymbol{\eta} = \mathbf{T}(g)\mathbf{P}^\top \boldsymbol{\eta}, \text{ and } \mathbf{Q}^\top \mathbf{w} = \mathbf{T}(g)\mathbf{Q}^\top \mathbf{w}, \quad (\text{A.22})$$

or

$$\boldsymbol{\eta} = \mathbf{P}\mathbf{T}(g)\mathbf{P}^\top \boldsymbol{\eta}, \text{ and } \mathbf{w} = \mathbf{Q}\mathbf{T}(g)\mathbf{Q}^\top \mathbf{w}, \quad (\text{A.23})$$

since $\mathbf{P}\mathbf{P}^\top = \mathbf{Q}\mathbf{Q}^\top = \mathbf{I}$. Now, because \mathbf{P} is constructed from the null-space associated with the irreducible representation $\mathbf{T}^\mathcal{N}$, it follows that $\mathbf{P}\mathbf{T}(g)\mathbf{P}^\top = \mathbf{T}^\mathcal{N}(g), \forall g \in G$. Furthermore, by Eq. (A.4), $\mathbf{T}^{\mathcal{N}^\perp}$ can be defined such that $\mathbf{T} = \mathbf{T}^\mathcal{N} \oplus \mathbf{T}^{\mathcal{N}^\perp}$. Then, it must be that $\mathbf{Q}\mathbf{T}(g)\mathbf{Q}^\top = \mathbf{T}^{\mathcal{N}^\perp}(g)$, for every $g \in G$. Equation (A.23) then becomes

$$\mathbf{T}^\mathcal{N}(g)\boldsymbol{\eta} = \boldsymbol{\eta} \text{ and } \mathbf{T}^{\mathcal{N}^\perp}(g)\mathbf{w} = \mathbf{w}, \quad (\text{A.24})$$

for every $g \in G, \boldsymbol{\eta} \in \mathcal{N}$, and $\mathbf{w} \in \mathcal{N}^\perp$. This implies that $\widehat{\mathcal{E}}$ is invariant under G as reflected by the following equation:

$$\widehat{\mathcal{E}}(\boldsymbol{\eta}, \mathbf{w}, \Delta\lambda) = \widehat{\mathcal{E}}(\mathbf{T}^\mathcal{N}(g)\boldsymbol{\eta}, \mathbf{T}^{\mathcal{N}^\perp}(g)\mathbf{w}, \Delta\lambda), \quad \forall g \in G. \quad (\text{A.25})$$

Now, recall from the LSK decomposition that the projection of the equilibrium equation onto \mathcal{N}^\perp implicitly defines the function $\mathbf{v}(\boldsymbol{\eta}, \Delta\lambda) \in \mathcal{N}^\perp$ such that $\widehat{\mathcal{E}}_{,\mathbf{w}}(\boldsymbol{\eta}, \mathbf{v}(\boldsymbol{\eta}, \Delta\lambda), \Delta\lambda) = \mathbf{0}$. Now, since $\widehat{\mathcal{E}}_{,\mathbf{w}} \in \mathcal{N}^\perp$, then from Eq. (A.25),

$$\mathbf{T}^{\mathcal{N}^\perp}(g)\widehat{\mathcal{E}}_{,\mathbf{w}}(\boldsymbol{\eta}, \mathbf{w}, \Delta\lambda) = \widehat{\mathcal{E}}_{,\mathbf{w}}(\mathbf{T}^\mathcal{N}(g)\boldsymbol{\eta}, \mathbf{T}^{\mathcal{N}^\perp}(g)\mathbf{w}, \Delta\lambda), \quad \forall g \in G. \quad (\text{A.26})$$

Substituting the function $\mathbf{v}(\boldsymbol{\eta}, \Delta\lambda) \in \mathcal{N}^\perp$ for \mathbf{w} , then Eq. (A.26), becomes

$$\widehat{\mathcal{E}}_{,\mathbf{w}}(\mathbf{T}^\mathcal{N}(g)\boldsymbol{\eta}, \mathbf{T}^{\mathcal{N}^\perp}(g)\mathbf{v}(\boldsymbol{\eta}, \Delta\lambda), \Delta\lambda) = \mathbf{0}, \quad \forall g \in G, \quad (\text{A.27})$$

Equation (A.27) then implicitly defines the function $\mathbf{v}(\mathbf{T}^\mathcal{N}(g)\boldsymbol{\eta}, \Delta\lambda)$ such that

$$\mathbf{v}(\mathbf{T}^\mathcal{N}(g)\boldsymbol{\eta}, \lambda) = \mathbf{T}^{\mathcal{N}^\perp}(g)\mathbf{v}(\boldsymbol{\eta}, \lambda), \quad \forall g \in G. \quad (\text{A.28})$$

The bifurcation equation is defined as $\mathbf{E}(\boldsymbol{\eta}, \Delta\lambda) \equiv \widehat{\mathcal{E}}(\boldsymbol{\eta}, \mathbf{v}(\boldsymbol{\eta}, \Delta\lambda), \Delta\lambda)$.

From Eq. (A.25) and Eq. (A.28), then

$$\mathbf{E}(\boldsymbol{\eta}, \Delta\lambda) = \mathbf{E}(\mathbf{T}^{\mathcal{N}}(g)\boldsymbol{\xi}, \Delta\lambda), \quad (\text{A.29})$$

where $\mathbf{E}(\mathbf{T}^{\mathcal{N}}(g)\boldsymbol{\eta}, \Delta\lambda) \equiv \widehat{\mathcal{E}}(\mathbf{T}^{\mathcal{N}}(g)\boldsymbol{\eta}, \mathbf{T}^{\mathcal{N}\perp}(g)\mathbf{v}(\boldsymbol{\eta}, \Delta\lambda), \Delta\lambda)$. Finally, taking the derivative of Eq. (A.29) with respect to $\boldsymbol{\eta}$, we obtain the desired result:

$$\mathbf{E}_{,\boldsymbol{\eta}}(\mathbf{T}^{\mathcal{N}}(g)\boldsymbol{\eta}, \Delta\lambda) = \mathbf{T}^{\mathcal{N}}(g)\mathbf{E}_{,\boldsymbol{\eta}}(\boldsymbol{\eta}, \Delta\lambda), \quad \forall g \in G, \quad (\text{A.30})$$

and therefore the bifurcation equation is equivariant under the action of G . \square

The equivariance of the bifurcation equations is useful since it implies that any m -th order coefficient in the Taylor expansion of the bifurcation equation is invariant under the action of the m -th order representation. That is if $\mathbf{E} \in \mathbb{R}^{\times_1^m N_{\mathcal{N}}}$ is an m -th order coefficient of the expansion of the bifurcation equation, then \mathbf{E} is invariant under the action of $\mathbf{T}^m(g)$ on $\mathbb{R}^{\times_1^m N_{\mathcal{N}}}$, $\forall g \in G$. Furthermore, recall in the discussion of Prop. 1 that any tensor that is invariant under the action of a representation of G must belong to the invariant subspace associated with the identity irreducible representation. This means that if we identify the invariant subspace associated with the identity irreducible representation for the representation \mathbf{T}^m , then the m -th order tensor in the expansion of the bifurcation equation must be an element of this subspace. Furthermore, if the invariant subspace is of dimension N_m , then we will have found the m -th order coefficients in the expansion of the bifurcation equation in terms of N_m unknowns instead of $(N_{\mathcal{N}})^m$ entries of an m -th order tensor (where $N_{\mathcal{N}}$ is the dimension of the irreducible representation associated with the null space). Therefore, the determination of the form of an m -th order coefficient in the bifurcation equation expansion is facilitated by finding the invariant subspace associated with the identity irreducible representation for the m -th order representation \mathbf{T}^m .

The next result shows that a representation \mathbf{T}^m on $\mathbb{R}^{\times_1^m N_{\mathcal{N}}}$ can be constructed given a representation $\mathbf{T}^{\mathcal{N}}$ on $W^{\mathcal{N}}$.

Proposition 3. *Suppose that $\{\mathbf{T}_i | i = 1, \dots, m\}$ are representations of G on the spaces $\{V_i | i = 1, \dots, m\}$ respectively, then it can be shown that $\bigotimes_{i=1}^m \mathbf{T}_i$ is a representation on $\bigotimes_{i=1}^m V_i$.*

The proof for this proposition can be found in *Miller (1972)*. The theorem can be applied in this manner: since an m -th order tensor is an element of $\bigotimes_{i=1}^m \mathbb{R}^{N_{\mathcal{N}}}$, let $V_i = \mathbb{R}^{N_{\mathcal{N}}}$ for $i = 1, 2, \dots, m$ and let $\mathbf{T}^{\mathcal{N}}$ be a representation of G on $\mathbb{R}^{N_{\mathcal{N}}}$. Therefore, applying the proposition we obtain that \mathbf{T}^m , the m -th order representation, is given by $\mathbf{T}^m = \bigotimes_{i=1}^m \mathbf{T}^{\mathcal{N}}$.

One final result from group representation theory that will be discussed here is regarding the computation of bifurcating paths at a symmetry-breaking bifurcation point. While the LSK decomposition can always be used to obtain an expression for a bifurcating path, the following result will greatly simplify the computation of some bifurcating paths.

Lemma 1. *Equivariant branching lemma: Let a system be described by the equilibrium equations $\mathbf{F}(\mathbf{u}, \lambda) = \mathbf{0}$. Let Σ_G be a G -symmetric path and $(\mathbf{u}^c, \lambda^c)$ be a G -symmetry-breaking bifurcation point. If there exists a proper subgroup $H \subset G$ such that*

$$\dim(\mathcal{N}^c) = 1, \tag{A.31}$$

where \mathcal{N}^c is the null space of $\nabla_{\mathbf{w}} \mathbf{F}_H(\mathbf{w}^c, \lambda^c)$ and $\mathbf{F}_H(\mathbf{w}^c, \lambda^c) \equiv (\Psi_H)^\top \mathbf{F}(\Psi_H \mathbf{w}^c, \lambda^c)$ is the H -symmetry reduced set of equations as in Eq. (A.17). Then there is an H -symmetric solution path Σ_H of $\mathbf{F}(\mathbf{u}, \lambda) = \mathbf{0}$ that bifurcates from Σ_G at $(\mathbf{u}^c, \lambda^c) = (\Psi_H \mathbf{w}^c, \lambda^c)$.

The equivariant branching lemma not only gives a criteria for the existence of a bifurcating path but also characterizes the nature of the bifurcating path in terms of its symmetry. Note that the while this lemma guarantees that certain solutions will bifurcate from the symmetry-breaking bifurcation point, other solutions not guaranteed by this lemma may also bifurcate at the bifurcation point. The proof for the equivariant branching lemma can be found in various resources such as *Healey* (1988).

Appendix B

A Branch-following and bifurcation study of a one dimensional phase transforming material model

The purpose of this appendix is to provide an explicit example of the branch-following and bifurcation (BFB) technique. Although the presentation of the theory behind the BFB method exists (e.g., Ch. 2, *Triantafyllidis and Peek (1992)*) a detailed example of the application of the method is heretofore lacking in the aforementioned works. It is the objective of this work to present such details, and in doing so, highlight some of the subtleties that may arise in the application of the method that may not be clear upon a first introduction to the method. These subtleties may prove to be important in BFB studies involving period-extending bifurcations such as the one performed in Ch. 3. In this appendix, the BFB method is demonstrated through the use of a prototypical example—a one dimensional model of a phase transforming material. The one dimensional model was developed by *Dobson et al. (2008)* and was used in the quasicontinuum study of an infinite bi-atomic crystal subject to axial loads. Furthermore, *Dobson et al. (2008)* augmented the quasicontinuum method with a technique referred to as the *Cascading Cauchy-Born Kinematics (CCB)* that is used to detect lattice period extensions and identify the minimum required periodic cell size. Using the CCB technique, *Dobson et al. (2008)* identified period extending bifurcations corresponding to deformations of the crystal with an essential unit cell consisting of up to eight atoms. In this work, we extend the investigation started by *Dobson et al. (2008)* by obtaining all period-extending bifurcations from non-essential unit cells of up to 24 atoms.

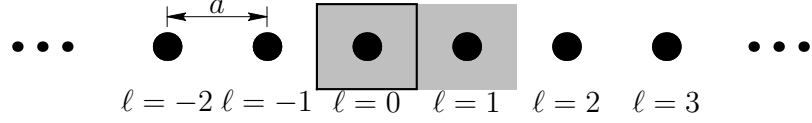


Figure B.1: A one dimensional lattice with reference lattice parameter a . The region enclosed by the black square is an example of an essential unit cell. The gray region is an example of a non-essential unit cell comprising of two atoms.

Note that the theoretical details of the BFB method will not be presented in this document. It will be assumed here that the reader is familiar with the method. For a more thorough description of the technique as well as a more detailed description of equilibrium solutions, critical points and bifurcation points, refer to *Jusuf (2010)*. This appendix will be presented as follows. In Sec. B.1 the details of the model used by *Dobson et al. (2008)* are presented. Here, the description of the one dimensional crystal is explained. The kinematics and equilibrium equations of the problem are also presented in this section. In Sec. B.2 simulation results using a 24 atom non-essential unit cell are presented. The section discusses the primary path and the paths bifurcating from the primary path. The discussion on the bifurcating paths is divided into two subsections which separates the bifurcation points into *general* and *special* symmetry-breaking bifurcation points. The two types of bifurcation points serve to highlight the subtleties that may arise in the application of the BFB method (especially in studies involving non-essential simulated unit cells to capture period-extending bifurcations¹). Sec. B.3 completes the bifurcation diagram by presenting the results of simulations using 10 and 14 atom non-essential unit cells. Unlike the previous (24 atom non-essential unit cell) simulation, the results in these simulations are obtained using a technique discussed in *Kiyohiro et al. (1991)*. The appendix concludes in Sec. B.4 which provides a summary of the work and draws conclusions based on its results.

B.1 A one dimensional perfect crystal description

An infinite one-dimensional crystal can be described by a *lattice* \mathcal{L} which defines the set of positions of the atoms in the crystal as follows

$$\mathcal{L} = \{X = \ell a \mid \ell \in \mathbb{Z}\}, \quad (\text{B.1})$$

where a is the reference lattice parameter (see Fig. B.1). The crystal can also be described

¹More generically the subtleties referred to here arises in BFB studies of systems with n -fold symmetries where $n > 5$.

as a multi-lattice \mathcal{L}^N which is defined as

$$\mathcal{L}^N = \{X_{(\ell,\alpha)} = \ell a_N + \Delta_\alpha a_N \mid \ell \in \mathbb{Z}\}. \quad (\text{B.2})$$

where $a_N = Na$ is the lattice parameter of the multi-lattice. $N \in \mathbb{N}$ is the number of *basis atoms* required in the multi-lattice description and $\Delta_\alpha \in [0, 1)$ gives the fractional positions of the basis atoms for atoms $\alpha = 0, 1, \dots, N-1$. Note that the lattice \mathcal{L}^N contains N times the number of basis atoms in the simple lattice \mathcal{L} (the basis atom for the simple lattice is the single atom $\alpha = 0$ with fractional position $\Delta_0 = 0$). One can view the concepts of lattices and multi-lattices as different labeling schemes to index the atoms in the crystal. Seen in this manner, the fundamental difference between the two descriptions is the periodicity \mathcal{P} between the first atoms ($\alpha = 0$) in the index. For example, the lattices \mathcal{L} and \mathcal{L}^N have periodicities equal to a and a_N respectively.

Related concepts in the description of crystals are the concepts of *essential* and *non-essential* description of a crystal. An essential description of a crystal is defined such that the periodicity \mathcal{P} of the lattice is commensurate with the physical periodicity of the crystal (otherwise it is a non-essential description). A unit cell² in the essential description of a crystal (referred to as the *essential unit cell*) always contains the smallest number of atoms necessary to define the crystal.

For computational efficiency it is often desirable to use the essential unit cell as the *representative* unit cell in atomistic simulations of crystalline materials. This is because the essential unit cell contains the smallest number of *degrees of freedom* (DOFs) needed in order to describe the associated crystal configuration. However, in order to capture translation-symmetry breaking (i.e., period-extending) phase transformations it is necessary to use a non-essential description of the crystal. The additional DOFs introduced in the non-essential description allow for more complicated deformations (many of which are relevant to phase transformations) to be captured. Since the objective of this work is to extend the BFB study by *Dobson et al. (2008)* by obtaining additional period-extending bifurcations, non-essential unit cells have been used for the atomistic simulations in this work.

²A unit cell is the region enclosed by the lattice parameter that can reproduce the crystal by translations of integer multiples of the lattice parameter.

B.1.1 Internal potential energy, Cauchy-Born kinematics, and equilibrium equations in a one dimensional biatomic crystal model by *Dobson et al. (2008)*

The internal potential energy density \mathcal{W} of an infinite biatomic crystal Ω is given by

$$\mathcal{W}(\mathbf{u}_N) = \frac{1}{2a} \sum_{\alpha \in \Omega_N} \sum_{\alpha' \in \Omega} \Phi(\alpha, \alpha', r(\alpha, \alpha')), \quad (\text{B.3})$$

where Ω_N denotes the chosen representative unit cell containing N atoms³ with lattice parameter a . $\mathbf{u}_N = (F, S_0, S_1, \dots, S_{N-1}) \in \mathbb{R}^{N+1}$ are the DOFs associated with the representative unit cell. Here, $N \in \mathbb{N}$ is the number of basis atoms in the representative unit cell⁴. The interaction contribution $\Phi(\alpha, \alpha', r(\alpha, \alpha'))$ is a function of the atoms α and α' and in general depends on the type of atoms in the crystal as well as the distance $r(\alpha, \alpha')$ between the two atoms given by

$$r(\alpha, \alpha') = |x_{(0,\alpha)} - x_{(\ell',\alpha')}|. \quad (\text{B.4})$$

Dobson et al. (2008) uses the Cauchy-Born (CB) kinematics which allows for the lattice to deform uniformly while the basis atoms are free to displace relative to one another during the deformation. The CB kinematics condition is given by the following equation that relates the current position $X_{(\ell,\alpha)}$ of atom α in unit cell ℓ to the uniform deformation F and shift S_α of atom α

$$x_{(\ell,\alpha)} = F(X_{(\ell,\alpha)} + S_\alpha), \quad (\text{B.5})$$

and hence

$$x_{(0,\alpha)} = F(X_{(0,\alpha)} + S_\alpha) = F(\Delta_\alpha a + S_\alpha), \quad (\text{B.6})$$

$$x_{(\ell',\alpha')} = F(X_{(\ell',\alpha')} + S_{\alpha'}) = F(\ell' a + \Delta_{\alpha'} a + S_{\alpha'}). \quad (\text{B.7})$$

Although in principle the sum in Eq. (B.3) is taken over all atoms in the crystal it is often the case that a cutoff distance r_{cut} is used such that

$$\Phi(\alpha, \alpha', r(\alpha, \alpha')) = 0, \quad \forall r > r_{cut}. \quad (\text{B.8})$$

³As mentioned in the preceding section, this work uses a non-essential representative unit cell in order to capture period-extending bifurcation points.

⁴For the biatomic crystal described in $N \geq 2$.

	ϵ	σ	r_{cut}/a_0
$(\alpha, \alpha') = (A, A)$	1.0	1.0	2.0
$(\alpha, \alpha') = (A, B)$	0.1	0.475	2.0
$(\alpha, \alpha') = (B, B)$	1.0	0.85	2.0

Table B.1: Parameters used in the LJ potential of *Dobson et al.* (2008). r_{cut} is the cutoff distance and a is the lattice parameter of the essential unit cell.

The atomic interaction used by *Dobson et al.* (2008) is a Lennard-Jones (LJ) potential which has the form

$$\Phi(\alpha, \alpha', r(\alpha, \alpha')) = 4\epsilon(\alpha, \alpha') \left[\left(\frac{\sigma(\alpha, \alpha')}{r} \right)^{12} - \left(\frac{\sigma(\alpha, \alpha')}{r} \right)^6 \right], \quad (\text{B.9})$$

where $\sigma(\alpha, \alpha')$ and $\epsilon(\alpha, \alpha')$ depends on the type of atoms α and α' and r is defined as in Eq. (B.4). The parameters $\sigma(\alpha, \alpha')$ and $\epsilon(\alpha, \alpha')$ for the biatomic crystal are given in Tbl. B.1 where the two atoms are labeled A and B . The applied load in this case is a tensile load of magnitude⁵ λ . The total potential energy density \mathcal{E} is given by the sum of the total internal potential energy density and the potential due to applied load and is given by

$$\mathcal{E}_{\Omega_N}(\mathbf{u}_N, \lambda) = \mathcal{W}(\mathbf{u}_N) - \lambda F, \quad (\text{B.10})$$

The equilibrium configurations $\mathring{\mathbf{u}}_N(\lambda)$ are extremums of the energy and is given by

$$\mathbf{F}_{\Omega_N}(\mathring{\mathbf{u}}_N(\lambda), \lambda) \equiv \nabla \mathcal{E}_{\Omega_N}(\mathring{\mathbf{u}}_N(\lambda), \lambda) = \mathbf{0}. \quad (\text{B.11})$$

The stability of a configuration is computed using the phonon stability and CB stability criteria discussed in *Elliott et al.* (2006a). Once the equilibrium equations have been obtained, the first step in the BFB method is to compute the equilibrium path associated with a reference configuration (referred to as the *principal path*). In the BFB method this is done using using a *predictor-corrector* method.

⁵The application of the tensile load can be interpreted as a tensile load at infinity. For a more thorough derivation on the application of the load here, refer to *Elliott et al.* (2006b).

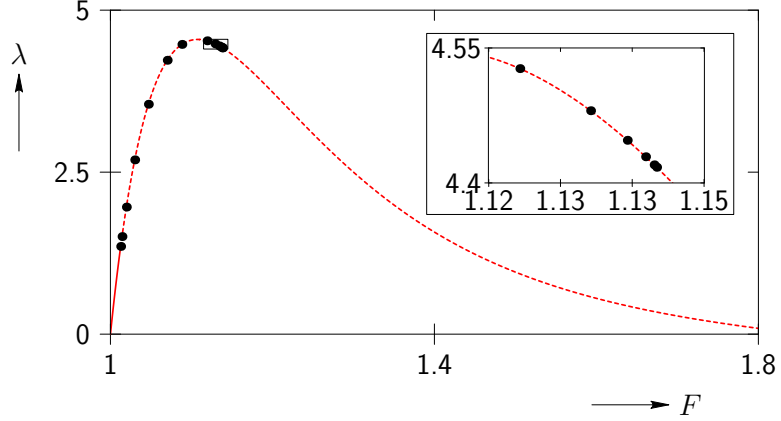


Figure B.2: The 2_m path is shown in red. Solid and dashed lines correspond to stable and unstable regions along the curve. Black dots along the curve indicate symmetry-breaking bifurcation points.

B.2 Simulation results using a 24 atom representative unit cell

For the first set of simulations, a simulated cell consisting of $N = 24$ atoms is chosen. As such, the equilibrium equations are given by

$$\nabla \mathcal{E}_{\Omega_{24}}(\mathbf{u}_{24}, \lambda) \equiv \nabla (\mathcal{W}(\mathbf{u}_{24}) - \lambda F), \quad (\text{B.12})$$

where $\mathbf{u}_{24} = (F, S_0, \dots, S_{23})$. Fig. B.2 shows a result of the BFB study obtained by solving Eq. (B.12). Here, solid and dashed lines correspond to stable and unstable segments, respectively. The configuration at zero load is stable and can be described as a multi-lattice comprising of two atoms in the essential unit cell. This reference configuration is described by one basis atom placed at the lattice site and the second basis atom placed at $0.5a_0$ as shown in Fig. B.3. Hence, the reference configuration is a biatomic crystal Ω with an essential description containing two atoms as shown in Fig. B.3. Equilibria described in this manner is referred to as principal solutions (the path associated with this configuration is the principal path). The reference position of atoms in a crystal described by the principal solution is given by

$$X_{(\ell, \alpha)} = \ell a_0 + \Delta_\alpha, \quad (\text{B.13})$$

where a_0 is the reference lattice parameter⁶ (see Fig. B.3), $\ell \in \mathbb{Z}$, $\alpha = 0, 1$ and $\Delta_0 = 0, \Delta_1 = 0.5$. This structure will be referred to as \mathbf{u}_{24}^{2m} where the subscript denotes the number of basis atoms used in the simulation (and represents the unit cell used in the

⁶Note that the non-essential representative unit cell has a lattice parameter of $a = 12a_0$, $\alpha = 0, \dots, 23$.

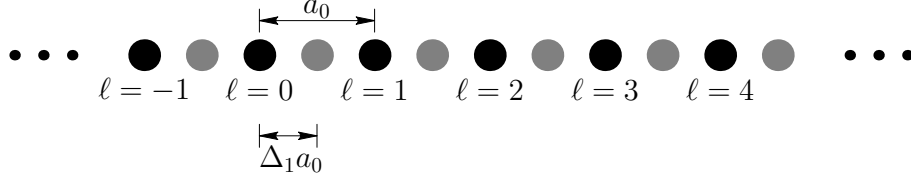


Figure B.3: A segment of a one dimensional biatomic crystal with reference lattice parameter a and fractional positions $\Delta_0 = 0$ and $\Delta_1 = 0.5$.

simulation). The M_m superscript describes the equilibrium structure where M indicates the number of atoms required to describe the essential unit cell and m indicates the presence of a mirror symmetry. An M -atom essential unit cell with no mirror symmetry will be denoted as M_o . This convention will be used to describe the structures obtained in the BFB study discussed later in this document.

As the load is increased the structure continues to stretch uniformly until a critical load is reached where the first bifurcation point is encountered. After this point the \mathbf{u}_{24}^{2m} path becomes unstable. A total of thirteen symmetry-breaking bifurcation points (indicated by black dots in Fig. B.2) are found along the \mathbf{u}_{24}^{2m} path. A *Lyapunov-Schmidt-Koiter (LSK) decomposition and asymptotic bifurcation analysis* is performed (as described in Ch. 2) on each of these points in order to obtain an asymptotic expression for the bifurcating paths. The asymptotic expression can then be used to seed a branch-following method to follow the bifurcating path.

B.2.1 Period-extending bifurcations from \mathbf{u}_{24}^{2m}

In the BFB method, the bifurcating paths are computed by performing the (LSK) decomposition and asymptotic bifurcation analysis at each bifurcation point. The first order terms ($\alpha^{(1)}$ and Λ_1) of the asymptotic expression obtained from the analysis is used to obtain the tangent of the bifurcating path at the bifurcation point. The tangent can then be used to seed the numerical branch-following method.

The set of bifurcation points found along the \mathbf{u}_{24}^{2m} path have here been divided into two categories: (1) general, and (2) special cases depending on what order of the bifurcation equation expansion is needed to determine $\alpha^{(1)}$ and Λ_1 .

B.2.1.1 General symmetry-breaking bifurcation points

General symmetry-breaking bifurcation points (GSBBPs) are symmetry-breaking bifurcation points where the determination of $\alpha^{(1)}$ and Λ_1 requires the third order (or less) equation

in the expansion of the bifurcation equation (i.e., $O(\zeta^n)$, where $n \leq 3$). The null spaces at GSBBPs are either one dimensional or two dimensional.

1. $\dim(\mathcal{N}) = 1$.

The one dimensional null space is the simplest case of obtaining the bifurcating tangent. $\alpha^{(1)}$ and Λ_1 are scalars and can be obtained by solving the $O(\zeta^2)$ equation in the expansion of the bifurcation equation which is generically given by the following equation:

$$\Lambda_1 C \alpha^{(1)} = 0, \quad (\text{B.14})$$

where $C \in \mathbb{R}$ is some constant. Eq. (B.14) implies that $\Lambda_1 = 0$ and $\alpha^{(1)} = \pm 1$. Therefore, if $\mathcal{N} = \text{span}\{\mathbf{V}^c\}$, where \mathbf{V}^c is the eigenvector with eigenvalue equal to zero, then *the tangent of the bifurcating path at the bifurcation point is given by $\pm \mathbf{V}^c$.*

2. $\dim(\mathcal{N}) = 2$.

All GSBBPs with $\dim(\mathcal{N}) = 2$ found along the \mathbf{u}_{24}^{2m} path are symmetric bifurcation points and hence $E_{,\xi\xi\xi} = \mathbf{0}$ (and so $\Lambda_1 = 0$). In fact, it can be shown (see App. (C)) that this is necessarily a result of the symmetry of the problem. Therefore, the first non-trivial order in the expansion of the bifurcation equation is $O(\zeta^3)$ which generically takes the following form:

$$C_1(\alpha_1^{(1)})^3 + C_2(\alpha_1^{(1)})^2\alpha_2^{(1)} + C_3\alpha_1^{(1)}(\alpha_2^{(1)})^2 + C_4(\alpha_1^{(1)})^3 + 3\Lambda_2 C_5\alpha_1^{(1)} = 0, \quad (\text{B.15})$$

$$-C_4(\alpha_1^{(1)})^3 + C_3(\alpha_1^{(1)})^2\alpha_2^{(1)} - C_2\alpha_1^{(1)}(\alpha_2^{(1)})^2 + C_1(\alpha_1^{(1)})^3 + 3\Lambda_2 + C_5\alpha_2^{(1)} = 0, \quad (\text{B.16})$$

where $\alpha^{(1)} = [\alpha_1^{(1)}, \alpha_2^{(1)}]^\top$. Equation (B.15) and Eq. (B.16) together with the condition that $\langle \alpha^{(1)}, \alpha^{(1)} \rangle = 1$ gives three equations to solve for the three unknowns $\alpha_1^{(1)}$, $\alpha_2^{(1)}$ and Λ_2 . A total of eleven paths coming out of seven GSBBPs were found along the \mathbf{u}_{24}^{2m} path. Figure (B.4) shows the plot of the load λ versus uniform deformation gradient F of each path bifurcating from the seven GSBBPs. The same convention is used here where solid lines and dashed lines indicate stable and unstable solutions respectively.

The following list describes the structure corresponding to each path in Fig. B.4: \mathbf{u}_{24}^{2m} (red), \mathbf{u}_{24}^{4m} (dark green), \mathbf{u}_{24}^{6m} (dark blue), \mathbf{u}_{24}^{8m} (gray), \mathbf{u}_{24}^{2o} (purple), \mathbf{u}_{24}^{8m} (light green),

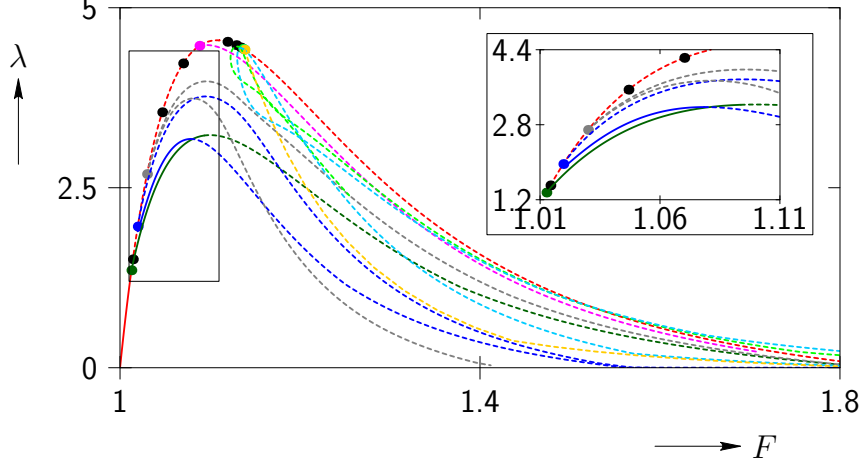


Figure B.4: A figure showing paths bifurcating from GSBBPs. Solid and dashed lines indicate stable and unstable solutions respectively.

\mathbf{u}_{24}^{6m} (light blue), \mathbf{u}_{24}^{4o} (orange). Note that only two stable bifurcating paths (\mathbf{u}_{24}^{4m} and \mathbf{u}_{24}^{6m}) were found along the $2m$ path and correspond to structures describable by a four atom and six atom essential unit cell with mirror symmetry.

B.2.1.2 Special symmetry-breaking bifurcation points

Special symmetry-breaking bifurcation points (SSBBPs) are all symmetric bifurcation points and as such are characterized by $E_{,\xi\xi\xi} = \mathbf{0}$ which implies that $\Lambda_1 = 0$. However, unlike GSBBPs, *SSBBPs require greater than $O(\zeta^3)$ equation in the expansion of the bifurcation equation* (i.e., $O(\zeta^n)$, where $n > 3$) in order to determine $\alpha^{(1)}$. These type of bifurcation points characterize the remaining bifurcations that have not yet been computed in the previous calculation (shown as black dots in Fig. B.4). SSBBPs involve subtle calculations in order to determine the bifurcating tangent. These calculations are presented here in order to highlight some of the subtleties that may arise in the application of the BFB method. These subtleties may prove to be important in BFB studies involving period-extending bifurcations. To illustrate the calculation involved for SSBBPs, consider the bifurcation point shown in plum in Fig. B.5. Recall that the $O(\zeta^3)$ term for a symmetric bifurcation point (see Ch. 2) is given by

$$E_{,\xi\xi\xi\xi} \alpha^{(1)} \alpha^{(1)} \alpha^{(1)} + 3\Lambda_2 E_{,\xi\xi\lambda} \alpha^{(1)} = \mathbf{0}. \quad (\text{B.17})$$

The null space at SSBBPs are all associated with a two-dimensional irreducible representation and as such, $\dim(\mathcal{N}) = 2$. At this point, we can take advantage of symmetry and

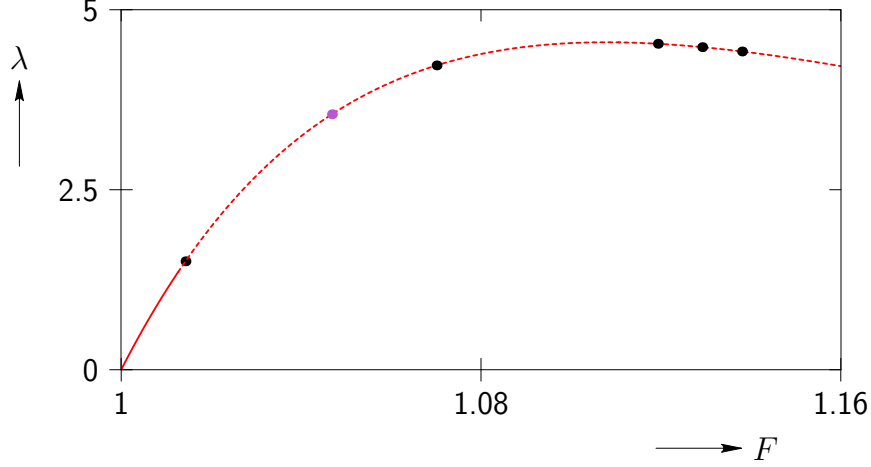


Figure B.5: A figure showing six SSBBPs found along the \mathbf{u}_{24}^{2m} path. The example calculation detailed in this section relates to the SSBBP shown in plum. The remaining five SSBBP are shown in black.

obtain the forms of $E_{,\xi\xi\xi\xi}$ and $E_{,\xi\xi\lambda}$ which is given by the following⁷:

1. $E_{,ij\lambda} = a\delta_{ij}$, $a \in \mathbb{R}$, $i = 1, 2$,
2. $E_{,1111} = E_{,2222} = 3E_{,1122} = b$, $b \in \mathbb{R}$, and other terms are equal to zero,

where $E_{,ijkl} \equiv \frac{\partial^4 E}{\partial \xi_i \partial \xi_j \partial \xi_k \partial \xi_l}$ and $E_{,ij\lambda} \equiv \frac{\partial^3 E}{\partial \xi_i \partial \xi_j \partial \lambda}$. Substituting the results from symmetry, Eq. (B.17) is given by the following equation in component form:

$$b(\alpha_1^{(1)})^3 + b\alpha_1^{(1)}(\alpha_2^{(1)})^2 + 3a\Lambda_2\alpha_1^{(1)} = 0, \quad (\text{B.18})$$

$$b(\alpha_1^{(1)})^2\alpha_2^{(1)} + b(\alpha_1^{(1)})^3 + 3a\Lambda_2\alpha_2^{(1)} = 0, \quad (\text{B.19})$$

where $\boldsymbol{\alpha}^{(1)} = [\alpha_1^{(1)}, \alpha_2^{(1)}]^\top$, $E_{,ijkl} = \frac{\partial^4 E}{\partial \xi_i \partial \xi_j \partial \xi_k \partial \xi_l}$ and $E_{,ij\lambda} = \frac{\partial^3 E}{\partial \xi_i \partial \xi_j \partial \lambda}$. Since $\|\boldsymbol{\alpha}^{(1)}\| = 1$ and assuming $\alpha_1^{(1)}, \alpha_2^{(1)} \neq 0$, Eq. (B.18) and Eq. (B.19) can be divided through by $\alpha_1^{(1)}$ and $\alpha_2^{(1)}$ respectively to obtain the same equation given by the following:

$$b \left(\alpha_1^{(1)2} + (\alpha_1^{(1)})^2 \right) + 3a\Lambda_2 = 0. \quad (\text{B.20})$$

Equation (B.20) is characteristic feature of the $O(\zeta^3)$ term in the expansion of the bifurcation equation is typical for a SSBBP. Note that equation (B.20) implies that an infinite

⁷Note that the coefficients can also be explicitly calculated. However, the presence of symmetry allows for the determination of the coefficients with less computational cost. An example of the process of determining the coefficients in the expansion of the bifurcation equation is detailed in App. C.

number of $\boldsymbol{\alpha}^{(1)}$ such that $\|\boldsymbol{\alpha}^{(1)}\| = \langle \boldsymbol{\alpha}^{(1)}, \boldsymbol{\alpha}^{(1)} \rangle^{\frac{1}{2}} = 1$, is a solution. As such, it is necessary to go to $O(\zeta^4)$ in order to uniquely determine $\boldsymbol{\alpha}^{(1)}$. However, $O(\zeta^3)$ can be used to solve for $\Lambda_2 = -b/3a$. The $O(\zeta^i)$ terms for $i \geq 4$ are given by $\mathbf{B}\boldsymbol{\alpha}^{(i-2)} = \mathbf{b}^{(i)}$, $\forall i \geq 4$, where $\mathbf{B} = \mathcal{E}_{,\xi\xi\xi\xi} \boldsymbol{\alpha}^{(1)} \boldsymbol{\alpha}^{(1)} + \Lambda_2 \mathcal{E}_{,\xi\xi\xi\lambda}$. In general, $\det(\mathbf{B}) \neq 0$ but note in this case, that \mathbf{B} has the following form:

$$\mathbf{B} = \begin{bmatrix} b(\alpha_1^{(1)})^2 + \frac{b}{3}(\alpha_1^{(1)})^2 + a\Lambda_2 & \frac{2b}{3}\alpha_1^{(1)}\alpha_2^{(1)} \\ \frac{2b}{3}\alpha_1^{(1)}\alpha_2^{(1)} & \frac{b}{3}(\alpha_1^{(1)})^2 + b(\alpha_2^{(1)})^2 + a\Lambda_2 \end{bmatrix}, \quad (\text{B.21})$$

where $\Lambda_2 = -b/3a$. It is easily verified that for this particular \mathbf{B} that $\det(\mathbf{B}) = 0$ with a null vector given by

$$\boldsymbol{\alpha}^{(1)\perp} = [-\alpha_1^{(2)}, \alpha_1^{(1)}]^\top. \quad (\text{B.22})$$

Additionally, $\mathbf{b}^{(i)}$ depends on Λ_j and $\boldsymbol{\alpha}^{(j)}$, for⁸ $j < i$. For the SSBBP in the current calculation $\mathbf{b}^{(4)}$ is given by

$$\mathbf{b}^{(4)} = -\frac{2}{3}\Lambda_3 \mathcal{E}_{,\xi\xi\xi\lambda} \boldsymbol{\alpha}^{(1)} = -\frac{2}{3}a\Lambda_3 \boldsymbol{\alpha}^{(1)}. \quad (\text{B.23})$$

At this point, we recall the following theorem:

Theorem 1. Solvability Theorem (The Fredholm alternative). Given $\mathbf{b} \in \mathbb{R}^N$, the equation

$$\mathbf{B}\boldsymbol{\alpha} = \mathbf{b} \quad (\text{B.24})$$

is solvable for $\boldsymbol{\alpha} \in \mathbb{R}^N$ if and only if $\mathbf{b} \perp \ker(\mathbf{B})$, where $\ker(\mathbf{B}) = \{\mathbf{y} \in \mathbb{R}^N \mid \mathbf{B}\mathbf{y} = \mathbf{0}\}$.

Applying the Fredholm alternative, we observe that

$$\langle \mathbf{b}^{(4)}, \boldsymbol{\alpha}^{(1)\perp} \rangle = \frac{-2\Lambda_3 a}{3} \begin{bmatrix} \alpha_1^{(1)} \\ \alpha_2^{(1)} \end{bmatrix} \cdot \begin{bmatrix} -\alpha_2^{(1)} \\ \alpha_1^{(1)} \end{bmatrix} = 0 \Rightarrow \mathbf{b}^{(4)} \perp \ker(\mathbf{B}), \quad (\text{B.25})$$

and therefore $\mathbf{B}\boldsymbol{\alpha}^{(2)} = \mathbf{b}^{(4)}$ is in fact solvable. However, since $\det(\mathbf{B}) = 0$, this implies that there are an infinite number of solutions $\boldsymbol{\alpha}^{(2)}$ such that $\mathbf{B}\boldsymbol{\alpha}^{(2)} = \mathbf{b}^{(4)}$.

Now, in order to determine the form of $\boldsymbol{\alpha}^{(2)}$, let $\boldsymbol{\alpha}^{(2)}$ be written as

$$\boldsymbol{\alpha}^{(2)} = c\boldsymbol{\alpha}^{(1)} + d\boldsymbol{\alpha}^{(1)\perp}, \quad c, d \in \mathbb{R}. \quad (\text{B.26})$$

⁸In the case of $O(\zeta^4)$, $\mathbf{b}^{(4)}$ is given by $\mathbf{b}^{(4)} = -\frac{1}{6}[\mathbf{E}_{,\xi\xi\xi\xi\xi} \boldsymbol{\alpha}^{(1)} \boldsymbol{\alpha}^{(1)} \boldsymbol{\alpha}^{(1)} \boldsymbol{\alpha}^{(1)} + 6\Lambda_2 \mathbf{E}_{,\xi\xi\xi\lambda} \boldsymbol{\alpha}^{(1)} \boldsymbol{\alpha}^{(1)} + 4\Lambda_3 \mathbf{E}_{,\xi\xi\lambda} \boldsymbol{\alpha}^{(1)}]$.

Note that the decomposition of $\boldsymbol{\alpha}^{(2)}$ as in Eq. (B.26) is possible here since \mathcal{N} is two dimensional (i.e., $\mathcal{N} = \text{span}\{\boldsymbol{\alpha}^{(1)}, \boldsymbol{\alpha}^{(1)\perp}\}$). Substituting this form of $\boldsymbol{\alpha}^{(2)}$ into $\mathbf{B}\boldsymbol{\alpha}^{(2)} = \mathbf{b}^{(4)}$, we obtain the following:

$$\mathbf{B}\boldsymbol{\alpha}^{(2)} = \mathbf{B}(c\boldsymbol{\alpha}^{(1)} + d\boldsymbol{\alpha}^{(1)\perp}) = \mathbf{b}^{(4)} = \frac{-2a\Lambda_3}{3}\boldsymbol{\alpha}^{(1)} \quad (\text{B.27})$$

$$= c\mathbf{B}\boldsymbol{\alpha}^{(1)} = \frac{-2a\Lambda_3}{3}\boldsymbol{\alpha}^{(1)}, \quad (\text{B.28})$$

since $\mathbf{B}\boldsymbol{\alpha}^{(1)\perp} = \mathbf{0}$ (by Eq. (B.22)). Dotting Eq. (B.28) with $\boldsymbol{\alpha}^{(1)}$ and noting that $\boldsymbol{\alpha}^{(1)} \cdot \boldsymbol{\alpha}^{(1)} = 1$, we see that

$$c = \left(\frac{-2a\Lambda_3}{3}\right) / (\mathbf{B}\boldsymbol{\alpha}^{(1)} \cdot \boldsymbol{\alpha}^{(1)}). \quad (\text{B.29})$$

Additionally, since $\langle \boldsymbol{\alpha}^{(j)}, \boldsymbol{\alpha}^{(1)} \rangle = 0, \forall j \geq 2$ then

$$0 = \langle \boldsymbol{\alpha}^{(2)}, \boldsymbol{\alpha}^{(1)} \rangle = \langle c\boldsymbol{\alpha}^{(1)} + d\boldsymbol{\alpha}^{(1)\perp}, \boldsymbol{\alpha}^{(1)} \rangle \quad (\text{B.30})$$

$$= c \langle \boldsymbol{\alpha}^{(1)}, \boldsymbol{\alpha}^{(1)} \rangle = c, \quad (\text{B.31})$$

and so $c = 0$ and $\boldsymbol{\alpha}^{(2)} = d\boldsymbol{\alpha}^{(1)\perp} = d[-\alpha_2^{(1)}, \alpha_1^{(1)}]^\top$. Substituting Eq. (B.29) for c then

$$0 = c = \left(\frac{-2a\Lambda_3}{3}\right) / (\mathbf{B}\boldsymbol{\alpha}^{(1)} \cdot \boldsymbol{\alpha}^{(1)}) \quad (\text{B.32})$$

$$= \frac{-2a\Lambda_3}{3} = \Lambda_3, \quad (\text{B.33})$$

and therefore $\Lambda_3 = 0$. At this point, what has been determined is the following: $\Lambda_3 = 0$ and $\boldsymbol{\alpha}^{(2)} = d\boldsymbol{\alpha}^{(1)\perp}$. However, note that $\boldsymbol{\alpha}^{(1)}$ has not yet been determined. Therefore, it is necessary to move on to $O(\zeta^5)$ given by $\mathbf{B}\boldsymbol{\alpha}^{(3)} = \mathbf{b}^{(5)}$ where \mathbf{B} is given by Eq. (B.21) and

$$\begin{aligned} \mathbf{b}^{(5)} = & \frac{-1}{10} [\mathbf{E}_{,\xi\xi\xi\xi\xi\xi} \boldsymbol{\alpha}^{(1)} \boldsymbol{\alpha}^{(1)} \boldsymbol{\alpha}^{(1)} \boldsymbol{\alpha}^{(1)} \boldsymbol{\alpha}^{(1)} + \mathbf{E}_{,\xi\xi\xi\xi} (15\boldsymbol{\alpha}^{(1)} \boldsymbol{\alpha}^{(2)} \boldsymbol{\alpha}^{(2)} \\ & + 10\Lambda_2 \boldsymbol{\alpha}^{(1)} \boldsymbol{\alpha}^{(1)} \boldsymbol{\alpha}^{(1)}) + 15\Lambda_2^2 \mathbf{E}_{,\xi\xi\lambda\lambda} \boldsymbol{\alpha}^{(1)} + 5\Lambda_4 \mathbf{E}_{,\xi\xi\lambda} \boldsymbol{\alpha}^{(1)}]. \end{aligned} \quad (\text{B.34})$$

Again, the forms of $\mathbf{E}_{,\xi\xi\lambda\lambda}$ and $\mathbf{E}_{,\xi\xi\xi\xi\xi}$ can be obtained from symmetry techniques. Applying the Fredholm alternative for the equation $\mathbf{B}\boldsymbol{\alpha}^{(3)} = \mathbf{b}^{(5)}$, we find that $\mathbf{b}^{(5)} \perp \ker(\mathbf{B})$ if and only if $\boldsymbol{\alpha}^{(1)}$ is equal to either of the following six vectors:

$$\pm \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \pm \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \pm \frac{1}{2} \begin{bmatrix} 1 \\ \sqrt{3} \end{bmatrix}, \pm \frac{1}{2} \begin{bmatrix} \sqrt{3} \\ 1 \end{bmatrix}, \pm \frac{1}{2} \begin{bmatrix} -1 \\ \sqrt{3} \end{bmatrix}, \text{ or } \pm \frac{1}{2} \begin{bmatrix} \sqrt{3} \\ -1 \end{bmatrix}. \quad (\text{B.35})$$

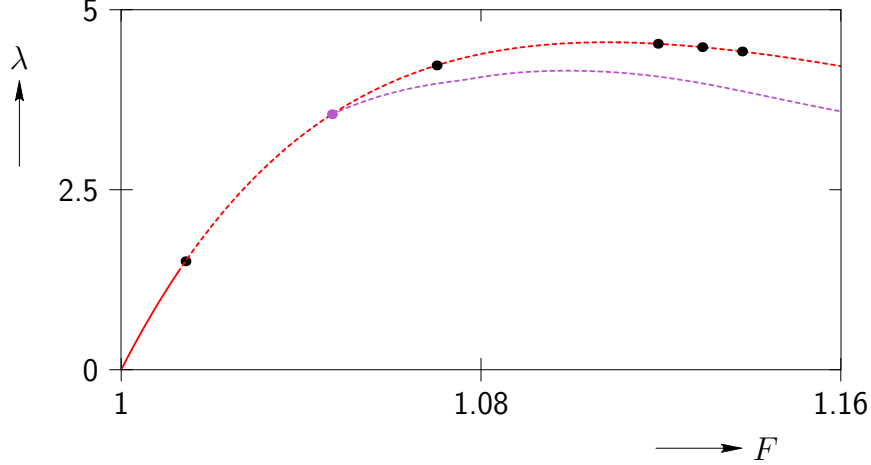


Figure B.6: A figure showing a path bifurcating away from a SSBBP (shown in plum).

Therefore, out of an infinite number of solutions ($\alpha^{(1)}$ such that $\|\alpha^{(1)}\| = 1$), the Fredholm alternative at this order ζ in the expansion of the bifurcation equation gives a condition by which $\alpha^{(1)}$ is determined to be either of the six vectors in Eq. (B.35). In addition, by applying the same process as in the determination of $\alpha^{(2)}$ and Λ_3 (in $O(\zeta^4)$), we find that $\alpha^{(3)} \propto \alpha^{(1)\perp}$ and $\Lambda_4 = k_1$, for some constant $k_1 \in \mathbb{R}$.

Even though at this point $\alpha^{(1)}$ has been determined, there remains the question as to what $\alpha^{(2)}$ is equal to. That is, what is $d \in \mathbb{R}$ such that $\alpha^{(2)} = d\alpha^{(1)\perp} = d[-\alpha_2^{(1)}, \alpha_1^{(1)}]^\top$. The next order equation in the expansion of the bifurcation equation is needed in order to determine the answer. It can be shown that at $O(\zeta^6)$ and with any of the six choices of $\alpha^{(1)}$ in Eq. (B.35) (and using the Fredholm alternative) that $d = 0$, $\Lambda_5 = k_2$, $k_2 \in \mathbb{R}$ and $\alpha^{(4)} \propto \alpha^{(1)\perp}$. This process continues indefinitely and so in general, at any $O(\zeta^i)$, $i \geq 7$, the following is determined: $\alpha^{(i-4)} = \mathbf{0}$, $\Lambda_{i-1} = k_{i-4} \in \mathbb{R}$ and that $\alpha^{(i-2)} \propto \alpha^{(1)\perp}$.

Note that a path can be computed with each choice of $\alpha^{(1)}$. However, all the paths associated with any of the $\alpha^{(1)}$'s are symmetry-related which means that any of the paths can be generated by a symmetry operation of a single path. This means that all the paths are physically equivalent and in this sense, only one “unique” path exists outside of a symmetry related difference. The path bifurcating away from this bifurcation point with $\alpha^{(1)} = [1, 0]^\top$ is shown in Fig. B.6. The following list summarizes the steps involved at any $O(\zeta^i)$ for this bifurcation point:

- $O(\zeta^3)$: $b \left((\alpha_1^{(1)})^2 + (\alpha_2^{(1)})^2 \right) + 3\Lambda_2 a = 0$
 $\Rightarrow \|\alpha^{(1)}\| = 1$ and $\Lambda_2 = -b/3a$.

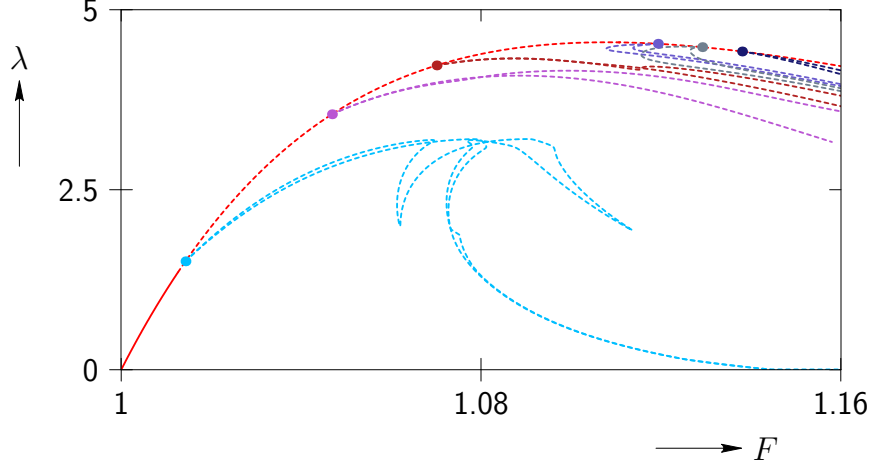


Figure B.7: All paths bifurcating from SSBBPs along the \mathbf{u}_{24}^{2m} path (shown in red). Note that none of these paths were found to be stable.

- $O(\zeta^4)$: $\mathbf{b}^{(4)} \propto \Lambda_3 \boldsymbol{\alpha}^{(1)} \Rightarrow$ solvable by Fredholm alternative.

$$\left. \begin{array}{l} \mathbf{B}\boldsymbol{\alpha}^{(2)} = \mathbf{b}^{(4)} \\ \langle \boldsymbol{\alpha}^{(2)}, \boldsymbol{\alpha}^{(1)} \rangle = 0 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \Lambda_3 = 0 \\ \boldsymbol{\alpha}^{(2)} \propto \boldsymbol{\alpha}^{(1)\perp} \end{array} \right.$$

- $O(\zeta^5)$: Solvable if $\mathbf{b}^{(5)} \perp \ker \mathbf{B} \Rightarrow \boldsymbol{\alpha}^{(1)}$ given by Eq. (B.35).

$$\left. \begin{array}{l} \mathbf{B}\boldsymbol{\alpha}^{(3)} = \mathbf{b}^{(5)} \\ \langle \boldsymbol{\alpha}^{(3)}, \boldsymbol{\alpha}^{(1)} \rangle = 0 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \Lambda_4 = k_1 \in \mathbb{R} \\ \boldsymbol{\alpha}^{(3)} \propto \boldsymbol{\alpha}^{(1)\perp} \end{array} \right.$$

- $O(\zeta^6)$: Solvable if $\mathbf{b}^{(6)} \perp \ker \mathbf{B} \Rightarrow \boldsymbol{\alpha}^{(2)} = \mathbf{0}$.

$$\left. \begin{array}{l} \mathbf{B}\boldsymbol{\alpha}^{(4)} = \mathbf{b}^{(6)} \\ \langle \boldsymbol{\alpha}^{(4)}, \boldsymbol{\alpha}^{(1)} \rangle = 0 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} \Lambda_5 = k_2 \in \mathbb{R} \\ \boldsymbol{\alpha}^{(4)} \propto \boldsymbol{\alpha}^{(1)\perp} \end{array} \right.$$

- $O(\zeta^i), i \geq 7$: $\mathbf{B}\boldsymbol{\alpha}^{(i-2)} = \mathbf{b}^{(i)} \Rightarrow \left\{ \begin{array}{l} \boldsymbol{\alpha}^{(i-4)} = \mathbf{0}, \\ \Lambda_{i-1} = \# \in \mathbb{R} \\ \boldsymbol{\alpha}^{(i-2)} \propto \boldsymbol{\alpha}^{(1)\perp} \end{array} \right.$

The above calculation gives an example of a typical calculation involved in obtaining an asymptotic expression for the bifurcating path from a SSBBP. As can be seen, the process involves many steps with many subtleties at every step in addition to the use of fairly sophisticated group-theoretic techniques. The remaining paths bifurcating from the six SSBBPs is shown in Fig. B.7 where the configuration of each bifurcating path is given in the following list: 24_m (sky blue), 12_m (plum), Brick: $24_m, 24_m$ (slate blue), 12_m (slate gray) and 24_m (midnight blue). Note that no more stable paths are present.

B.3 Simulation results using a 10 and 14 atom representative unit cell

In the previous section, paths corresponding to a structure with a 4, 6, 8, 12, and 24-atom essential unit cell were found bifurcating from a \mathbf{u}_{24}^{2m} path. These paths were discovered by solving the equilibrium equations using a simulated cell consisting of a 24 atom non-essential unit cell ($\nabla \mathcal{E}_{\Omega_{24}} = \mathbf{0}$). Using this non-essential unit cell only paths corresponding to essential unit cells that are factors of 24 (and ≥ 2 since the crystal is biatomic) can be computed. Therefore to complete the bifurcation diagram and obtain structures with 10 and 14 atom essential unit cells, an appropriate simulated cell must be used to define the equilibrium equations to be solved ($\mathcal{E}_{\Omega_{10}}$ and $\mathcal{E}_{\Omega_{14}}$). Using the respective equilibrium equations, bifurcation points corresponding to structures with an essential description of 10 and 14 atoms is then detected. At this point, the procedure to compute the bifurcating paths from these simulated cells can proceed in much the same way as is the case using Ω_{24} described above. However, as can be seen in the previous section, the subtleties involved in the computation makes the process rather difficult and time consuming. Fortunately, there are better methods to obtain the symmetry-breaking bifurcation paths out of the principal path. This is because one important feature about the reference configuration in an Ω_N representative unit cell is that the energy \mathcal{E}_{Ω_N} defined in this manner exhibits the symmetry of the dihedral group of order $n = N/2$; denoted by D_n . This allows for the determination of the bifurcating tangent using a technique found in *Kiyohiro et al.* (1991). The following summarizes the results from that work.

B.3.1 Bifurcations from a D_n -symmetric path

The main tools that were used by *Kiyohiro et al.* (1991) is an LSK asymptotic bifurcation analysis coupled with the equivariant branching lemma. This technique applies to all problems that exhibit the D_n symmetry such as the \mathbf{u}_N^{2m} configuration described in this work and reflected in the following equation

$$\mathcal{E}_{\Omega_N}(\mathbf{T}(g)\mathbf{u}_N^{2m}, \lambda) = \mathcal{E}_{\Omega_N}(\mathbf{u}_N^{2m}, \lambda), \quad (\text{B.36})$$

for any $g \in D_n$, where g is an element of the dihedral group D_n of order $n = N/2$. \mathbf{u}_N^{2m} denotes the principal solution obtained using an N atom representative unit cell. $\mathbf{T}(g) \in \mathbb{R}^{N+1} \times \mathbb{R}^{N+1}$ is a matrix representation of the element g of the symmetry group D_n on the

carrier space \mathbb{R}^{N+1} . The dihedral group D_n of order n is given by

$$D_n = \{E, R_n, \dots, R_n^{n-1}, S, SR_n, \dots, SR_n^{n-1}\} = \{R_n^k, SR_n^k \mid k = 0, 1, \dots, n-1\}, \quad (\text{B.37})$$

where R_n represents a rotation by $360/n$ degrees and S represents a mirror operation (so $R_n^n = S^2 = (SR_n)^2 = E$). Note that Eq. (B.37) implies that the entire group D_n can be generated from the two elements R_n^k and S , for any $k = 0, 1, \dots, n-1$. Subgroups of D_n consist of dihedral and cyclic groups whose degree m divides n (i.e., the family of the subgroups of D_n) is given by

$$\{D_m^j \mid j = 1, \dots, n/m; m \text{ divides } n\} \text{ and } \{C_m \mid m \text{ divides } n\}, \quad (\text{B.38})$$

where

$$D_m^j = \{R_n^{kn/m}, SR_n^{j-1+kn/m} \mid k = 0, 1, \dots, m-1\}, \quad (\text{B.39})$$

$$C_m = \{R_n^{kn/m} \mid k = 0, 1, \dots, m-1\}. \quad (\text{B.40})$$

Note that $D_m^1 = D_m$ and $C_1 = \{E\}$. In this work, the nonlinear equilibrium equations of a multi-dimensional system is generically given by

$$\mathbf{F}_{\Omega_N}(\mathbf{u}_N, \lambda) = \mathbf{0}, \quad (\text{B.41})$$

where $\mathbf{u}_N \in \mathbb{R}^{N+1}$ describes the state of the system and $\lambda \in \mathbb{R}$. Additionally, let the bifurcation equation as defined through the LSK decomposition be given by

$$\mathbf{f}(\mathbf{w}, \lambda) = \mathbf{0}, \quad (\text{B.42})$$

where $\mathbf{f}, \mathbf{w} \in \mathbb{R}^M$, $M < N$. Assuming that the original equilibrium (Eq. (B.41)) is equivariant under D_n then, it was shown that Eq. (B.42) is equivariant to D_n with respect to the irreducible representation (associated with the null space $\mathcal{N} \equiv \ker(\nabla \mathbf{F})$ at the bifurcation point) of D_n on \mathbb{R}^M . Additionally, since the irreducible representation of D_n are either one or two-dimensional, then it must be that $M = 1$ or $M = 2$ (referred to as simple or double bifurcation points respectively). With respect to D_n , the decomposition of \mathbb{R}^{N+1} into mutually orthogonal subspaces is given by:

$$\mathbb{R}^N = \begin{cases} X_{n+} \oplus X_{n-} \oplus X_{n/2+} \oplus X_{n/2-} \oplus \bigoplus_{k=1}^{n/2-1} X_k, & \text{for } n \text{ even,} \\ X_{n+} \oplus X_{n-} \oplus \bigoplus_{k=1}^{(n-1)/2} X_k, & \text{for } n \text{ odd,} \end{cases} \quad (\text{B.43})$$

where X_{n+} , X_{n-} , $X_{n/2+}$ and $X_{n/2-}$ corresponds to irreducible representations of dimension one (in particular X_{n+} corresponds to the identity irreducible representation) and X_k for $1 \leq k < n/2$ corresponds to two-dimensional irreducible representations of D_n . Each subspaces $X \subset \mathbb{R}^{N+1}$ is associated with a subgroup $G[X]$ (referred to as the *isotropy subgroup* of X) of D_n such that

$$G[X] = \{g \in D_n \mid \mathbf{T}(g)\mathbf{x} = \mathbf{x}, \forall \mathbf{x} \in X\}. \quad (\text{B.44})$$

With the definition given in Eq. (B.44), the isotropy subgroup associated with the subspaces of \mathbb{R}^N is given by the following:

$$G[X_{n+}] = D_n; \quad G[X_{n-}] = C_n; \quad G[X_{n/2+}] = D_{n/2}; \quad G[X_{n/2-}] = D_{n/2}^2; \quad (\text{B.45})$$

$$\text{and } G[X_k] = C_m, \quad (\text{B.46})$$

where m is the greatest common divisor of n and k . At this point, we identify two cases:

1. $\dim(\mathcal{N}) = 1$ (simple bifurcation points).

According to the decomposition of \mathbb{R}^{N+1} given by Eq. (B.43), the null vector for a simple bifurcation point is invariant under either D_n , C_n , $D_{n/2}$ or $D_{n/2}^2$. At simple bifurcation points, a path bifurcates in the positive and negative directions of its null vector. The subgroup of the critical eigenvector is the isotropy subgroup associated with the (in this case one dimensional) null space \mathcal{N} . Note that for the one dimensional crystal described in this chapter, all simple bifurcation points are GSBBPs. In these case, bifurcating paths can be computed in a straightforward manner using the $O(\zeta^2)$ equation in the expansion of the bifurcation equation.

2. $\dim(\mathcal{N}) = 2$ (double bifurcation points).

Consider a double bifurcation point of a D_n -equivariant system and denote by C_m , the symmetry group associated with $\mathcal{N} = \text{span}\{\mathbf{u}_1, \mathbf{u}_2\}$ (this is possible because of Eq. (B.46)). In the case of a D_n -equivariant system, *there exist a finite number of critical eigenvectors with symmetry higher than C_m . I.e., those which are D_m^j -symmetric for some $j = 1, \dots, n/m$.* These eigenvectors \mathbf{e}^* turns out to give the directions of the bifurcating paths. An eigenvector $\mathbf{e}_1 \in \mathcal{N}$ can be defined in the following way such that it is D_m^1 invariant:

$$\mathbf{e}_1 = \frac{\mathbf{T}(g)\mathbf{u}_1 + \mathbf{u}_1}{\|\mathbf{T}(g)\mathbf{u}_1 + \mathbf{u}_1\|}, \quad \text{for any } g \neq E, \quad (\text{B.47})$$

where \mathbf{T} is a representation of D_m^1 on \mathbb{R}^N and $g \in D_m^1$ is not the identity element. Additionally, a unit vector \mathbf{e}_2 can be defined to be orthogonal to \mathbf{e}_1 by the Gram-Schmidt process and is given by the following:

$$\mathbf{e}_2 = \frac{[\mathbf{u}_2 - (\mathbf{u}_2 \cdot \mathbf{e}_1)\mathbf{e}_1]}{\|\mathbf{u}_2 - (\mathbf{u}_2 \cdot \mathbf{e}_1)\mathbf{e}_1\|}, \quad (\text{B.48})$$

where \mathbf{e}_1 is defined by Eq. (B.47). An arbitrary D_m^j -symmetric vector \mathbf{e}^* in this subspace can be expressed as

$$\mathbf{e}^* = 2w_1\mathbf{e}_1 + 2w_2\mathbf{e}_2. \quad (\text{B.49})$$

With the use of the complex variable $z = w_1 + iw_2$ (where $i = \sqrt{-1}$), Eq. (B.49) is rewritten as

$$\mathbf{e}^* = (\mathbf{e}_1 + i\mathbf{e}_2)z + (\mathbf{e}_1 - i\mathbf{e}_2)\bar{z}, \quad (\text{B.50})$$

where \bar{z} denotes the complex conjugate of z . Let $\mathbf{f}(\mathbf{w}, \lambda) = [f_1(\mathbf{w}, \lambda), f_2(\mathbf{w}, \lambda)]^\top = \mathbf{0}$ be the two-dimensional bifurcation equation defined as in Eq. (2.14). Since $\mathbf{w} \in \mathcal{N}$ and by choosing (z, \bar{z}) as the coordinates of the null space \mathcal{N} , then we can define $H(z, \bar{z}, \lambda)$ as follows:

$$H(z, \bar{z}, \lambda) = f_1(z, \bar{z}, \lambda) + f_2(z, \bar{z}, \lambda). \quad (\text{B.51})$$

Then, the bifurcation equation $\mathbf{f}(\mathbf{w}, \lambda) = \mathbf{0}$ is equivalent to the complex equation

$$H(z, \bar{z}, \lambda) = 0, \quad (\text{B.52})$$

where $H(z, \bar{z}, \lambda)$ is defined as in Eq. (B.51). It can be shown that the D_n -equivariance of the bifurcation equation is equivalent to the following:

$$\omega H(z, \bar{z}, \lambda) = H(\omega z, \omega \bar{z}, \lambda), \quad (\text{B.53})$$

$$H(z, \bar{z}, \lambda) = \overline{H(z, \bar{z}, \lambda)}, \quad (\text{B.54})$$

where $\omega = \exp(i2\pi m/n)$. The equilibrium equations (Eq.(B.41)) has the trivial solution with $z = 0$ that corresponds to the D_n -symmetric \mathbf{u}_N^{2m} solution. The non-trivial solution is determined from $H/z = 0$. Assuming that Eq. (B.52) has an expansion

given by

$$H(z, \bar{z}, \lambda) = \sum_{k=0}^{\infty} \sum_{q=0}^{\infty} A_{kq}(\lambda) z^k \bar{z}^q, \quad (\text{B.55})$$

and defining

$$\tilde{H}(r \exp(i\theta), r \exp(-i\theta), \lambda) = H(r \exp(i\theta), r \exp(-i\theta), \lambda) / (r \exp(i\theta)) = H(z, \bar{z}, \lambda) / z, \quad (\text{B.56})$$

where $z = r \exp(i\theta)$, then Eq. (B.54) implies that

$$\text{im}(\tilde{H}) = \sum_{k=1}^{\infty} \sum_{q=0}^{\infty} [A_{q+1+kn', q}(\lambda) r^{2q+kn'} - A_{q, q-1+kn'}(\lambda) r^{2(q-1)+kn'}] \sin(kn'\theta) = 0, \quad (\text{B.57})$$

where $\text{im}(\tilde{H})$ is the imaginary part of \tilde{H} . In general the coefficients in Eq. (B.57) are non-zero and so, Eq. (B.57) implies $\sin(n'\theta) = 0$ or

$$\theta = \pi(j-1)m/n, \quad j = 1, \dots, 2n/m. \quad (\text{B.58})$$

The form of the right hand side of Eq.(B.57) is derived in detail in *Kiyohiro et al.* (1991) and is reproduced here in order to show the condition that determines Eq.(B.58), which ultimately determines (w_1, w_2) in Eq. (B.49).

B.3.2 Period-extending bifurcations from \mathbf{u}_{10}^{2m} and \mathbf{u}_{14}^{2m}

The technique discussed in *Kiyohiro et al.* (1991) and summarized in the preceding subsection was used for to determine the bifurcating tangents of the bifurcation points found along the \mathbf{u}_{10}^{2m} and \mathbf{u}_{14}^{2m} paths. The technique greatly simplifies the determination of the tangent to the bifurcating path. Instead of going through the various high order ζ terms, the problem of finding the tangent reduces to first finding the subgroup C_m from which the D_m^j -symmetric tangent can then be obtained. Using these techniques, the following simulation results are obtained using a simulated cell consisting of 10 atom (Fig. B.8) and 14 atom (Fig. B.9) non-essential unit cell. The following is a list of the structure corresponding to each path in Fig. B.8: \mathbf{u}_{10}^{10m} (tan), \mathbf{u}_{10}^{10m} (navy), \mathbf{u}_{10}^{10m} (dark olive), \mathbf{u}_{10}^{10m} (olive), \mathbf{u}_{10}^{10m} (sienna), and the following gives a list of the structure corresponding to each path in Fig. B.9: \mathbf{u}_{14}^{14m} (coral), \mathbf{u}_{14}^{14m} (pink), \mathbf{u}_{14}^{14m} (dodger blue), \mathbf{u}_{14}^{14m} (aquamarine), \mathbf{u}_{14}^{14m} (cadet blue), \mathbf{u}_{14}^{14m} (chartreuse), \mathbf{u}_{14}^{14m} (gold). Observe that a stable path corresponding to a \mathbf{u}_{10}^{10m}

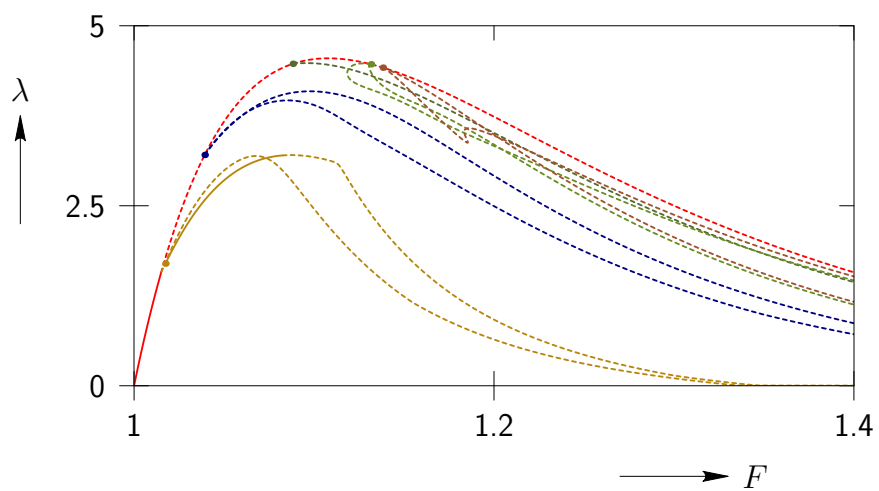


Figure B.8: Simulation result using a 10 atom non-essential unit cell.

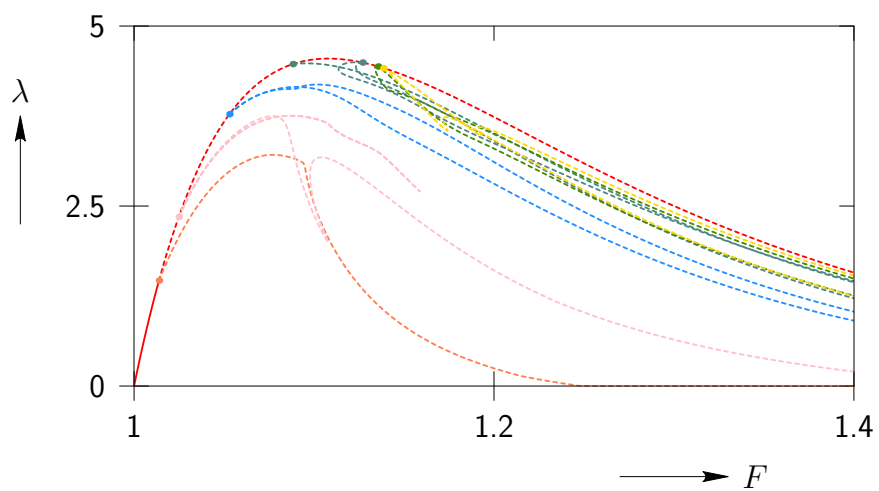


Figure B.9: Simulation result using a 14 atom non-essential unit cell.

structure is found using a Ω_{10} representative unit cell (Fig. B.8).

B.4 Summary

This chapter discusses the simulation of a phase transforming one dimensional biatomic crystal under tension using a BFB technique discussed in Ch (2). The crystal is modeled using the model developed by *Dobson et al.* (2008) which describes the atomic interactions through a Lennard-Jones potential. The DOFs of the system is described using the CB kinematics that relates the current configuration of the crystal to the uniform deformation and inter-atomic shift.

The first set of simulations uses a 24 atom simulated cell to capture period-extending phase transformations corresponding to M -atom essential unit cell structures where M divides 24 (and $M \geq 2$ for a biatomic crystal). This simulation establishes the reference structure (denoted by \mathbf{u}_{24}^{2m}) as a stable two-atom essential unit cell with mirror symmetry. Upon (quasi-static) application of a tensile load, the structure continues to stretch uniformly until a critical load is reached at which point the \mathbf{u}_{24}^{2m} structure becomes unstable (see Fig. B.2). A total of thirteen bifurcation points were found along the \mathbf{u}_{24}^{2m} path. Out of the thirteen bifurcation points, seven corresponds to a typical symmetry-breaking bifurcation point where a straightforward application of the LSK asymptotic bifurcation analysis was used to determine the tangents to the bifurcating paths. The tangents were then used to seed the numerical branch-following method to compute paths bifurcating from the seven typical symmetry-breaking bifurcation points. Only two bifurcating paths (corresponding to a \mathbf{u}_{24}^{4m} and \mathbf{u}_{24}^{6m} structure) were found to be stable (see Fig. B.4). The remaining six bifurcation points involve subtle complications in the LSK asymptotic analysis. Group-theoretic techniques and the Fredholm alternative were then used in order to obtain the asymptotic expression to the bifurcating paths (see Fig. B.7). Two additional sets of simulations were performed using 10 atom and a 14 atom simulated cell. For these two cases, a technique developed by *Kiyohiro et al.* (1991) was used to expediate the process of identifying the tangents to the bifurcating paths (see Fig. B.8 and Fig. B.9). One additional stable path \mathbf{u}_{10}^{10m} was found.

The objective of this work to present the details of the BFB method in order to highlight some of the complications that may arise in the application of the method. However, these complications may be unavoidable in certain cases (such as investigations of period-extending bifurcations). It is our belief that this document is useful in illustrating the nature of the steps involved in such cases.

Appendix C

An example of the determination of the forms of the coefficients in the expansion of the bifurcation equation

This specific example relates to the bifurcation points shown in blue in Fig. (C.1). This

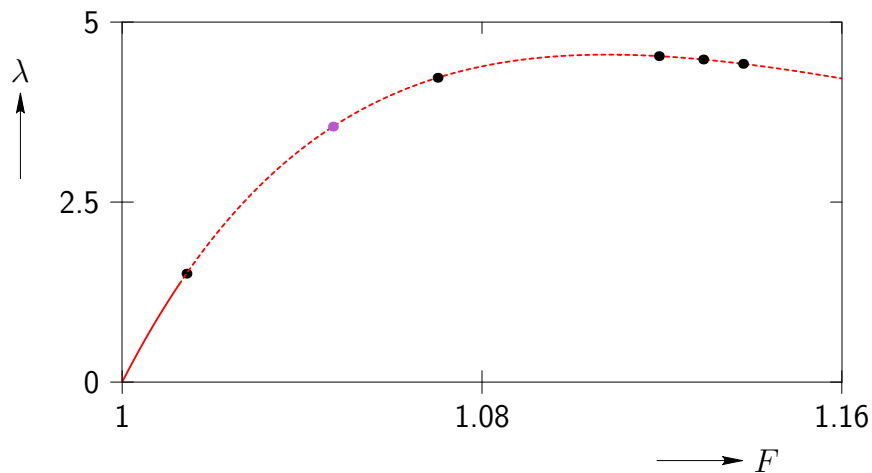


Figure C.1: A figure showing six SSBBPs found along the principal path. The example calculation detailed in this section relates to the SSBBP shown in blue. The remaining five SSBBP are shown in black.

bifurcation point is a symmetry-breaking bifurcation point found along a D_{12} -symmetric path.

The group D_n is the dihedral group of order n and represents the symmetry of a regular

n -gon and is defined as

$$D_n = \{E, R_n, \dots, R_n^{n-1}, S, SR_n, \dots, SR_n^{n-1}\} = \{R_n^k, SR_n^k \mid k = 0, 1, \dots, n-1\}, \quad (\text{C.1})$$

where R_n represents a rotation by $360/n$ degrees and S represents a mirror operation (so $R_n^n = S^2 = (SR_n)^2 = E$). Note that Eq. (C.1) implies that the entire group D_n can be generated from the two elements R_n^k and S , for any $k = 0, 1, \dots, n-1$.

For the bifurcation point shown in blue in Fig. (C.1), the null space \mathcal{N} is two-dimensional and is associated with the following irreducible representation \mathcal{T} of D_{12} (where only the generators R_{12}^k and S are shown):

- $\mathcal{T}(R_{12}^k) = \begin{bmatrix} \cos(2\pi k/12) & \sin(2\pi k/12) \\ \sin(2\pi k/12) & \cos(2\pi k/12) \end{bmatrix},$
- $\mathcal{T}(S) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$

Now, recall from Ch. (2), that the bifurcation equation is shown to be equivariant under the irreducible representation \mathbf{T} . This implies that any m -th order coefficient in the expansion of the bifurcation equation is invariant under the action of the m -th order representation. Therefore, the form of an m -th order coefficient in the bifurcation equation expansion can be found by finding the invariant subspace associated with the identity irreducible representation for the m -th order representation $\mathbf{T}^{(m)} = \otimes_{i=1}^m \mathcal{T}$. The remaining part of this appendix will present an example of this process to obtain the form of $E_{,\xi\xi\lambda}$ and $E_{,\xi\xi\xi\xi}$.

C.1 Form of $E_{,\xi\xi\lambda}$

First, note that \mathcal{T} is a representation on \mathbb{R}^2 . Now, since $\dim(\mathcal{N}) = 2$, then $\boldsymbol{\xi} = [\xi_1, \xi_2]^\top \in \mathcal{N}$. Therefore, in component form, $E_{,\xi\xi\lambda}$ is given by $E_{,\xi\xi\lambda} = [E_{,11\lambda}, E_{,12\lambda}, E_{,21\lambda}, E_{,22\lambda}]^\top \in \mathbb{R}^2 \otimes \mathbb{R}^2$. Additionally since, $E_{,\xi\xi\lambda}$ is symmetric, and by Prop. (3), a representation $\mathbf{T}^{(2)} \in \mathbb{R}^{2 \times 2}$ on the carrier space $\mathbb{R}^2 \otimes \mathbb{R}^2$ of $E_{,\xi\xi\lambda}$ is given by the fully symmetric tensor product constructed from \mathcal{T} and is given by

$$[\mathbf{T}^{(2)}(g)]_{ijkl} = \frac{1}{2} ([\mathcal{T}(g)]_{ik}[\mathcal{T}(g)]_{jl} + [\mathcal{T}(g)]_{il}[\mathcal{T}(g)]_{jk}), \quad \forall g \in G, \quad (\text{C.2})$$

where in this case, $G = D_{12}$. Once the representation $\mathbf{T}^{(2)}(g)$ is known, then the projection operator $\mathbf{P}^{(m)}$ associated with the identity irreducible representation can be constructed

using

$$\mathbf{P}^{(m)} = \frac{1}{|G|} \sum_{g \in G} \mathbf{T}^{(m)}(g), \quad (\text{C.3})$$

In this case, $\mathbf{P}^{(2)} \in \mathbb{R}^{2 \times 2}$ is given by

$$\mathbf{P}^{(2)} = 0.5 \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \quad (\text{C.4})$$

The basis for the invariant subspace is then the eigenvector of $\mathbf{P}^{(2)}$ associated with the eigenvalue equal to one and is given by $\mathbf{e} = [1, 0, 0, 1]^\top$. Finally, since $E_{,\xi\xi\lambda}$ is invariant under $\mathbf{T}^{(2)}$ (i.e., $\mathbf{T}^{(2)}(g)E_{,\xi\xi\lambda} = E_{,\xi\xi\lambda}$, for all $g \in G$), then it must be that $E_{,\xi\xi\lambda} \in \text{span}\{\mathbf{e}\}$ or $E_{,\xi\xi\lambda} = [E_{,11\lambda}, E_{,12\lambda}, E_{,21\lambda}, E_{,22\lambda}]^\top = a[1, 0, 0, 1]^\top$ or $E_{,ij\lambda} = a\delta_{ij}$ where $a \in \mathbb{R}$ and $E_{,ij\lambda} \equiv \frac{\partial^3 E}{\partial \xi_i \partial \xi_j \partial \lambda}$.

C.2 Form of $E_{,\xi\xi\xi\xi}$

The process to determine the form of $E_{,\xi\xi\xi\xi}$ is similar to the two preceding case. Since $E_{,\xi\xi\xi\xi}$ must inherit the symmetry of D_{12} due to the equivariance of the bifurcation equation, then $\mathbf{T}^{(4)}(g)E_{,\xi\xi\xi\xi} = E_{,\xi\xi\xi\xi}$, for all $g \in G$ where $\mathbf{T}^{(4)}(g)$ is a representation of D_{12} on the carrier space $\mathbb{R}^2 \otimes \mathbb{R}^2 \otimes \mathbb{R}^2 \otimes \mathbb{R}^2$ of $E_{,\xi\xi\xi\xi}$ and is given by

$$[\mathbf{T}^{(4)}(g)]_{ijklmnop} = \frac{1}{2} ([\mathbf{T}^{(2)}(g)]_{ijmn}[\mathbf{T}^{(2)}(g)]_{klop} + [\mathbf{T}^{(2)}(g)]_{ijop}[\mathbf{T}^{(2)}(g)]_{klmn}), \quad \forall g \in G, \quad (\text{C.5})$$

where $\mathbf{T}^{(2)}$ is constructed using Eq. (C.2). The projection operator $\mathbf{P}^{(4)} \in \mathbb{R}^{2 \times 2 \times 2 \times 2}$ is constructed using Eq. (A.13). Obtaining the eigenvector of $\mathbf{P}^{(4)}$ associated with eigenvalue equal to one, we find that $E_{,1111} = E_{,2222} = 3E_{,1122} = b$, $b \in \mathbb{R}$, and other terms are equal to zero.

C.3 A remark regarding symmetric bifurcation points

Note that while the bifurcation point shown in blue is known to be a symmetric bifurcation and therefore $E_{,\xi\xi\xi}$ is known to be such that $E_{,\xi\xi\xi} = \mathbf{0}$, this is in fact the result of the null space being associated to this particular irreducible representation \mathcal{T} . Since $E_{,\xi\xi\xi}$ must

inherit the symmetry of G , then $\mathbf{T}^{(3)}(g)E_{,\xi\xi\xi} = E_{,\xi\xi\xi}$, for all $g \in G$, where $\mathbf{T}^{(3)}$ is a representation of D_{12} on the carrier space of $E_{,\xi\xi\xi} \in \mathbb{R}^2 \otimes \mathbb{R}^2 \otimes \mathbb{R}^2$. As before a projection operator $\mathbf{P}^{(3)} \in \mathbb{R}^{2 \times 2 \times 2}$ is constructed using Eq.(C.3) and is here found to be $\mathbf{P}^{(3)} = \mathbf{0}$ and therefore $E_{,\xi\xi\xi} = \mathbf{0}$.

Appendix D

The application of a kinematic compatibility constraint

A set of calculations were performed in order to examine the effect of an austenite kinematic compatibility constraint on the energies of the structures discussed in Ch. 3. This appendix will present the details of how the BFB method was augmented to accommodate the application of this constraint.

As mentioned in Ch. 3, the structures were simulated subject to the constraint that the middle eigenvalue λ_2 of \mathbf{U} for the structures were equal to that found for B2 at $\theta = 1.5$. Additionally, recall that the bulk free energy density ($W(\mathbf{u}, \theta)$) is given by Eq.(3.4), where the various interactions in AuCd is given by the temperature dependent Morse EIPs detailed in *Guthikonda and Elliott (2011)*. The application of the kinematic compatibility constraint was done here by adding a penalty term to the free energy density such that deviations away from the constraint manifests itself as an increase in the bulk free energy of the crystal. This is represented in the following equation,

$$\widetilde{W}(\mathbf{u}, \theta) = W(\mathbf{u}, \theta) + \frac{1}{\epsilon} \left(\lambda_2^{(\text{B2})} - \lambda_2 \right)^2, \quad (\text{D.1})$$

where, $\lambda_2^{(\text{B2})}$ is the desired constraint corresponding to the middle eigenvalue of \mathbf{U} for the B2 structure at $\theta = 1.5$. λ_2 is the middle eigenvalue of \mathbf{U} for the structure of interest and ϵ is an adjustable parameter used to control ease of convergence. The equilibrium equations

are then given by derivatives of Eq. (D.1) given by

$$\frac{\partial \widetilde{W}}{\partial \mathbf{u}} = \mathbf{0} \begin{cases} \frac{\partial \widetilde{W}}{\partial \mathbf{U}} = \mathbf{0}, \\ \frac{\partial \widetilde{W}}{\partial \mathbf{s}_\alpha} = \mathbf{0}, \quad \alpha = 0, 1, \dots, N - 1. \end{cases} \quad (\text{D.2})$$

Lastly, by starting from a stress-free configuration associated with each structure of interest (obtained through the BFB study for modulated martensites detailed in Ch. 3), the predictor-corrector routine in the BFB method will yield solutions to Eq. (D.2) that describe the constrained behavior of the corresponding structure.

Appendix E

The M^4 algorithm

This appendix presents an algorithm for the construction of a modulated martensite according to M^4 . The M^4 algorithm presented here provides the deformed configuration corresponding to a modulated martensite by giving the current position of the atoms in the primitive unit cell of the modulated martensite. Note that the algorithm assumes that the B2, B19, and $L1_0$ input data has been obtained from the BFB study of the model. The BFB study performed in this work uses the CB kinematics to describe the deformation of a given crystal (see Ch. 1 or Ch.3 for the definition of the CB kinematics). As such, the input data obtained in this study is the uniform deformation \mathbf{U} of a particular structure and the fractional shift vectors \mathbf{S} of all the atoms in the simulated cell.

E.1 B2, B19, and $L1_0$ input data

The M^4 algorithm constructs a modulated martensite of modulation equal to $2N$. The even numbered modulation is because the M^4 assumes a B2 reference configuration with lattice vectors of the form given by:

$$\mathbf{e}_1 = N \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad (\text{E.1})$$

and hence uses a non-primitive description of a B2 structure which comprises of at least two $(110)_{\text{cubic}}$ planes. The first set of input data is obtained from the B2 structure and is given by the following:

- $\Delta = \frac{1}{2N}\mathbf{e}_1$, which is a vector from the geometric centroid of adjacent $(110)_{\text{cubic}}$.
- $\delta_{(p,\alpha)}$, which is a vector from the centroid of plane p to atom α in plane p in the modulated martensite (which is of the form $\delta_{(p,\alpha)} = \delta_{(p,\alpha)}^2\mathbf{e}_2 + \delta_{(p,\alpha)}^3\mathbf{e}_3$). Note that $p = 1, \dots, 2N$ since there are $2N$ planes and $\alpha = 1, \dots, 4N$ since there are 2 atoms per plane.

The input data for B19 is given by

- The energies $\mathcal{E}_{B^-} = \mathcal{E}_{B^+} = \mathcal{E}_{B19}$
- \mathbf{U}_{B19} , which is the uniform deformation gradient corresponding to the B2 to B19 transformation. In the \mathcal{B}_{M^4} basis, this is given by

$$\mathbf{F}_{B^+} = \mathbf{F}_{B^-} = \begin{bmatrix} a + \gamma & 0 & 0 \\ 0 & a - \gamma & 0 \\ 0 & 0 & b \end{bmatrix}, \quad (\text{E.2})$$

- $\mathbf{S}_{(p,\alpha)}$ which is the fractional shift vector in of atom α in plane p given in the \mathcal{B}_{M^4} basis. $\mathbf{S}_{(p,\alpha)}$ can be written as a sum of the in-plane centroidal shift \mathbf{s}_p and an in-plane atomic displacement $\mathbf{ds}_{(p,\alpha)}$ as follows

$$\mathbf{S}_{(p,\alpha)} = \mathbf{s}_p + \mathbf{ds}_{(p,\alpha)} \quad (\text{E.3})$$

where, the centroidal shift may be computed as

$$\mathbf{S}_p = \frac{1}{2} [\mathbf{S}_{(p,1)} - \mathbf{S}_{(p,2)}] = s_p \tilde{\mathbf{g}}_2 \quad (\text{E.4})$$

where $\mathbf{S}_{(p,1)}$ and $\mathbf{S}_{(p,2)}$ are the shift vectors of the first and second atom in plane p respectively. s_p may be positive or negative depending on whichever plane is defined as the first plane. Here, a negative s_p will correspond to \mathbf{S}_{B^+} and a positive s_p will correspond to \mathbf{S}_{B^-} . That is if $s_p < 0$,

$$\mathbf{S}_{B^+} = s_p[-1, 1, 0]^\top, \quad (\text{E.5})$$

$$\mathbf{S}_{B^-} = -s_p[-1, 1, 0]^\top. \quad (\text{E.6})$$

The input data for L1₀ is given by

- The energies $\mathcal{E}_{L^-} = \mathcal{E}_{L^+} = \mathcal{E}_{L1_0}$

- Uniform deformation corresponding to the transformation of B2 into two variants of L1₀ and given (in the \mathcal{B}_{M^4} basis) by,

$$\mathbf{F}_{L1_0^-} = \frac{1}{2} \begin{bmatrix} \ell_1 & -\ell_2 & 0 \\ -\ell_2 & \ell_1 & 0 \\ 0 & 0 & 2d \end{bmatrix}, \quad \mathbf{F}_{L1_0^+} = \frac{1}{2} \begin{bmatrix} \ell_1 & \ell_2 & 0 \\ \ell_2 & \ell_1 & 0 \\ 0 & 0 & 2d \end{bmatrix}, \quad (\text{E.7})$$

where $\ell_1 = c + d$ and $\ell_2 = c - d$. The final deformation gradient is obtained from Eq. (4.12) and Eq. (4.13) and is given by

$$\mathbf{F}_{L^+} = \mathbf{R}^+ \mathbf{F}_{L1_0^+}, \quad \mathbf{F}_{L^-} = \mathbf{R}^- \mathbf{F}_{L1_0^-}, \quad (\text{E.8})$$

where \mathbf{R}^\pm is given by

$$\mathbf{R}^\pm = \begin{bmatrix} \cos(\pm\theta) & -\sin(\pm\theta) & 0 \\ \sin(\pm\theta) & \cos(\pm\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{where } \tan(\theta) = \frac{\ell_1}{\ell_2}. \quad (\text{E.9})$$

- L1₀ is characterized only by a uniform deformation and as such,

$$\mathbf{S}_{L^+} = \mathbf{S}_{L^-} = \mathbf{d}\mathbf{s}_{L^+} = \mathbf{d}\mathbf{s}_{L^-} = \mathbf{0}, \quad (\text{E.10})$$

Once the input data has been obtained (for a given temperature), the following algorithm can be used to obtain the deformed position vectors $\mathbf{x}_{(p,\alpha)}$ of the atoms in the primitive unit cell of the modulated martensite. Additionally, the average deformation gradient and the energy of the modulated martensite is given by

$$\mathbf{U}_{M^4} = \frac{1}{N} \sum_{p=p_1}^{p_{2N}} \mathbf{U}_p, \quad (\text{E.11})$$

$$\mathcal{E}_{M^4} = \frac{1}{N} \sum_{p=p_1}^{p_{2N}} \mathcal{E}_p, \quad (\text{E.12})$$

where $\mathcal{P} = (p_1, \dots, p_{2N})$ gives the M^4 sequence for a particular modulated martensite.

Algorithm 2 M^4 algorithm to construct a modulated martensite of modulation equal to $2M$

```

1: Input:  $M^4$  sequence  $\mathcal{P} = (p_1, \dots, p_{2M})$  %  $p_i$  is plane type of plane  $i$ 
    $\Delta, \delta_{(p,\alpha)}, \mathbf{U}_{L^-}, \mathbf{U}_{L^+}, \mathbf{U}_{B^-}, \mathbf{U}_{B^+}, \mathbf{S}_{L^-}, \mathbf{S}_{L^+}, \mathbf{S}_{B^-}, \mathbf{S}_{B^+}, \mathbf{ds}_{L^-}, \mathbf{ds}_{L^+}, \mathbf{ds}_{B^-}, \mathbf{ds}_{B^+}$ 
2:  $\mathbf{G} = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ 
3:  $\mathbf{U}_{\text{plane}} = \frac{1}{2M} \sum_{p=p_1}^{p_{2M}} \mathbf{U}_p$ 
4:  $\mathbf{U} = \frac{1}{2}(\mathbf{U}_{p_1} + \mathbf{U}_{p_{2M}})$ 
5:  $\mathbf{C}_{p_1} = \mathbf{UGS}_{p_1}$  % Deformed position of centroid of plane 1
6: for  $\alpha = 1, 2$  do
7:    $\mathbf{x}_{(p_1,\alpha)} = \mathbf{C}_{p_1} + \mathbf{U}_{\text{plane}}\mathbf{G} [\delta_{(p_1,\alpha)} + \mathbf{ds}_{(p_1,\alpha)}]$  % Deformed positions of atoms  
in plane 1
8: end for
9:  $\mathbf{O}_{\text{shift}} = [0, 0, 0]^T$  % Set centroid of plane 1 to origin
10: for  $p = p_2, \dots, p_{2M}$  do
11:    $\mathbf{U} = \frac{1}{2}(\mathbf{U}_p + \mathbf{U}_{p-1})$ 
12:    $\mathbf{O}_{\text{shift}} = \mathbf{O}_{\text{shift}} + \mathbf{UG}\Delta$ 
13:    $\mathbf{C}_p = \mathbf{O}_{\text{shift}} + \mathbf{UGS}_p$  % Deformed position of centroid  
of planes 2 to  $2M$ 
14:   for  $\alpha = 1, 2$  do
15:      $\mathbf{x}_{(p,\alpha)} = \mathbf{C}_p + \mathbf{U}_{\text{plane}}\mathbf{G} [\delta_{(p,\alpha)} + \mathbf{ds}_{(p,\alpha)}]$  % Deformed positions of atoms  
in planes 2 to  $2M$ 
16:   end for
17: end for

```
