Minimal Surfaces and Topology of Martensitic Phase Transformations

Imperial College London

Author: Yin Mengdi
Department of Physics

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Physics and the Diploma of Imperial College London

06/09/2024

Abstract

Martensitic phase transformations and the shape memory effect have been extensively studied, with various theories proposed to understand their mechanisms. This work develops a model based on differential geometry to distinguish martensitic phase transformations from ordinary phase transformations and shape memory alloys from ordinary martensites, excluding cases involving magnetism.

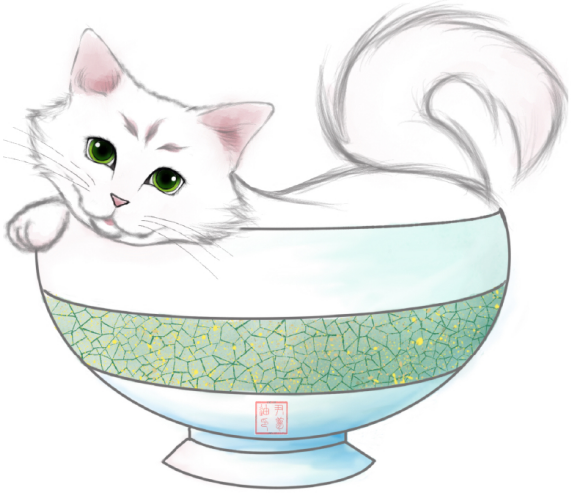
First, we analyse the surface S of a given charge density distribution in crystals using density functional theory (DFT) for Cu, Al, Na, Zr, and NiTi in different phases.We find that S corresponding to different crystals in different phases converge to different triply periodic minimal surfaces (TPMS). The point group of a crystal is a subgroup of its corresponding TPMS, similar to its relationship with the point group of its corresponding Bravais lattice. Notably, the TPMS of two end phases in a martensitic transformation share the same genus, unlike in ordinary phase transformations.

Next, we explore the topological continuity between the two end phases. Since martensitic transformations are diffusionless, the connections between lattice points and their nearest neighbours remain invariant. These connections can be represented by skeletal graphs derived from TPMS. Hence, the invariance of connections means the topological equivalence between skeletal graphs (TPMS) of the two end phases.

Finally, we investigate the geometric reason that distinguishes shape memory alloys from ordinary martensites. Since shape memory alloys undergo reversible martensitic transformations, we introduce the concept of "geometric entropy,"analogous to thermodynamic entropy but determined by the geometry of S and the arrangement of lattice points. DFT calculations reveal that only shape memory alloys can have zero-barrier transformation paths, which coincide with paths where geometric entropy remains unchanged. Thus, the shape memory effect is possible if a martensitic phase transformation occurs without variation in its geometric entropy.

Acknowledgements

I would like to thank the following people for helping me with patience, encouragement and support throughout my project. Thank you to my supervisor, Professor Dimitri Vvedensky, for steering me through this project with his knowledge and insight into this subject and thoughtful comments on this thesis. Thank you to Professor Jing Zhang for taking the time to discuss this project with me and inspire me from a new perspective. Thank you to my parents for funding my study. And thank you to my cats, Fourier (white) and Schrödinger (black), for being models in illustrations at the beginning of each chapter.





Declaration

I declare that except references specifically made to work completed by others, this thesis is my original work and has not been submitted in any other institutions for the consideration of an academic degree. This thesis is completed by myself, without collaboration with others, except these specified in the Acknowledgements and the thesis.

Yin Mengdi

June 2024

Copyright

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Glossary

Notation	Meaning
$\otimes$	Direct (tensor) product
$\rtimes$	Semidirect product
$\oplus$	Direct sum
$A \trianglelefteq B$	A is an invariant subgroup of B
$A \triangleleft B$	A is an invariant subgroup of B (only proper)
$\lvert A \rvert$	Oder of a group A
$\circ$	Operating on
ε_{ijk}	Levi-Civita symbol
δ_i^j	Kronecker symbol
$\Gamma^i{}_{jk}$	Christoffel symbol
g_{ij}	Metric
g_N	Genus
$\forall$	For all
$\exists$	Exist
$\succ$	Being positive defined

Contents

1 Introduction **13**

1.1 Outline . . . 19

1.2 Publications . . . 21

2 Martensitic Phase Transformations and the Shape Memory Effect **24**

2.1 Definition of Martensite and Shape Memory Effect . . . 26

2.2 Proposed Theories and Models . . . 29

2.2.1 Landau's Theory of Crystal Phase Transformations . . . 29

2.2.2 Reconstructive Phase Transformations . . . 42

2.2.3 Phonon Softening . . . 44

2.2.4 Deformation Models . . . 46

2.2.5 Continuum Theory Based on Energy Minimisation . . . 49

3 Triply Periodic Minimal Surfaces **56**

3.1 Zero Equipotential Surfaces . . . 57

3.2 Background of Differential Geometry . . . 61

3.3 Mathematical Description of Minimal Surfaces . . . 65

3.3.1 Gauss Map, Fundamental Forms and Curvatures . . . 65

3.3.2 Geometric Interpretation of Gaussian Curvature . . . 68

3.3.3 Geometric Interpretation of Mean Curvature and Minimal Surface . . . 70

3.3.4 Projection of TPMS onto $\mathbb{C}^*$ and The Weierstrass Representation . . . 71

3.3.5 Expression of Triply Periodic Minimal Surfaces . . . 76

4 Visualisation of Minimal Surfaces Corresponding to Crystals **80**

4.1 Introduction to Density Functional Theory . . . 82

4.2 Method and Results . . . 87

4.2.1 Sodium . . . 89

4.2.2 Aluminium . . . 90

4.2.3 Copper . . . 91

4.2.4 Zirconium . . . 92

4.2.5 Nitinol . . . 93
4.3 Conclusion . . . 94

5 Reversibility and Continuous Deformations **97**
5.1 A Glance at Topology and Manifold . . . 98
5.2 Gauss-Bonnet Theorem . . . 99
5.3 Conservation of Genus and Martensitic Phase Transformation . . . 101
5.3.1 FCC−BCC Phase Transformation . . . 108
5.3.2 HCP−BCC and HCP−Hexagonal Phase Transformations . . . 109
5.3.3 B2−B19′ Phase Transformations in NiTi . . . 110
5.4 Conclusion . . . 113

6 Convergence to Triply Periodic Minimal Surfaces in Crystals **115**
6.1 Vanishing of Energy Barrier and Shape Memory Effect . . . 116
6.1.1 Density-Functional Theory Calculations of Energy Barriers . . . 116
6.2 Helfrich Hamiltonian . . . 118
6.3 Geometric Description of System Energy . . . 120
6.3.1 Revision of Lattice Points . . . 120
6.3.2 Geometric Description of Ground State . . . 122
6.3.3 Minimal Surface and Charge Density . . . 127
6.4 Geometric Explanation of the Shape Memory Effect . . . 131
6.4.1 Model Construction, Geometric Entropy . . . 131
6.4.2 Shape Memory Effect in NiTi . . . 134
6.4.3 ΔS_g and Energy Barrier . . . 141
6.5 Conclusion . . . 142

7 Conclusion and Outlook **145**

A Group Theory and Crystallography **151**
A.1 Group Theory and Representation Theory . . . 151
A.1.1 Abstract Groups . . . 151
A.1.2 Representation Theory . . . 153
A.1.3 Crystallography and Crystallographic Groups . . . 154
A.2 Lifshitz Criterion . . . 155
A.3 Powers of Irreducible Representations . . . 156
A.4 Group Theoretic Expression of the Landau Criterion and the Lifshitz Criterion . . . 158
A.4.1 Irreducible Representation X_5^- . . . 160

B Differential Geometry Background **161**
B.1 Quadratic Form and Congruence . . . 161
B.2 Tensor Algebra and Differential Geometry . . . 162
B.2.1 Definition of Metric Space . . . 162
B.2.2 Metric Differentiation . . . 164
B.2.3 Derivation of Weierstrass Representation . . . 166
B.2.4 Gaussian Curvature in Weierstrass Representation . . . 169

C Density Functional and Root-Mean-Square Calculations **174**
C.1 Calculating Bloch States in Reciprocal Space . . . 174
C.2 RMS Data Sets . . . 175
C.2.1 Na . . . 175
C.2.2 NiTi . . . 176
C.2.3 Zr . . . 177
C.2.4 Cu . . . 179
C.2.5 Al . . . 181

D Topology and Gauss–Bonnet Theorem **182**
D.1 Mathematical Definition of Topology . . . 182
D.2 Embedding and Immersion . . . 183
D.3 Gaussian Curvature as Area Ratio . . . 184
D.4 Spherical Triangle and Triangulation of Polygons . . . 185
D.5 Euler Characteristic of a Planar Graph . . . 186
D.6 Euler Characteristic and Genus . . . 187
D.7 Derivation of Gauss-Bonnet Theorem . . . 189
D.8 Calculating Gaussian Curvature on Minimal Surfaces . . . 193
D.8.1 P Surface . . . 193
D.8.2 I-WP Surface . . . 194
D.8.3 F-RD Surface . . . 195
D.8.4 H′-T Surface . . . 195
D.8.5 oPb Surface Describing B19′, 109-B19′ and R NiTi . . . 195
D.8.6 oPa Surface Describing B33 NiTi . . . 196

E Motion of Particles on A Manifold **197**
E.1 Metric Space and Christoffel Symbols . . . 197
E.1.1 Christoffel Symbol and Connection . . . 197
E.2 Generalised Hamiltonian Method . . . 200
E.3 Details of Hamiltonian Calculation . . . 202
E.3.1 Simplification Using Gauss Normal Coordinates . . . 203

E.3.2 Commutation Relation and Christoffel Symbol 204
E.3.3 Coordinate Transformation of Commutation Relation 205

References **206**

List of Figures

1.1 Sketch of a lattice and the correspnoding deformed lattice under a shear 13
1.2 Sketch of a crushed antenna and how it recovers to its original shape. 14
1.3 Sketch of the mechanism of a martensitic phase transformation and the mechanism of the shape memory effect. 15
1.4 Sketch of BCC and FCC lattices and their fittings to the P and the D surfaces, respectively. 17

2.1 Sketch of the phase diagram of phase transformation of a magnetic system. . . . 25
2.2 Sketch of how variants are symmetry-related in a cubic to tetragonal lattice deformation. 28
2.3 Shear mechanisms of martensitic phase transformations. 28
2.4 Sketch of a cube deformed under a simple shear and a pure shear. 29
2.5 Sketch of a 3D space divided into the direct sum of 2 subspaces W and U. . . . 32
2.6 Sketch of a simple cubic lattice and an orthogonal lattice. 37
2.7 Hierarchy of 32 point groups [86]. 41
2.8 Phase diagram of a phase transformation from the HCP lattice to the hexagonal lattice. 44
2.9 Sketch of the core-shell model. 45
2.10 Phonon frequency of SnS in *Pnma* space group and *Cmcm* space group. 47
2.11 Sketch of the Bain deformation, the Nishiyama–Wassermann deformation and the Bogers–Burgers deformation. 48
2.12 Sketch of the Burgers mechanism and the BCC–HCP deformation model of Ti-Ta/Ti-Tb alloys. 49
2.13 Sketch of deformation $\boldsymbol{y}$ that brings a configuration on left to a new configuration to right. 50
2.14 Energy wells of $p = |P_a|/|P_m|$ variants and the austenite phase. 52

3.1 Plots of a sea urchin plate and its corresponding P surface arrangement. 57
3.2 Sketch of triply periodic minimal surfaces. 59
3.3 Diffraction of a plane wave $\boldsymbol{k}_0$ by atoms labelling by O and Q in a crystal. . . . 59

3.4 Sketch of the P* surface. . . . 61
3.5 Sketch of a normal sections α and β. . . . 62
3.6 Sketch of principal sections α and β. . . . 64
3.7 Sketch of Gauss map from a surface S to a unit sphere. . . . 66
3.8 Sketch of a Möbius strip and a sphere. . . . 66
3.9 Sketch to illustrate in what case the Gaussian curvature equal to zero, being larger than zero, and being smaller than zero. . . . 69
3.10 Gauss map of a plane and a unit sphere. . . . 69
3.11 Sketch of a spherical triangle on a sphere. . . . 70
3.12 Sketch of a bubble film and a bubble hemisphere. . . . 71
3.13 Sketch of the stereographic projection. . . . 72
3.14 Sketch of projecting a line of 3D objects onto a 2D plane such that the translational symmetry is preserved. . . . 73
3.15 Sketch of two different projections. . . . 73
3.16 Sketch of a parallelogram whose opposite edges are pairwise identified. . . . 75
3.17 Sketch of surfaces belonging to the oPa family and oPb family. . . . 77

4.1 Flow chart of how a SCF calculation is realised in DFT calculations. . . . 84
4.2 Sketch of the Monkhorst–Pack sampling on a square lattice in the first Brillouin zone and a hexagonal lattice in the first Brillouin zone. . . . 86
4.3 Surfaces of a given charge density of Na in the BCC lattice and the HCP lattice. 89
4.4 Surfaces of a given charge density of Al in the FCC lattice and the BCC lattice. 90
4.5 Surfaces of a given charge density of Cu in the FCC lattice and the BCC lattice. 91
4.6 Surfaces of a given charge density of Zr in the HCP lattice and the BCC lattice. 92
4.7 Surfaces of a given charge density of Zr in the hexagonal lattice and NiTi in the the BCC lattice. . . . 93
4.8 Plots of TPMS and their corresponding ZEP. . . . 94

5.1 Sketch of a continuous deformation from a mug to a torus. . . . 98
5.2 Sketch of a 2D paper in which a cat lives. . . . 99
5.3 Sketch of deforming a hollow cube to form a three-torus. . . . 101
5.4 Sketch of graphene and its corresponding lattice points. . . . 102
5.5 Sketch of lattices and their corresponding structures formed by lattice points. . . . 103
5.6 Sketch of iron in different magnetic states and how magnetism changes its symmetry. . . . 104
5.7 Sketch of how the connection between one lattice point in a simple cubic lattice and its nearest neighbours changes under a shear. . . . 105

5.8 Sketch of arms connecting one lattice point to its neighbours in a BCC lattice, an FCC lattice and a hexagonal/HCP lattice. . . . 105
5.9 Example of obtaining a skeletal graph of a TPMS by a continuous shrinking and expanding. . . . 106
5.10 Sketch of skeletal graphs (red sticks) of the P surface, the I-WP surface, the F-RD surface and the H′-T surface. . . . 106
5.11 Sketch of continuous deformations of the skeletal graph of a P surface. . . . 107
5.12 Sketch of how arms connecting one lattice point to two of its nearest neighbours change under a Bain deformation, a Nishiyama–Wassermann deformation and a Bogers–Burgers deformation. . . . 107
5.13 Sketch of nearest neighbours of a lattice point. . . . 108
5.14 Sketch of the Bravais lattices of the BCC and the FCC crystal lattices and their corresponding skeletal graphs. . . . 109
5.15 Sketch of the transformation from a hexagonal lattice to a BCC lattice. . . . 110
5.16 Sketch of hexagonal lattices and their corresponding skeletal graphs. . . . 111
5.17 Sketch of NiTi tranforming from a BCC lattice to a monoclinic lattice. . . . 111
5.18 Formation of a P surface by taking symmetry operations on its fundamental piece. 112
5.19 Sketch of the B2 NiTi lattice against the P surface from different views. . . . 112
5.20 Sketch of the crystal lattice and the Bravais lattice of B2 NiTi against the P surface and the corresponding skeletal graph. . . . 113
5.21 Sketch of the crystal lattice and the Bravais lattice of B19′ NiTi against the deformed P surface and its corresponding skeletal graph. . . . 113

6.1 Sketch of crystal lattices and the choice of lattice points. . . . 121
6.2 Sketch of the evolution of a surface of a given charge density ρ when ρ goes towards zero. . . . 127
6.3 Sketch of regions within a B2 NiTi unit cell divided by the P surface and illustration of how charges “flowing” through this surface. . . . 128
6.4 Sketch of a small piece on a surface S and a pill box enclosing a small piece of S 128
6.5 Sketch of viewing the surface piece along two orthonormal directions. . . . 129
6.6 Conclusion of results obtained in previous sections for the minimal surface modelling. . . . 132
6.7 Sketch of how lattice points of B2 NiTi locate on the P surface. . . . 135
6.8 Sketch of the Bravais lattice of the B19′ NiTi against the oPb TPMS and its corresponding skeletal graph. . . . 136
6.9 Sketch of B2 NiTi deformed under a $\langle 100\rangle_{\mathrm{B2}}\{011\}_{\mathrm{B2}}$ single-layer shear and a $\langle 100\rangle_{\mathrm{B2}}\{011\}_{\mathrm{B2}}$ bilayer shear. . . . 137

6.10 Sketch of TPMS against deformed lattice of B2 NiTi under a $\langle 100\rangle_{B2}\{011\}_{B2}$ basal shear and a $\langle 100\rangle_{B2}\{011\}_{B2}$ bilayer shear. . . . 138
6.11 Plots of the Bravais lattices of 109-B19′ NiTi and B19 NiTi against their corresponding minimal surfaces and the shear under which B2 NiTi can deform into these lattices. . . . 139
6.12 Sketch of how B2 NiTi deforms into R NiTi. . . . 139
6.13 Sketch of how the P surface corresponding to the B2 NiTi deforms into the one describing the R NiTi. . . . 140
6.14 Sketch of how B2 NiTi deforms into B33 NiTi and their corresponding minimal surfaces. . . . 140
6.15 Sketch of how B2 NiTi deforms into B33 NiTi under . . . 141
6.16 Sketch of the crystal lattice and Bravais lattice of B2 NiTi and B19 NiTi against their corresponding minimal surfaces. . . . 142

7.1 Conclusion of this thesis. . . . 149

A.1 Sketch of the homomorphism and the isomorphism. . . . 153

B.1 Sketch of a one-to-one map and an onto map. . . . 163
B.2 Sketch of the stereographic projection. . . . 166
B.3 Sketch of how points moving along the parallel and the meridian on a sphere. . 171
B.4 Stereographic projection of p on to the equatorial plane. . . . 171
B.5 Sketch of points on a Gauss map sphere and their images under a stereographic projection. . . . 173

D.1 Mapping from a circle to an eight-knot in a 2D plane. . . . 183
D.2 Sketch of a pair of spherical lunes on a sphere and three pairs of spherical lune covering a sphere. . . . 185
D.3 Sketch of a planar graph and a tree graph. . . . 186
D.4 Sketch of loops can be drawn on a sphere and a torus. . . . 188
D.5 Sketch of opening a hole (torus) around a loop C. . . . 188
D.6 Parallel transportation of a vector along a path $A \to B \to C \to D$ marked red on a sphere. . . . 190
D.7 Example of projecting a polyhedron (cube) onto a sphere. . . . 190
D.8 Planar graph formed by projecting Fig. D.7(c) onto a plane. . . . 191
D.9 Sketch of embedded Fig. D.8(d) on a torus. . . . 192

E.1 Sketch of how the direction a flag pointing into is changed when the ship is moving along a straight line and a curved line. . . . 199

List of Tables

1.1 Part list of martensites and shape memory alloys. Checkmark (resp. cross mark) means that this material is (resp. is not) a martensite or a shape memory alloy. . 15
1.2 Conclusion of proposed theories studying martensitic phase transformations and the shape memory effect. 18

2.1 Crystal systems, possible point groups corresponding to each system and basis vectors of each system. 38
2.2 Table of characters used in calculating Landau criterion and Lifshitz criterion. . 39
2.3 List of orientation relationships 48

3.1 List of ZEP and the TPMS they resemble. 58
3.2 Weierstrass representations $(G_s, \mathrm{d}h)$ of different TPMS in $\mathbb{R}^3$. 76

4.1 Comparison between XCFs, arranged according to "Jacob's ladder" proposed by J. P. Perdew [144, 145]. 81
4.2 Symmetry groups of materials studied in this work with their corresponding TPMS. SG and PG stand for space group and point group, respectively. 88
4.3 Comparison of convergence of surfaces of a given charge density calculated using DFT and their corresponding ZEP to TPMS using the minimal RMS. 95

5.1 Genus of TPMS and their corresponding materials [178, 179]. 102

6.1 Energy barriers calculated along different transformation paths. 117
6.2 Space groups and point groups of triply periodic minimal surfaces corresponding to different NiTi phases. 118
6.3 Parameters for the TPMS corresponding to phases along martensitic transformation paths of NiTi. 136

A1 List of point group operations 154
A2 Table of principal directions in different crystal systems [122, 226]. 155
A3 Matrices of the X_5^- representation [226] 160



"*By the end of the peach-blossom grove was a cave. At first the cave opening was very narrow, but after a dozen steps it opened into a flood of light. He saw before his eyes a wide, level valley, with houses, bamboos and mulberries...*"[1] —-Tao Yuanming

[1] Selected from *The Peach Colony*, translated by Lin Yutang [1].

Chapter 1

Introduction

EVIDENCE of people hardening iron by quenching that indicates the application of phase transformations in steel is documented in *Odyssey*, written 3000 years ago [2]. However, it was not until the 19th century that people began to call such phase transformations "martensitic phase transformations" and study them systematically [3]. Adolf Martens is one of the founders of metallography in the 19th century. He devoted himself to studying the structure of metals and observed a structure showing a shear-like pattern [4]. A shear is a mechanism that is important in martensitic phase transformations. Hence, martensitic phase transformations are named in honour of Martens. Generally, martensitic phase transformations are diffusionless, first-order phase transformations that involve no switch between atoms and their neighbourhoods. The initial (high temperature phase) and final (low temperature phase) states of a martensitic phase transformation are named an "austenite" and a "martensite", respectively [2,3,5,6]. If we consider solid-solid phase transformations only, martensitic phase transformations can now be characterised by spontaneous deformations of the lattice structure, realised by shears [3,5]. A shear is a relative displacement between two adjacent atomic layers along opposite directions and this mechanism is diffusionless (Fig. 1.1). More details will be provided in Sec. 2.1. One typical martensitic phase transformation is the α–γ iron phase transformation. When quenched from 912°C, the α-iron in a BCC lattice will shift rapidly to an FCC lattice (γ-iron). After intensive studies, such phase transformations

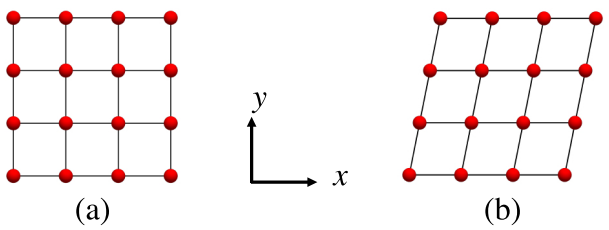


Figure 1.1 Sketch of (a) the original lattice and (b) the lattice under a shear along the x-direction between adjacent atomic layers.

were also observed in Cu-Al in 1926 [7] and in Cu-Sn in 1938 [8]. The early 1950s witnessed a bumper harvest in searching for non-ferrous martensite experimentally. Materials such as AuCd are proven to be able to undergo martensitic phase transformations [9]. Later, researchers also found martensitic phase transformations in ceramics such as zirconia and in perovskites such as $BaTiO_3$ [3, 10–18]. Additionally, in 1982, Olson and Hartman [19] treated protein transformations as displacive phase transformations, inaugurating the application of martensitic phase transformations in understanding the evolution of protein structures.

Studies of martensites shed light on the awareness of its various interesting properties, such as the shape memory effect. The shape memory effect can be understood as a "reversible" martensitic phase transformation that enables the material to return to its original shape/size when heated. For instance, as shown in Fig. 1.2, an antenna [Fig. 1.2(a)] made of NiTi crushed at room temperature into a ball [Fig. 1.2(b)] will expand back to a hemisphere [Fig. 1.2(c,d)] when heated above 77°C [20]. Different from martensitic phase transformations in steels that are irreversible, alloys such as NiTi and CuAl are observed to exhibit a shape memory effect and thus are named shape memory alloys (SMA) [3, 5, 14, 21]. A comparison of the mechanism of ordinary martensites against the mechanism of the shape memory effect is given in Fig. 1.3. Indeed, there is evidence for the shape memory effect being noticed in 1932 [13, 15, 22, 23]. However, it was in the 1950s that the discovery of the shape memory effect in alloys was declared [6, 13, 15]. Buehler and his colleagues [3, 24] discovered that NiTi is an SMA when studying heat shielding materials. Since this discovery was conducted in the Naval Ordnance Laboratory (NOL), NiTi is also named as NiTiNOL (nitinol) [3]. By the 1990s, researchers discovered another type of SMA called the magnetic shape memory alloys, such as $BaTiO_3$, whose magnetic ordering will change during a martensitic phase

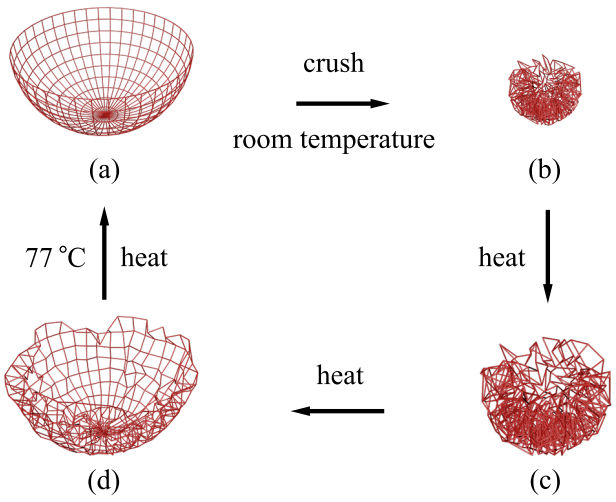


Figure 1.2 Sketch of crushing (a) an antenna to (b) a ball. After heating, (b) will gradually unfold, as shown in (c) and (d), then recover its original shape (a).

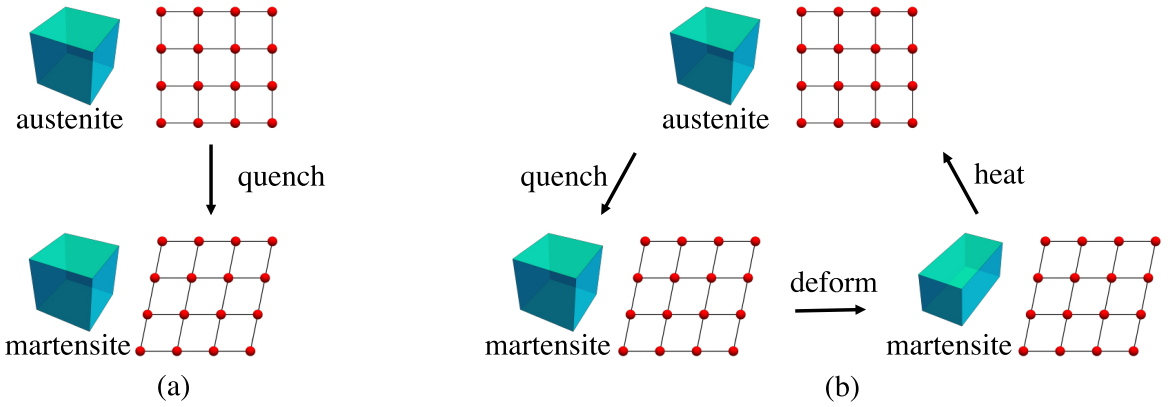


Figure 1.3 Sketch of (a) the mechanism of a martensitic phase transformation and (b) the mechanism of the shape memory effect.

transformation [12, 16, 17, 25–27]. Table 1.1 provides a list of a part of SMA and martensites.

Martensites and SMA have a broad application in industries. SMA was first applied commercially in coupling hydraulic pipes in the F-14 fighter aircraft. The manufacturer wanted to find a material to fabricate pipes lying close to the aluminium skin of the aircraft. They found pipes made of NiTi can shrink to fit the skin tightly when heated above the martensitic-transformation temperature [20]. However, before assembly, those NiTi pipes need to be stored at a temperature below $-120°C$. Getting rid of this low-temperature requirement encouraged the development of high-temperature SMA such as TiPd and TiPt, whose transformation temperature is above $100°C$ [3]. Moreover, SMA can also be used to fabricate actuators and sensors, such as the shape memory spring in a thermostatic valve. One can set the temperature by turning a knob that will change the compression exerted on a spring made of SMA. Increasing in temperature will lead the spring to expand gradually to its original shape. Hence, the valve will stop heating when the set temperature is reached. In addition to the aforementioned hardening of steel that benefits the fabrication of stainless steels, martensitic phase transformations can also contribute to the enhancement of permittiv-

Table 1.1 Part list of martensites and shape memory alloys. Checkmark (resp. cross mark) means that this material is (resp. is not) a martensite or a shape memory alloy.

Alloys	Martensite	SMA
Fe	✓	×
Zr	✓	×
Na	✓	×
NiTi, NiTiNb, NiAl TiNiCu,TiPt, TiPd	✓	✓
AgCd, AuCd	✓	✓
CuAl, CuSn, CuAuZn	✓	✓
InTl, InCd	✓	✓

ity of HfO_2 [28] and find their applications in the field of micro-/nano-scaled systems [29,30].

The tremendous potential of industrial applications of martensitic phase transformations and the shape memory effect triggers intensive studies of such phenomena both theoretically and experimentally. In 1924, Bain and Dunkirk [31] proposed the "Bain deformation" to explain the $\gamma \rightarrow \alpha$ martensitic phase transformation in steels, followed by the Nishiyama–Wassermann model proposed in 1933 [32]. Cohen *et al.* [33,34] differentiated martensitic phase transformations from other displacive phase transformations by examining whether the phase transformation is dominated by the strain energy kinetically and morphologically or not. Furthermore, Krumhansl [6,35] classified martensitic phase transformations into three different classes:

1. First-close-to-second-order martensitic phase transformations;
2. Phase transformations with appreciable elastic strain effects between martensite and austenite;
3. Phase transformations in which the dominance of shear lattice deformation leads to difficulties in introducing order parameters into the analysis.

The first class are first order phase transformations witnessing a group-subgroup (G-sub G) relation in the two end phases, such as the martensitic phase transition in NiTi that goes from the $Pm\bar{3}m$ group to the $P2_1/m$ group. Although the last two classes are reconstructive phase transformations [6,36], class 2 can still be solved in the framework of Landau's theory of phase transformations, similar to class 1. To be specific, the phase transformations belonging to class 1 should have G-sub G relations in their space groups, while for class 2 are point-group G-sub G relations. Therefore, order parameters can be found in these systems. For instance, Falk's method using strain tensor components as order parameters will only be valid in systems belonging to the first two classes [6,10]. We will elaborate the Landau's theory in Sec. 2.2.1. For class 3, it was suggested that Landau's theory should fail here. However, Tolédano and Dmitriev [37] used atomic shifts as order parameters and studied martensitic phase transformations, including class 3, in the language of Landau's theory. Cochran [38,39] also proposed a soft-mode theory, supposing the variation can not only be expanded in terms of irreducible representations, as Landau's theory, but also eigenvectors of the crystal lattice.

Besides theories focusing on general explanations, there is also interest in approaches yielding "accurate" results. Density functional theory (DFT) calculations, as well as the relationship between the phonon spectrum and phase transformations [38–40], contributed to the proposal of the first-principles alloy theory due primarily to de Fontaine [41,42]. This theory inaugurated the application of computational strategies to studying precursors of martensitic phase

transformations, *e.g.*, softening of a phonon mode. In 2009, Hatcher *et al.* [32,43] searched for possible martensitic phase transformation paths of NiTi using DFT and calculated energy barriers along each path. Table 1.2 briefly concludes the proposed theories and models that will be introduced in the next chapter. The Landau's theory, the reconstructive theory and the continuum theory are phenomenological. They can give a general explanation of martensitic phase transformations and the shape memory effect from the perspective of symmetry. The soft phonon mode theory can be applied to specific materials with the help of DFT calculations but can be time-consuming if the system is large.

Different from the methods mentioned above, Hyde and Andersson [44–46] attempted to use the concept of minimal surface to understand the martensitic phase transformation in iron. They interpreted the $\gamma \rightarrow \alpha$ Fe martensitic phase transformation as a continuous Bonnet transformation between the Schwarz D (D) surface [Fig. 1.4(d,e)] and the Schwarz P (P) surface [Fig. 1.4(a,b)]. Both P surface and D surface are minimal surfaces. According to Hyde and Andersson, a phase transformation is characterised by symmetry variations. If we locate a crystal on a surface, then the symmetry variations can all be represented by a surface

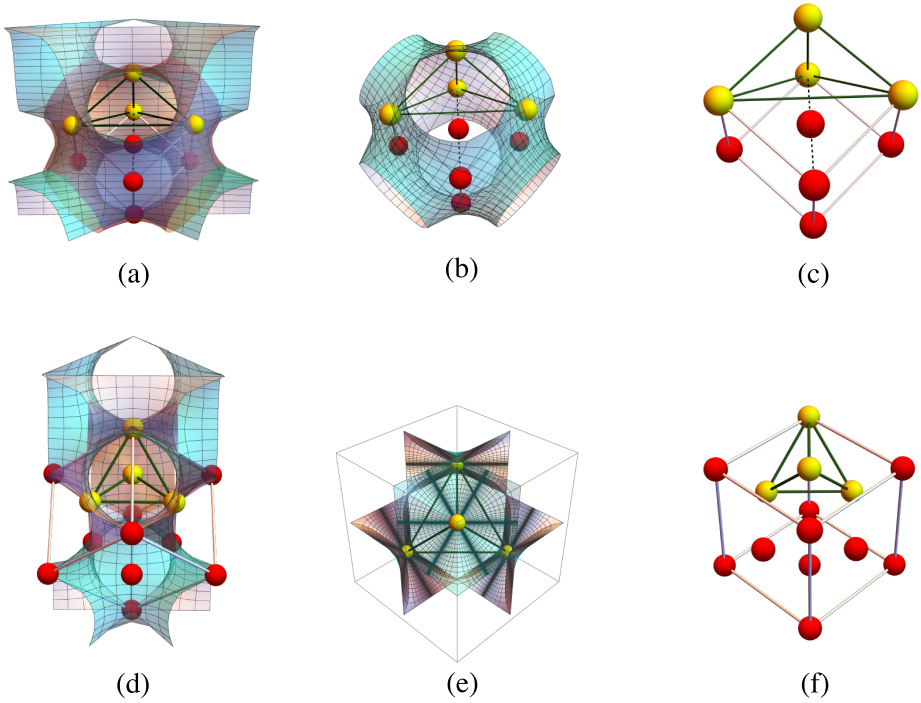


Figure 1.4 Sketch of (a,b) fitting a BCC lattice to the P surface and (c) a BCC lattice. Sketch of (d,e) an FCC lattice to the D surface and (e) an FCC lattice. Purple area in (a) (resp. (d)) is depicted in (b) (resp. (c)). (a,b) are the P surface and (d,e) are the D surface.

Table 1.2 Conclusion of proposed theories studying martensitic phase transformations and the shape memory effect.

Theory	Details
Landau's theory	1 Phase transformations are accompanied by symmetry breakings. 2 A general, phenomenological theory, does not give explanations to a specific material. 3 Applied to martensitic phase transformations with two end phases linked by a G-sub G relation.
Reconstructive theory	1 Applied to martensitic phase transformations subjected to no G-sub G relation. 2 Adopted from Landau's theory by choosing other order parameters that can vary continuously in first-order phase transformations instead of the charge density. 3 A general, phenomenological theory, do not give explanations to a specific material.
Soft Phonon Mode	1 Treating atoms as oscillators. 2 Applied mostly in studying structural phase transformations such as martensitic phase transformations. 3 A vanishing of phonon frequency can be observed at critical temperature. 4 Can give explanations to a specific material, not a phenomenological theory.
Deformation Models	1 Studying shears that can realise a phase transformation. 2 Observable. 3 Can give general explanations to lots of systems.
Continuum Theory	1 Focusing on shape memory alloys. 2 Calculating shear plane and shear direction based on the continuity of the deformation gradient at the boundary. 3 A general, phenomenological theory, do not give explanations to a specific material.

deformation. They attempted to locate the γ-Fe and α-Fe crystal lattices at flat points[1] on the D and P surface (Fig. 1.4). Moreover, Schnering and Nesper [48–50] discovered that the surface of zero charge density in an electrostatic field resembles a minimal surface that decided by the arrangement of atoms in the corresponding crystal. As inspired by Schnering and Nesper, Hyde and Andersson, we wonder that whether the introduction of electron-electron interactions will influence the convergence of the surface of equal charge density in crystals to a minimal surface, and if it is possible to build a both phenomenological and specific model to study the martensitic phase transformation and the shape memory effect using minimal surfaces.

The concept of "minimal surface" was born when studying the formation of stable soap films based on the principle of economy of nature [51]. A soap film around a specific contour is regarded to be stable when its surface free energy, proportional to the surface area, is minimised. In other words, the stability of a soap film is related to the minimisation of its surface area [51]. To be specific, the area of the formed soap film should be a local minimum of the area functional, which makes soap films one of the most vivid exhibitions of the concept "minimal surface" [51]. If a minimal surface is triply periodic, then it is named a triply-periodic minimal surface (TPMS). Here, being triply-periodic means such an object has translational symmetries along three mutually independent directions, respectively. For instance, crystals are triply-periodic. The Schwarz P (P) and the Schwarz D (D) surfaces are examples of TPMS [52, 53]. It can be seen that TPMS and crystals share similarities: 1) possessing point group and space group symmetries and 2) possessing triply-periodicity. Therefore, we can expect the crystallographic martensitic phase transformation to be studied from the perspective of TPMS, which is the main idea of this work. Our work aims to construct a theory that is both phenomenological and specific, such that this theory can not only explain martensitic phase transformations and the shape memory effect, but also can tell whether a material can undergo a martensitic phase transformation or exhibit the shape memory effect or not in a quick way. In the next section, we will provide an outline of this work.

1.1 Outline

This thesis focuses on building a model that explains the martensitic phase transformation and the shape memory effect from a geometry perspective. We introduce the physical and mathematical background in the first two chapters to help readers understand our modelling. In Chapter 2, we will give an introduction to the martensitic phase transformation and the

[1] **Flat point** [47]: points with Gaussian curvature equal to zero. More details will be given in Chapter 3.

shape memory effect, as well as the relevant physical property, shear, and deformation models based on shearing. Also, we will review the theories proposed to study the mechanism behind the martensitic phase transformation and the shape memory effect. Including Landau's theory of phase transformations, the reconstructive phase transformation theory, and a continuum theory based on the minimisation of free energy. All these theories are constructed using the group theory. Our model is not based on the group theory and can distinguish shape memory alloys from ordinary martensites. Inspired by Hyde and Andersson [54], we find that differential geometry may be a possible method.

Chapter 3 introduces the mathematical background of basic concepts we employed in building our theory. We will first introduce the zero equipotential surface in an electrostatic field calculated by Schnering and Nesper [48] to make the concept of TPMS more physical. Next, we will introduce principal curvatures on a parameterised surface, the mean curvature and Gaussian curvature, as well as the definition of the TPMS and its expression.

With the background clarified, Chapter 4 to Chapter 6 focus on the model construction. Chapter 4 aims to give a physical understanding of the TPMS by calculating a surface S of a given charge density in crystals using DFT. After an introduction to DFT, we will calculate the charge density distribution ρ using DFT, then set ρ to different values to find S. We will then calculate the shortest distance of points on S from its corresponding TPMS, hence the root-mean-square (RMS) of this data set. The deviation of S from the TPMS is thus evaluated by their corresponding RMS. By choosing different real space mesh grid densities in the DFT calculation, S converges to a TPMS. Moreover, the point group of a crystal should be a subgroup of the point group of its corresponding TPMS, and the end phases of a martensitic phase transformation can be described by TPMS having the same genus. In other words, a martensitic phase transformation is topologically continuous.

Based on the result obtained, Chapter 5 explains why a martensitic phase transformation can witness a geometric continuity. The key point is that a martensitic phase transformation is diffusionless, thus allowing no breaking of connections between lattice points. The connection between lattice points of a crystal can be represented by a skeletal graph obtained by continuously shrinking or stretching the corresponding TPMS of this crystal. No breaking in lattice point connections means no discontinuity transformation between skeletal graphs. Hence, the genus is conserved. We will use examples of BCC–FCC, BCC–HCP and B2–B19′ phase transformations to illustrate our theory.

Chapter 6 further studies the geometric property that distinguishes shape memory alloys from

martensites and other ordinary materials. Although we have shown that martensitic phase transformations are topologically continuous in previous chapters, the mechanism that differentiates shape memory alloys from ordinary martensites is not yet clear. By comparing the energy barriers along different transformation paths, it can be seen that shape memory alloys have vanishing barriers while ordinary martensites and non-martensites do not. Therefore, we assume that there are additional geometric conditions that determine whether a martensitic phase transformation is reversible or not. Before discussing this, we first analyse the value of ρ that can make S a TPMS by treating ρ as a continuous fluid and using the continuum equation. Given that the quantity we propose later is a function of ρ and we find for a TPMS, ρ should be zero. Then, we propose a term E_{g} having the dimension of energy that consists of two parts: 1) the bending energy of S and 2) the gliding of lattice points on S. When a phase transformation occurs, atomic positions change, S will thus deform accordingly, and so will the lattice points.

We employ the Helfrich Hamiltonian [55] to express the surface bending energy and find that the energy of lattice points confined on S is related to the mean curvature of S (Sec. 6.4.1). To show the "reversibility" of phase transformations, we use the term "geometric entropy", in analogy with the thermodynamic entropy, which is defined as $S_g = T|\Delta E_g|$, where ΔE_g is the variation of E_g during a phase transformation and T is the quenching temperature. We then will see when the phase transformation: 1) has no gliding of lattice points during a phase transformation and 2) can be described by a path living in the space of TPMS that connects the topological equivalent end phases, $\Delta E_g = 0$. This means $\Delta S_g = 0$, and this progress is geometrically reversible. By looking at examples of NiTi, we will see paths with $\Delta S_g = 0$ also have their energy barrier vanishing. Thus, we can conclude that the vanishing of the energy barrier, which is special to shape memory alloys, is equivalent to the martensitic phase transformation being reversible.

1.2 Publications

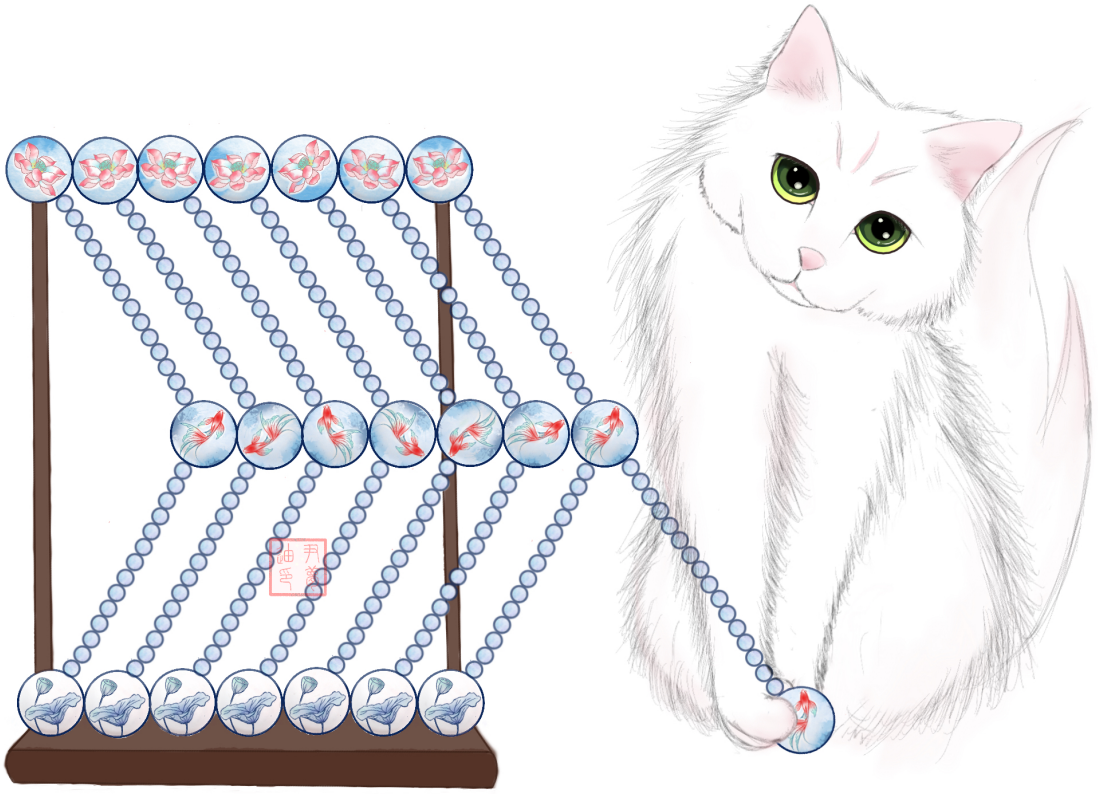
Here is the list of publications on which this work is based:

1. M. Yin and D. D. Vvedensky. Shape-Memory Effect and the Topology of Minimal Surfaces. *Symmetry*. 16(9):1187, 2024.

2. M. Yin, J. Zhang and D. D. Vvedensky. Density-Functional Theory and Triply-Periodic Minimal Surfaces. (in preparation)

3. M. Yin, D. D. Vvedensky*. Topology and Martensitic Phase Transformations. (in preparation)

4. M. Yin and D. D. Vvedensky. Mathematical Explanation to Topological Continuity of Shape Memory Alloys. (in preparation)



Sleeping like a snowball resting on the meadow,
Cavorting like a snowflake flying between the brocade curtain,
Running like a rabbit bouncing in the forest,
Capturing insects like a warrior fighting for a conquest.[2]

—- Zhu Zhongmei

[2]Selected from *Xi Jiang Yue, Ode to The White Cat*, translated by Yin Mengdi.

Chapter 2

Martensitic Phase Transformations and the Shape Memory Effect

MARTENSITIC phase transformations are displacive phase transformations realised by shear. A shear can be understood as a relative displacement between adjacent atomic layers along opposite directions parallel to these atomic layers. Models based on different shear mechanisms have been proposed to describe martensitic phase transformations. For example, the Bain deformation, the Nishiyama–Wasswemann deformation and the Bogers–Burgers deformation that aim to explain an FCC–BCC (and *vice versa*) lattice deformation. There has also been interest in modelling martensitic phase transformations using group theory. According to Landau's theory, a phase transformation is accompanied by a symmetry breaking [56]. Take the magnetic phase transformation as an example (Fig. 2.1). When the temperature is above the Curie temperature, and no external magnetic field is applied, the system is non-magnetic, because now magnetic moments will point in all directions randomly. When the temperature is lowered below the Curie temperature, magnetic moments will align in a preferred direction, which means the symmetry of magnetic moments is broken (only one direction exists), and the system will transform into a magnetic phase. The symmetry breaking is revealed by a change in the order parameter, which will become a non-zero value when the system goes from a high-symmetry phase to a low-symmetry phase. In a magnetic system, the order parameter can be chosen to be the magnetisation. The system free energy inherits the system symmetry. Hence, by expanding the free energy in terms of the order parameter around the transition point, one can find the high-symmetry phase free energy being zero at the critical temperature and becoming larger than zero when there is a symmetry breaking. Also, one can plot the free energy against the order parameter and find which phase has the lower free energy when the order parameter is tuned to a particular value. As introduced before in Chapter 1, some martensitic phase transformations having a continuous variation in the charge density can be studied in the framework of Landau's theory, with

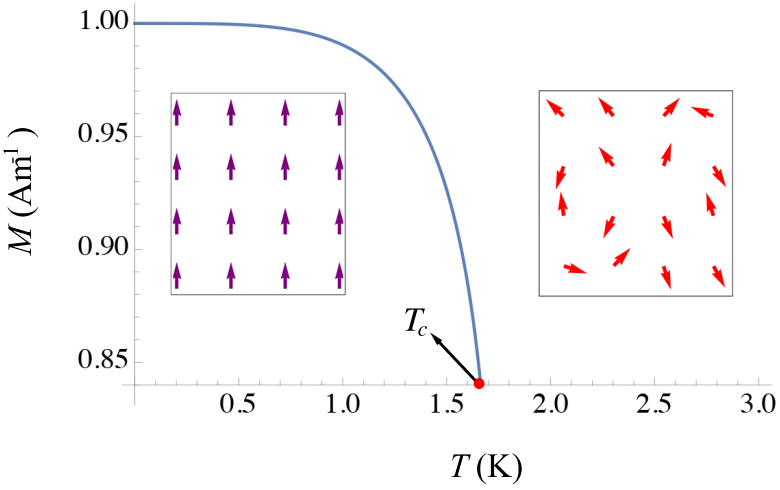


Figure 2.1 Sketch of the phase diagram of phase transformation of a magnetic system. Below the Curie temperature, the magnetisation is non-zero, as shown by the purple arrows. When heated above the Curie temperature T_c, the magnetisation of the system becomes zero, and magnetic moments will point in various directions, as shown by the red arrows.

the order parameter chosen to be the charge density. We will explain this later in Sec. 2.2.1. The rest of martensitic phase transformations, however, can still be solved using an adapted Landau's theory, *i.e.*, the reconstructive theory. This theory was proposed by Tolédano and Dimitriev [37] and will be introduced in Sec. 2.2.2.

The soft phonon mode is a phenomenon observed during structural phase transformations, stating that when the critical temperature is approached, a phonon frequency will vanish [38]. In addition to martensitic phase transformations, such a phenomenon is also observable in other phase transformations such as the ferroelectric phase transformation [57, 58] and the charge density wave phase transformaiton [59, 60]. Although Ginzburg and Levanyuk [38, 61–63] explained this phenomenon in the framework of Landau's theory, Cochran [39], however, built a lattice dynamic model by treating ions as oscillators to explain the soft phonon mode. Cowley [64,65] proposed his microscopic anharmonic model by introducing a damping term and a driving force. This model also shows that the phonon frequency should vanish at the critical temperature when a structural phase transformation takes place.

In addition to Landau's theory, Bhattacharya [5] proposed a theory that can study the shear and variants of a shape memory alloy. This theory is based on the minimisation of free energy, which is a function of the deformation gradient (caused by a shear) and temperature. According to Bhattacharya, below the critical temperature, all phases that can co-exist should have the same free energy and the same normal vector at the phase boundary if they are adjacent. Since different phases are formed by different deformation gradients, the same free

energy means these gradients are symmetry-related to each other. Therefore, by performing symmetry operations on one deformation gradient, we can find all possible phases that can co-exist in a crystal, which are variants of a martensite. The shear is then calculated using the normal vector shared by two phases possessing different gradients at their common boundary. Details about this theory will be introduced in Sec. 2.2.5.

2.1 Definition of Martensite and Shape Memory Effect

As mentioned, martensitic phase transformations are defined as diffusionless, first-order phase transformations [5,6,21]. The earliest classification of martensitic phase transformations was given by Cohn in 1979 [33] based on energetic and dynamical factors. According to Cohn, a martensitic phase transformation is described as a displacive phase transformation with 1) undistorted lines in the body and 2) its kinetics and morphology being dominated by strain energy but not shuffle displacements. In 1989, however, Krumhansl [6,21] viewed martensitic phase transformation from the perspective of symmetry characteristics. He suggested that, in general, there are two types of martensitic phase transformations, *viz.*, those can be explained in the framework of Landau's theory of phase transformations, and those that cannot. The first type can be categorised further into two classes:

1. first-to-second phase transformations during which the space groups (SGs) of the *parent* (austenite) phase and the *daughter*(martensite) phase are related by a group-subgroup (G-sub G) relation;

2. phase transformations have G-sub G relations not in their SGs but in their point groups (PGs).

Hence, martensitic phase transformations belonging to this type can be studied using Landau's theory. For the second type, however, no G-sub G relation can be found to relate either the SGs or PGs of the parent phase and the daughter phase to each other. According to Landau's theory [37,38,56,66], we need first to find proper order parameters that inherent all symmetry information of the parent phase, then do a Taylor expansion of the free energy in terms of these order parameters around zero. Here, in Landau's theory, the order parameters are set to be zero for the parent phase. For instance, if we set the atomic displacement to be the order parameter, then in the parent phase, we see that the atomic displacement, hence the order parameter, is zero.

Macroscopically, martensitic phase transformations are considered to be completed by shear,

with an interface between the parent phase and the daughter phase such that the average lattice distortion on it is zero [5]. There are two properties of shear to be considered: shear strain ε and shear stress σ. Generally speaking, stress can be understood as force per unit area exerted on an object, while strain ε is the ratio of variation in length caused by force to original length [67]. In real 3D space, we usually express both stress and strain in a 3×3 tensor form. But for the shear itself, we are more interested in its direction and the *invariant plane*, *i.e.*, **the plane that remains unchanged under a shear and parallel to which the atomic displacements take**. Hence a shear can be written using the direction of the shear and the invariant plane, in the form $[uvw](hkl)$ (or equivalently, $\langle uvw \rangle \{hkl\}$), where $[uvw]$ are the *direction indices* [1] that give the shear direction and (hkl) are the *Miller indices* [2] that give the direction of the invariant plane (or equivalently, *habit plane*).

Experimentally, an invariant plane that performs as the phase boundary can be observed, from which we can conclude the *orientation relationships* that gives **relations between planes and directions of the two end phases** (examples are given in Sec. 2.2.4). **Martensites co-existing in the same daughter phase having different orientations** are said to be different *variants* of the martensite [5]. To understand this, we use the concept of the *isogonal group*. *Isogonal group* is defined as **the group formed by all point group operations presented in all operations of the corresponding space group** [68]. More details are in Appendix A.1.3. If we assume that the *isogonal group* P_m of the daughter phase is a subgroup of the *isogonal group* P_a of the parent phase, by performing operations in P_a onto the transformation matrix describing the corresponding shear[3] of one variant, we can obtain the transformation matrices of other variants. Figure 2.2 gives an example of how variants are symmetry-related. Let the austenite phase have a cubic lattice and the martensite phase have a tetragonal lattice, then by elongating the cubic lattice along different directions, *i.e.*, x, y, $z-$axis, we will obtain three variants [Fig. 2.2(b,c,d), accordingly]. Rotating Fig. 2.2(b) along $z-$axis by 90°, we will obtain the variant given in Fig. 2.2(c). Hence, variants are symmetry-related to each other and the number of variants is calculated by $|P_a|/|P_m|$ [5].

The shear has two different mechanisms: 1) twinning [Fig. 2.3(c)], 2) slip [Fig. 2.3(b)], among which twinning is responsible for a reversible deformation while the slip is responsible for a plastic strain and is thus a permanent deformation [24, 69]. Additionally, there are several different types of shear, among which *simple shear* and *pure shear* are mostly applied. *Simple*

[1] **Direction indices**: indices expressing vectors in the basis of lattice vectors of a Bravais lattice corresponding to a crystal, written in square brackets.

[2] **Miller indices**: indices describing a lattice plane obtained by calculating the inverse intercepts of this plane with the lattice vectors.

[3] Details are provided in Sec. 2.2.5

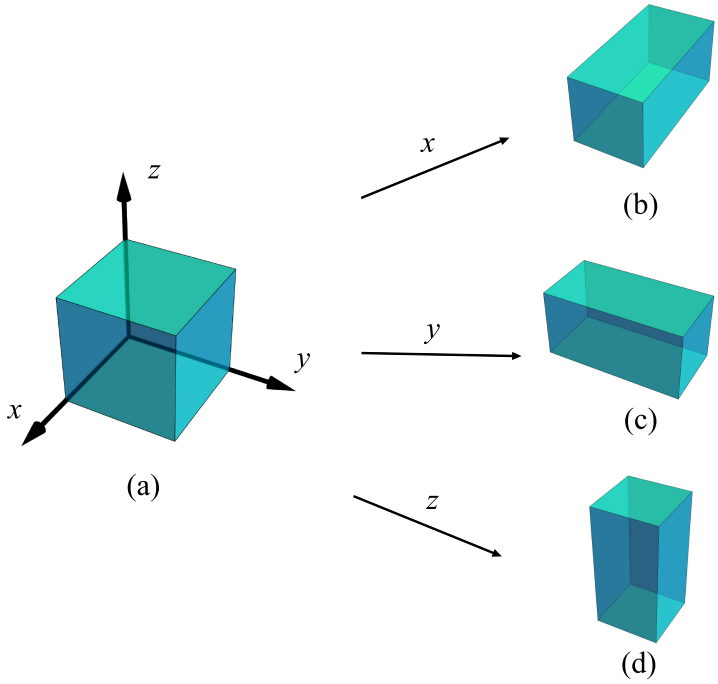


Figure 2.2 Sketch of how variants are symmetry-related in a cubic to tetragonal lattice deformation. (a) The original cubic lattice. (b) The tetragonal lattice obtained by elongating (a) along the x direction. (c) The tetragonal lattice obtained by elongating (a) along the y direction. (d) The tetragonal lattice obtained by elongating (a) along the z direction.

shear can be understood as a sliding of atoms lying in the same atomic layer along a direction parallel to this layer [Fig. 2.4(b)]. A *pure shear*, however, can be understood as a deformation caused by equal but opposite forces along orthogonal directions [Fig. 2.4(c)].

If a martensitic phase transformation is reversible, or equivalently, thermoelastic, we say that the shape memory effect occurs. The shape memory effect is a phenomenon in which an alloy can recover its original shape when heated after plastic deformation. Materials exhibiting a shape memory effect are said to be, in addition to shape memory alloys, "thermoelastic". This effect was first observed in Au-Cd [23] and In-Tl [70] alloys. Although some disordered

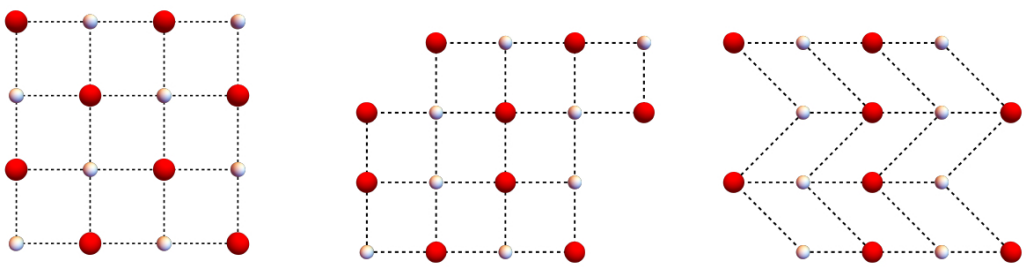


Figure 2.3 Shear mechanisms of martensitic phase transformations. (a) Austenite phase, (b) slip, (c) twinning.

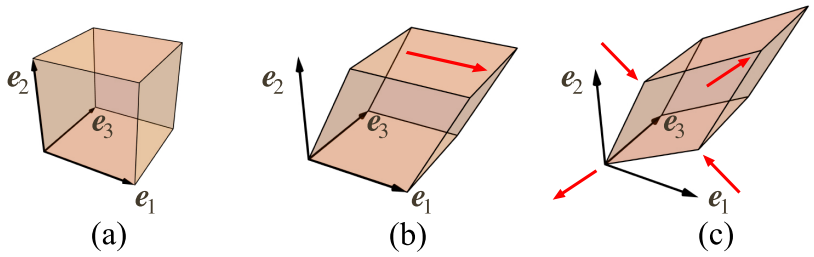


Figure 2.4 Sketch of a (a) cube deformed under (b) a simple shear and (c) a pure shear. In (b), the red arrow shows the direction of the shear and in (c), the red arrows show the direction of forces applied.

alloys, such as Mn-Cu, can exhibit the shape memory effect, this effect occurs mainly in ordered alloys because the atomic disorder will make it difficult to restore the original atomic order in the parent phase [24].

Martensitic phase transformations and shape memory alloys can have various applications in industries. For instance, martensitic stainless steels, due to their good corrosion resistance and high hardness, are applied widely in chemical plants and the oil and gas industry [71]. The shape memory alloy NiTi is used in pipe coupling, such that NiTi is stored at a deformed state and then installed by heating to its austenite state [24]. Moreover, in the 1970s, NiTi was also used to make orthodontic treatment devices since it can generate constant force when recovering its original shape during the positioning [72,73]. Shape memory alloys are also used to fabricate self-expandable vascular stents [72,74,75] and venous "Simon filters" [76,77], as well as coils and microguidewires in the neurosurgery [72]. A vascular stent is a tubular mesh used to treat narrowing in arteries to ensure that oxygen-rich blood can be provided to the heart. During surgery, the stent will be crimped into a catheter using a sheath and then inserted into the vessel. Since the transformation temperature is lower than the body temperature, the stent will expand to its original shape after removing the sheath. A "Simon filter" is a device designed to prevent pulmonary embolism [78]. It straightens at low temperature (4°C–10°C) and will recover its designed shape, *i.e.*, an umbrella dome of seven loops connected to six legs, at body temperature [78]. Microguidewires are applied in positioning stents, due to its high strain recovery that can reduce the wire bending problems [72].

2.2 Proposed Theories and Models

2.2.1 Landau's Theory of Crystal Phase Transformations

The whole discussion using Landau's theory of phase transformation pivots around two properties: (1) breaking of symmetry and (2) variation in the order parameter. That is to say, the

appearance of a new phase is accompanied by the vanishing of symmetry elements that can be measured by the variation in some physical properties [79]. Such physical properties are chosen to be *order parameters*, depending on the system. For example, atomic displacements, strain tensor components, *etc*. According to Landau's theory, the difference between a first-order and a second-order phase transformation is that the former can witness a discontinuity in its order parameter variation while the latter cannot. Taking a Taylor expansion of the system free energy φ around the critical point, we see:

$$\varphi = \varphi_0 + \alpha_1\eta + \alpha_2\eta^2 + \alpha_3\eta^3 + \alpha_4\eta^4 + \cdots, \tag{2.1}$$

where η is the order parameter, φ_0 is the free energy of the original phase (or, ordered phase) and α_i, $i = 1,2,3,4,\cdots$ are parameters relying on the environment (such as temperature). φ_0 is usually set to be zero. Additionally, α_2 depends linearly on temperature and will vanish at the critical temperature. When away from the critical temperature, α_2 is non-zero. According to the stability condition $\partial\varphi/\partial\eta = 0$ at the critical point, we have $\alpha_1 \equiv 0$. Ignoring terms of order higher than 4 because they are too small, we see that at the critical temperature, system stability requires:

$$\frac{\partial\varphi}{\partial\eta} = 0, \qquad \frac{\partial^2\varphi}{\partial\eta^2} \geq 0. \tag{2.2}$$

The second equation is to ensure the minimum and when the equal sign is taken, it means the the phase at critical point is stable. Plugging Eq.2.1 into the first equation in Eq. 2.2, we see:

$$3\alpha_3\eta^2 + 4\alpha_4\eta^3 = 0, \tag{2.3}$$

the two solutions, $\eta = 0$ is for the parent phase and $\eta \propto \alpha_3/\alpha_4$ is for the daughter phase. Given the second-order phase transformation requires no jump in η at the critical point, we have $\alpha_3 = 0$, which means a vanishing of the η^3 term. **The requirement that cubic terms η^3 should vanish at the critical point** is the *Landau criterion* for **second-order phase transformations**. It is noticeable that according to the Landau criterion, we have the free energy as a function of the η but not the gradient of η, which means it is assumed that the ordered phase is homogeneous. Lifshitz [79] took a step further by looking at under which condition the ordered phase can be homogeneous. He considered $\varphi = \varphi(\eta, \nabla_r\varphi)$ and discovered that the inhomogeneity occurs during a phase transformation that can be characterised by the *Lifshitz invariant*, which is a linear term defined as **an anti-symmetric combination of the partial differential of η to coordinates** [80]:

$$\eta_i\frac{\partial\eta_j}{\partial x_p} - \eta_j\frac{\partial\eta_i}{\partial x_p}, \tag{2.4}$$

where $i, j = 1, 2, 3, \cdots, d$, d is the dimension of the order parameter and $p = 1, 2, 3$. And the **Lifshitz criterion** is defined as (Einstein summation convention is used here) [56, 80]:

$$A_{i,j,p} \int \mathrm{d}\boldsymbol{r} \left(\eta_i \frac{\partial \eta_j}{\partial x_p} - \eta_j \frac{\partial \eta_i}{\partial x_p} \right) = 0, \qquad A_{i,j,p} = \frac{\partial^2 \phi}{\partial \eta_i \partial (\partial_{x_p} \eta_j)}, \tag{2.5}$$

ϕ is the free energy density such that the total free energy $\varphi = \int \phi \, \mathrm{d}\boldsymbol{r}$. More details of the corresponding derivation is in Appendix A.2. Hence, we find *second-order phase transformations* can be described as **phase transformations satisfying the Landau criterion and the Lifshitz criterion**.

Explanation to Continuity in Order Parameter

With the help of the Landau criterion and the Lifshitz criterion, we are able to find possible low-symmetry phases of a given high-symmetry phase. Now we would like to explain the G-sub G relation required by a second-order phase transformation with the help of the representation theory (Appendix A.1.2). A representation $d_\uparrow$ of G_0 can be *induced* by a representation d acting on a vector space V_d of $G \subset G_0$. An *induced representation* is defined as follows [68]. Given a group G_0 and its subgroup G, then let us label the corresponding coset representatives by $\{g_i^0\}$, $i = 1, 2, \cdots, |G_0|/|G|$ (according to the *Lagrange's theorem*, see Appendix A.1.1). Let d be a representation of G and let its basis be denoted by $\{\boldsymbol{e}_j\}$, $j = 1, 2, \cdots n_d$, n_d is the dimension of d, then we have:

$$h \circ \boldsymbol{e}_i = \sum_{j=1}^{n_d} \boldsymbol{e}_j d_{ij}(g), \qquad \forall h \in G. \tag{2.6}$$

Let us write:

$$\boldsymbol{e}_{qi} = g_q^0 \boldsymbol{e}_i, \qquad g_q^0 \in \{g_i^0\}. \tag{2.7}$$

Then the space $V_\uparrow^d$, whose basis is written as $\{g_i^0 \boldsymbol{e}_j\}$, is invariant under G_0, which means this space corresponds to an n-dimensional representation $d_\uparrow$ of G_0, $n = n_d(|G_0|/|G|)$. $d_\uparrow$ is called an *induced representation* of d in G_0, satisfying:

$$g\boldsymbol{e}_{qi} = g g_q^0 \boldsymbol{e}_i = g_k^0 h \boldsymbol{e}_i = g_k^0 \sum_{j=1}^{n_d} \boldsymbol{e}_j d_{ij}(h) = \sum_{j=1}^{n_d} \boldsymbol{e}_{kj} d_{ij}(h) = \sum_{j=1}^{n_d} \boldsymbol{e}_{kj} d_{\uparrow kj,qi}(g). \tag{2.8}$$

Here we use the rearrangement theorem (Appendix A.1.1). Take the 3D space S as an example (Fig. 2.5). Let us divide this 3D space into 2 subspaces U, a 2D plane, and W, an axis perpendicular to U. Hence, $S = U \oplus W$. The basis vector of W is $\boldsymbol{e}_z$ and the basis vectors of U are $\{\boldsymbol{e}_x, \boldsymbol{e}_y\}$. Within W, we can take symmetry operations such as C_{2z} (a 2-fold rotation around the z axis, more details are in Table A1) and M_z (a mirror reflection about the plane

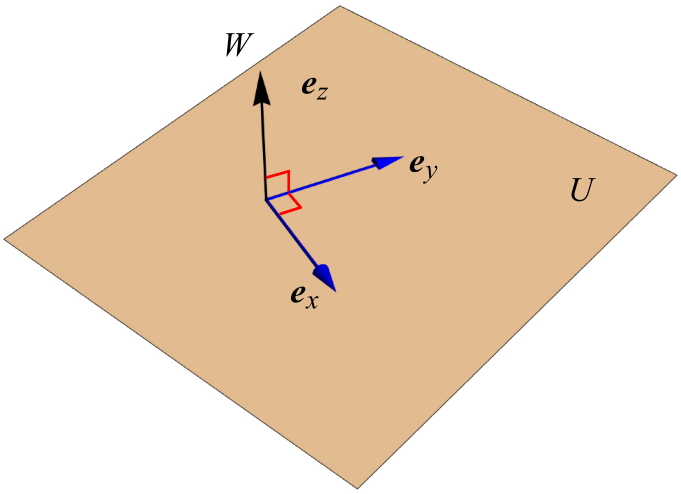


Figure 2.5 Sketch of a 3D space divided into the direct sum of 2 subspaces W and U, where U is a 2D plane whose basis vectors are $\{e_x, e_y\}$ and W is a 1D space whose basis vector is e_z. $\{e_x, e_y, e_z\}$ are orthonormal basis vectors of the 3D space.

whose normal direction is along the z axis, more details are in Table A1) that will not bring vectors inside W out of W or vectors out of W into W. However, if we perform C_{4x} (C_{4y}), a 4-fold rotation around the x (resp. y) axis, onto W, this operation will bring e_z to e_y ($-e_x$). Hence, a C_{4x} (C_{4y}) operation will bring e_z out of W. If we want to construct a space invariant under C_{2z}, C_{4x} and C_{4y} symmetries, this space should include vectors e_x, e_y, e_z, which satisfy:

$$C_{2z}e_z = e_z, \qquad C_{4x}e_z = e_y, \qquad C_{4y}e_z = -e_x. \tag{2.9}$$

Hence, a space equal to:

$$e_z \oplus C_{4x}e_z \oplus C_{4y}e_z = W \oplus U = S, \tag{2.10}$$

is invariant under C_{2z}, C_{4x} and C_{4y}.

Extending Eq. 2.10 to our discussion about V^d and $V^d_\uparrow$, we see that:

$$V^d_\uparrow = \oplus_j (g^0_j \circ V^d), \qquad g^0_j \in \{g^0_i\}. \tag{2.11}$$

Hence, by performing G_0 onto the basis $\{e_j\}$ of V^d, we have a set of vectors expanding a vector space $V^d_\uparrow$ invariant under G_0, on which the representation $d_\uparrow$ induced by d acts. For instance, in the example given in Fig. 2.5, we can take W invariant under a group $G = \{E, C_{2z}\}$ as V^d. Let $G_0 = \{E, C_{4x}, C_{2x}, C^3_{4x}, C_{2z}, C_{2y}, C_{2,011}, C_{2,01\bar{1}}\}$, $G \subset G_0$. The space $S = U \oplus W$ is induced by operating G_0 on V^d. Explanation to those notations are given in Appendix A.1.3.

Now we can explain the continuity in ρ, the charge density, using the concept of induced representation. Intuitively, it can be understood as follows. Let G be the group of the daughter

phase with charge density ρ and G_0 be the group of the parent phase with charge density ρ_0. Hence, ρ is invariant under G and ρ_0 is invariant under G_0. Given ρ_0 has all symmetries of G_0, then we can find a vector space V_0 that is invariant under G_0 and express ρ_0 using basis vectors of this space. If there is a G-sub G relation between G and G_0, saying $G \subset G_0$, then we can find a subspace V^G of V_0 that is invariant under G. Also, ρ can be expressed using vectors in V^G. Because $G \subset G_0$, $V^G \subset V_0$, ρ can be expressed using vectors in G_0. Therefore, the variation from ρ_0 to ρ is continuous given they are contained in V_0. Because V^G is an invariant subspace of V_0 under G, we can obtain basis vectors of G by taking a projection of vectors in V_0 onto V^G. Such a projection can be done by referring to an irreducible representation (IRREP) of G_0. More details about projection operators can be found in Ref. [68, 81, 82]. The reason why we use IRREP here is because we do not need vectors that do not have the symmetry of G_0. If there is no such a G-sub G relation, then the vector space V^G which is invariant under G always contains vectors that is not invariant under G_0, hence not contained in V_0. Thus, the additional vectors make it impossible for us to find a projection that can bring ρ_0 to ρ, which means the variation from ρ_0 to ρ is discontinuous. Also, given $V^G \subset V_0$, the corresponding IRREP of G_0 cannot be trivial, otherwise we will see no variation in ρ_0 and the choice of the trivial IRREP will conflict with the Landau criterion (will be introduced later in Sec. 2.2.1).

A more mathematical explanation is given below. Let $G \subset G_0$. Let us construct a vector space V^G invariant under G spanned by basis $\{\boldsymbol{e}_j^G\}$, $j = 1, 2, \cdots, \dim V^G$, where $\dim V^G$ is the dimension of V^G. Then we see that ρ can be expressed in terms of $\{\boldsymbol{e}_j^G\}$. Performing G_0 on the basis of V^G (*i.e.*, $\{\boldsymbol{e}_j^G\}$), we now have a vector space $V_\uparrow^G$ that is invariant under G_0. Hence, ρ_0 can be expressed using vectors in $V_\uparrow^G$. Because $V^G \subset V_\uparrow^G$, both ρ and ρ_0 can be expanded using vectors in $V_\uparrow^G$. Thus, we see that in $V_\uparrow^G$, the variation $\delta\rho$ can be written as $\delta\rho = \rho - \rho_0$ and is continuous. If there is no G-sub G relation between G_0 and G, then we cannot construct a space invariant under G_0 using V^G, which means ρ_0 and ρ exist in two mutually independent vector spaces. Hence, $\delta\rho = \rho_0 - \rho$ is discontinuous in this case. Moreover, for a phase transformation holds no G-sub G relation, there must be vectors describing ρ that is not invariant under G_0. Hence, we see no continuity in the variation in ρ. Therefore, a requirement of order parameters being continuous means that in a second order phase transformation, the two end phases are related by a G-sub G relation.

Application of Landau's Theory and Example

Since $\delta\rho$ transforms according to an IRREP D of G_0 (SG of the parent phase), we can write $\delta\rho$ as a combination of basis vectors $\{\psi_i\}$ of D, such that:

$$\delta\rho = \sum_i \eta_i \psi_i, \qquad i = 1, 2, \cdots, n_D, \tag{2.12}$$

where n_D is the dimension of D. The parameters $\{\eta_i\}$ are the chosen order parameters; when all of them go to zero, then the phase is the parent phase. Given $\{\psi_i\}$ are invariant under $G \subset G_0$, performing an operation $g \in G_0$, $g \notin G$, $\{\psi_i\}$ will be transformed to $\{\psi_i'\}$, such that:

$$g \circ \delta\rho = \sum_i \eta_i \psi_i'. \tag{2.13}$$

Equivalently, we can fix $\{\psi_i\}$, then g will transform $\{\eta_i\}$ to $\{\eta_i'\}$. For instance, let $i = 1, 2$, $g \circ \psi_1 = \psi_1$, $g \circ \psi_2 = \exp(\mathrm{i}\pi/4)\psi_2$, then fix $\{\psi_i\}$, we will see:

$$\begin{aligned} g \circ \delta\rho &= \eta_1(\psi_1) + \eta_2[\exp(\mathrm{i}\pi/4)]\psi_2, \\ &= (\eta_1)\psi_1 + [\exp(\mathrm{i}\pi/4)\eta_2]\psi_2. \end{aligned} \tag{2.14}$$

Hence, all η_i can form a vector in an *order parameter spaces* V_η that has the symmetry of $\delta\rho$. The dimension of V_η is also n_D. The action of elements in G_0 on vectors in V_η can be treated as a **geometric transformation** that preserves the origin of the space and the length of the vector. Moreover, the IRREP of G_0 according to which $\delta\rho$ transforms is said to be "*active*", as long as this **IRREP satisfies both the Landau criterion and the Lifshitz criterion** [66, 80]. The concept of active representation plays a vital role in studying second-order phase transformations and first-to-second martensitic phase transformations. From the perspective of group theory, one can tell whether an IRREP is active or not by calculating its symmetric and anti-symmetric powers (details are in the Appendix A.3) [83]. Here, $[D]^n$ and $\{D\}^n$ are the symmetric and anti-symmetric nth power of the IRREP D, respectively. D_1 is the identity representation (which is trivial), V is the *vector representation* of G_0. A *vector representation* V of a space group G_0 is defined as **a representation whose carrier space is spanned by performing operations in the isogonal group of G_0 on a vector $\boldsymbol{r} \in \mathbb{R}^3$** [80]. Hence, V is a 3D representation.

1. Landau criterion requires no identity representation D_1 contained in $[D]^3$, the symmetric third power of the IRREP D, *i.e.*, $D_1 \not\subset [D]^3$, which means we have, in terms of character (Appendix A.1.2):

$$\sum_g \left[\frac{1}{6}\chi^3(g) + \frac{1}{2}\chi(g^2)\chi(g) + \frac{1}{3}\chi(g^3)\right] = 0, \tag{2.15}$$

2. Lifshitz criterion requires no identity representation D_1 contained in the tensor product of V, the vector representation, and $\{D\}^2$, the anti-symmetric second power of the

IRREP D, *i.e.*, $D_1 \not\subset V \otimes \{D\}^2$, which means we have, in terms of the characters (Appendix A.1.2):

$$\sum_g \left\{ \frac{1}{2} \chi_V(g) [\chi^2(g) - \chi(g^2)] \right\} = 0. \tag{2.16}$$

The sum runs over elements in the isogonal group of *group of* $\boldsymbol{k}$ at $\boldsymbol{k}_0$ in the Brillouin zone, and *group of* $\boldsymbol{k}$ is defined as a set of **space group operators that transform $\boldsymbol{k}$ to itself or an equivalent vector $\boldsymbol{k}+\boldsymbol{K}$**, where $\boldsymbol{K}$ is a reciprocal lattice vector [81]. $\boldsymbol{k}_0$ is a high-symmetry point (details are in Ref. [80]). The key to solving second-order phase transformations using Landau's theory is in expanding the free energy around the critical point (order parameter vanishes) using polynomials invariant under the daughter phase (low-symmetry phase) symmetries as order parameters; then, the phase diagram can be drawn by plotting the free energy against the order parameter. Briefly speaking, the free energy expansion can be realised by: 1) finding symmetries remaining in the low-symmetry phase, 2) finding polynomials of order parameters invariant under these symmetries, and 3) writing the free energy as a sum of these polynomials. In step two, the symmetries are selected by looking at IRREPs of G_0. Here is the specific procedure of searching for the expansion of the free energy and possible low-symmetry phases [80]:

1. Choosing active representations from all IRREPs of G_0 according to the Landau criterion and the Lifshitz criterion.

2. Considering one of these active representations, D, whose dimension is n_D. Finding **all distinct matrices of D, which form a subgroup of** $O(n_D)$ (group of all $n_D \times n_D$ real orthogonal matrices). This set of matrices forms an *image group* I_0 of G_0 [84] and is isomorphic to a group of rotation operations that preserves the space origin and the vector length. Choosing a vector $\boldsymbol{\eta}$ in the order parameter space V_η whose dimension is also n_D. η is invariant under the image group of the low-symmetry phase. **The complete set of elements in (the point group that is isomorphic to) I_0 that leaves η invariant** is defined as the *isotropy group* of $\boldsymbol{\eta}$, labelled by I_η.

3. The free energy is expanded as following:

   - Finding the *normal series* of the group I_0, which is **an invariant subgroup array**:

   $$E = I_n \trianglelefteq \cdots \trianglelefteq I_3 \trianglelefteq I_2 \trianglelefteq I_0, \tag{2.17}$$

   where E is the identity. I_m is an invariant subgroup of I_{m-1}, $m = 2, 3, \cdots, n$ and I_2 is an invariant subgroup of I_0. Beginning with I_n, let us find a set of polynomials consisting of components of $\boldsymbol{\eta}$ that are invariant under I_n, and next use this set of polynomials to find a new set of polynomials invariant under I_{n-1}. Repeating

this process until we arrive at the set of polynomials invariant under I_0. Also, according to Noether's theorem [85], the maximum order of these polynomials should not exceed the order of the symmetry group under which they should be invariant.

- Expanding free energy in terms of the polynomials obtained in the previous steps, using stability conditions to find conditions for parameters of each term:

$$\begin{aligned} &\frac{\partial \varphi}{\partial \eta_i} = 0, \qquad i = 1, 2, \cdots, n_D, \\ &\mathsf{H} = \begin{pmatrix} \frac{\partial^2 \varphi}{\partial \eta_1 \partial \eta_1} & \frac{\partial^2 \varphi}{\partial \eta_1 \partial \eta_2} & \cdots & \frac{\partial^2 \varphi}{\partial \eta_1 \partial \eta_{n_D}} \\ \frac{\partial^2 \varphi}{\partial \eta_2 \partial \eta_1} & \frac{\partial^2 \varphi}{\partial \eta_2 \partial \eta_2} & \cdots & \frac{\partial^2 \varphi}{\partial \eta_2 \partial \eta_{n_D}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \varphi}{\partial \eta_{n_D} \partial \eta_1} & \frac{\partial^2 \varphi}{\partial \eta_{n_D} \partial \eta_2} & \cdots & \frac{\partial^2 \varphi}{\partial \eta_{n_D} \partial \eta_{n_D}} \end{pmatrix} \succ 0, \end{aligned} \tag{2.18}$$

  where $\mathsf{H} \succ 0$ means the matrix H is positive definite.

4. For finding the space group of the low-symmetry phase, we need to:

   - Determine the point group (isogonal group) with the help of I_η;
   - Determine the Bravais lattice with the help of the point group;
   - Determine the non-primitive translations.

   The isogonal group $\tilde{G}$ of the low-symmetry phase is given by:

$$\tilde{G} = I_\eta \otimes \tilde{K}, \tag{2.19}$$

   where $\tilde{K}$ is the isogonal group of the kernel of the homomorphism D (Appendix A.1.2), *i.e.*, operations of the G_0 that are mapped to the identity by the representation D. Then we can determine the Bravais lattice using $\tilde{G}$. Correspondence of point groups, Bravais lattices and basic translations are given in Table 2.1. Since in the Bloch wave basis a translation will contribute a phase factor [81], then an invariance of the free energy under a translation $\boldsymbol{t}$ requires $\exp(\mathrm{i}\boldsymbol{k}_j^* \cdot \boldsymbol{t}) = 1$, where $\boldsymbol{k}_j^*$, $j = 1, 2, \cdots, s$ is the star of $\boldsymbol{k}$. The star of $\boldsymbol{k}$, $\boldsymbol{k}^*$, is defined as **a set of k vectors that are obtained by carrying out all point group operations of a space group. Here, k is a general vector in reciprocal space and vectors belonging to the same k^* are inequivalent to each other** [81]. Elements in $\boldsymbol{k}^*$ are called *arms* of the star [81]. Therefore, the invariance of free energy under a translation in the Bloch wave basis requires:

$$\exp(\mathrm{i}\boldsymbol{k}_j^* \cdot \boldsymbol{t}) = 1 \Rightarrow \boldsymbol{k}_i^* \cdot \boldsymbol{t} = 2\pi n, n \in \mathbb{Z}, i = 1, 2, \cdots, s, \tag{2.20}$$

to be satisfied by possible translational symmetries of the new phase. A relevant example is given below and the illustration is provided in Fig. 2.6.

As for first-order phase transformations, it is the reconstructive group theory that was developed upon the Landau theory to address these problems. This will be discussed in the next section. Here, we provide an example to explain the method stated above. Consider a simple cubic lattice in the $Pm\bar{3}m$ SG [as shown in Fig. 2.6(a), more details of the symbol are given in Appendix A.1.3]; the basis vectors of its primitive cell in Cartesian coordinates are:

$$\boldsymbol{a}_1 = a(1,0,0), \qquad \boldsymbol{a}_2 = a(0,1,0), \qquad \boldsymbol{a}_3 = a(0,0,1), \tag{2.21}$$

and the basis vectors of its reciprocal lattice are:

$$\boldsymbol{b}_1 = \frac{2\pi}{a}(1,0,0), \qquad \boldsymbol{b}_2 = \frac{2\pi}{a}(0,1,0), \qquad \boldsymbol{b}_3 = \frac{2\pi}{a}(0,0,1), \tag{2.22}$$

a is the lattice parameter. Take the point $\boldsymbol{k} = \frac{1}{2}\boldsymbol{b}_1$ in the Brillouin zone (X point). Now, the corresponding isogonal group of the group of $\boldsymbol{k}$ is D_{4h}, and the star of $\boldsymbol{k}$ has three arms: (i) $\boldsymbol{k} = \frac{2\pi}{a}(0,\frac{1}{2},0)$; (ii) $\boldsymbol{k} = \frac{2\pi}{a}(\frac{1}{2},0,0)$; (iii) $\boldsymbol{k} = \frac{2\pi}{a}(0,0,\frac{1}{2})$. Then, according to Eq. 2.20, possible translation groups of the low-symmetry phase can be generated by, for instance, $\boldsymbol{v}_1 = 2a(1,\frac{1}{2},\frac{1}{2}), \boldsymbol{v}_2 = 2a(1,1,\frac{1}{2}), \boldsymbol{v}_3 = 2a(1,1,1)$. Here, $\boldsymbol{v}_i$, $i = 1,2,3$ can be expressed as combinations of basis vectors of the high-symmetry phase, and the volume of the unit cell of their corresponding new phase is 2,4,8 times of the one of the old phase, respectively. For instance:

$$\boldsymbol{v}_1 = 2\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3, \qquad (2\boldsymbol{a}_1 \times \boldsymbol{a}_2)\cdot\boldsymbol{a}_3 = 2(\boldsymbol{a}_1 \times \boldsymbol{a}_2)\cdot\boldsymbol{a}_3 = 2\Omega_{\text{old}}, \tag{2.23}$$

where Ω_{old} is the volume of the high-symmetry phase unit cell.

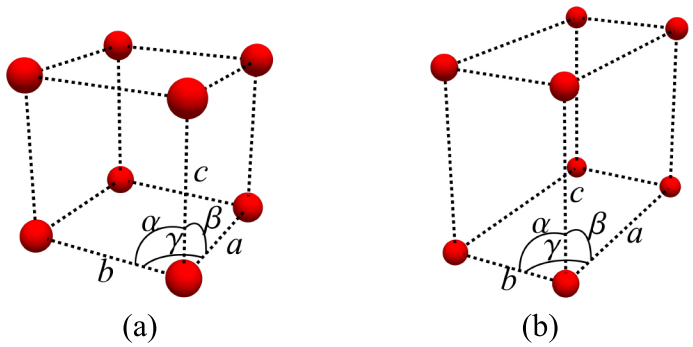


Figure 2.6 Sketch of (a) a simple cubic lattice in $Pm\bar{3}m$ SG, lattice parameters are: $\alpha = \beta = \gamma = \pi/2$, $a = b = c$; (b) an orthogonal lattice in $Cmcm$ SG, lattice parameters are:$\alpha = \beta = \gamma = \pi/2$, $a \neq b \neq c$.

Table 2.1 Crystal systems, possible point groups corresponding to each system and basis vectors of each system. P, I, F, C, R are Hermann–Mauguin notations standing for: simple cubic, body-centred, face-centred, base-centred and rhombohedral lattices [68].

Crystal System	Point Groups	Bravais Lattice	Basis Vectors
Triclinic	C_1, C_i	P	arbitrary
Monoclinic	C_s, C_2, C_h	P	$(0,-b,0), (a\sin\gamma, -a\cos\gamma, 0), (0,0,c)$
		C	$(0,-b,0), \frac{1}{2}(a\sin\gamma, -a\cos\gamma, -c), \frac{1}{2}(a\sin\gamma, -a\cos\gamma, c)$
Orthorhombic	D_{2h}, D_2, C_{2v}	P	$(0,-b,0), (a,0,0), (0,0,c)$
		C	$\frac{1}{2}(a,-b,0), \frac{1}{2}(a,b,0), (0,0,c)$
		I	$\frac{1}{2}(a,b,c), \frac{1}{2}(-a,-b,c), \frac{1}{2}(a,-b,-c)$
		F	$\frac{1}{2}(a,0,c), \frac{1}{2}(0,-b,c), \frac{1}{2}(a,-b,0)$
Tetragonal	$D_{4h}, C_{4v}, C_{4h}, D_4, C_4, D_{2d}, S_4$	I	$\frac{1}{2}(-a,a,c), \frac{1}{2}(a,-a,c), \frac{1}{2}(a,a,-c)$
		P	$(a,0,0), (0,a,0), (0,0,c)$
Trigonal	$D_{3d}, D_3, C_{3v}, C_3, C_{3i}$	R	$(0,-a,c), \frac{1}{2}(\sqrt{3}a,a,2c), \frac{1}{2}(-\sqrt{3}a,a,2c)$
Hexagonal	$D_6, D_{6h}, D_{3h}, C_6, C_{6h}, C_{6v}, C_{3h}$	P	$(0,-a,0), \frac{1}{2}(\sqrt{3}a,a,0), (0,0,c)$
Cubic	T, T_d, T_h, O, O_h	P	$(a,0,0), (0,a,0), (0,0,a)$
		I	$\frac{1}{2}(-a,a,a), \frac{1}{2}(a,-a,a), \frac{1}{2}(a,a,-a)$
		F	$\frac{1}{2}(a,a,0), \frac{1}{2}(a,0,a), \frac{1}{2}(0,a,a)$

Taking the X_5^- IRREP given in Table A3 as an example, we have the character calculated in Table 2.2. Plugging results in Table. 2.2 into Eq. 2.15 and Eq. 2.16, we can see:

1. Landau criterion:

$$\sum_g \left[\frac{1}{6}\chi^3(g)+\frac{1}{2}\chi(g^2)\chi(g)+\frac{1}{3}\chi(g^3)\right]=0 \tag{2.24}$$

2. Lifshitz criterion:

$$\sum_g \left\{\frac{1}{2}\chi_V(R)[\chi^2(g)-\chi(g^2)]\right\}=0 \tag{2.25}$$

Therefore, the X_5^- IRREP is an active representation. Since X_5^- is 2-dimensional (2D), the corresponding order parameter space V_η is also 2D, and the basis of V_η transforms in accordance with the basis of X_5^- under symmetry operations of $Pm\bar{3}m$ (similar to what is shown in Eq. 2.14). According to Table A3, the image group I_0 is isomorphic to C_{2v}. Choosing the vector $\boldsymbol{\eta}=(a,0)$ in V_η as an example, we see that $\boldsymbol{\eta}$ is invariant under C_{2v}. The isotropy group I_η is thus C_{2v}. The kernel of the map from the high symmetry space group to its isotropy group is $K=\{\{E|000\},\{M_y|000\}\}$. Therefore, we find the point group symmetry corresponding to the low-symmetry phase is:

$$\tilde{G}=C_{2v}\otimes\{E,M_y\}=D_{2h}. \tag{2.26}$$

Table 2.2 Table of characters used in calculating Landau criterion and Lifshitz criterion.

Operation	$\chi(g)$	$\chi(g^2)$	$\chi(g^3)$	$\chi_V(R)$
$\{E\|000\}$	2	2	2	3
$\{C_{2z}\|000\}$	0	2	0	−1
$\{C_{2y}\|000\}$	−2	2	−2	−1
$\{C_{2x}\|000\}$	0	2	0	−1
$\{C_{4y}\|000\}$	0	−2	0	1
$\{C_{2,101}\|000\}$	0	2	0	−1
$\{C_{4y}^3\|000\}$	0	−2	0	1
$\{C_{2,\bar{1}01}\|000\}$	0	2	0	−1
$\{I\|000\}$	−2	2	−2	−3
$\{M_z\|000\}$	0	2	0	1
$\{M_y\|000\}$	2	2	2	1
$\{M_x\|000\}$	0	2	0	1
$\{IC_{4y}\|000\}$	0	−2	0	−1
$\{M_{101}\|000\}$	0	2	0	1
$\{IC_{4y}^3\|000\}$	0	−2	0	−1
$\{M_{\bar{1}01}\|000\}$	0	2	0	1

Hence, possible Bravais lattices are P,C,I,F (Appendix A.1.3). Taking the new phase to be twice the volume of the old one as an example (see Eq. 2.23), then the basic translation that forms the translation group of the new phase should: 1) give the new phase a twice volume of the old one and 2) satisfy Eq.2.20. Given in this example we consider $\boldsymbol{k} = \frac{2\pi}{a}(\frac{1}{2},0,0)$, such that $\boldsymbol{v}_1 \cdot \boldsymbol{k} = 2\pi$, one set of basic translations can be:

$$\boldsymbol{t}_1 = 2\boldsymbol{a}_1 + \boldsymbol{a}_2 + \frac{1}{2}\boldsymbol{a}_3 = 2a(1,\tfrac{1}{2},\tfrac{1}{4}), \tag{2.27}$$

$$\boldsymbol{t}_2 = \boldsymbol{a}_2 + \boldsymbol{a}_3 = 2a(0,\tfrac{1}{2},\tfrac{1}{2}), \tag{2.28}$$

$$\boldsymbol{t}_3 = 2\boldsymbol{a}_1 + \frac{1}{2}\boldsymbol{a}_3 = 2a(1,0,\tfrac{1}{4}). \tag{2.29}$$

It can be proven that $\boldsymbol{k}\cdot\boldsymbol{t}_i = 2\pi n$, $n \in \mathbb{Z}$, $i = 1,2,3$. To correspond these translations to those listed in Table 2.1, we need to find a combination of $\boldsymbol{t}_1,\boldsymbol{t}_2,\boldsymbol{t}_3$, such that:

$$\boldsymbol{t}_1' = \boldsymbol{t}_3 - \frac{1}{2}\boldsymbol{t}_2 = 2a(1,\overline{\tfrac{1}{4}},0), \tag{2.30}$$

$$\boldsymbol{t}_2' = \boldsymbol{t}_1 - \frac{1}{2}\boldsymbol{t}_2 = 2a(1,\tfrac{1}{4},0), \tag{2.31}$$

$$\boldsymbol{t}_3' = \boldsymbol{t}_3 - \boldsymbol{t}_1 + \boldsymbol{t}_2 = 2a(0,0,\tfrac{1}{2}). \tag{2.32}$$

According to Table 2.1, this describes a C lattice [as shown in Fig. 2.6(b)]. Given $\boldsymbol{\eta} = (a,0)$ and the generators of D_{2h} are $\{IC_{2x}, IC_{2y}, IC_{2z}\}$, we expect the skew axis to be along the z- or y- directions to ensure a change in the unit cell volume, which means all possible elements are:

$$\{E|000\}, \quad \{IC_{2x}|000\}, \quad \{IC_{2y}|00\tfrac{1}{2}\}, \quad \{IC_{2z}|00\tfrac{1}{2}\}, \tag{2.33}$$

$$\{I|000\}, \quad \{C_{2x}|000\}, \quad \{C_{2y}|00\tfrac{1}{2}\}, \quad \{C_{2z}|00\tfrac{1}{2}\}, \tag{2.34}$$

which yields a space group *Cmcm* [Fig. 2.6(b)]. The invariant polynomial here can be derived using C_{2v} (I_0). According to Fig. 2.7, the normal series is:

$$\{E\} \trianglelefteq C_2 \trianglelefteq C_{2v} \tag{2.35}$$

Then the derivation of the integral rotational basis of invariants is given as follows.

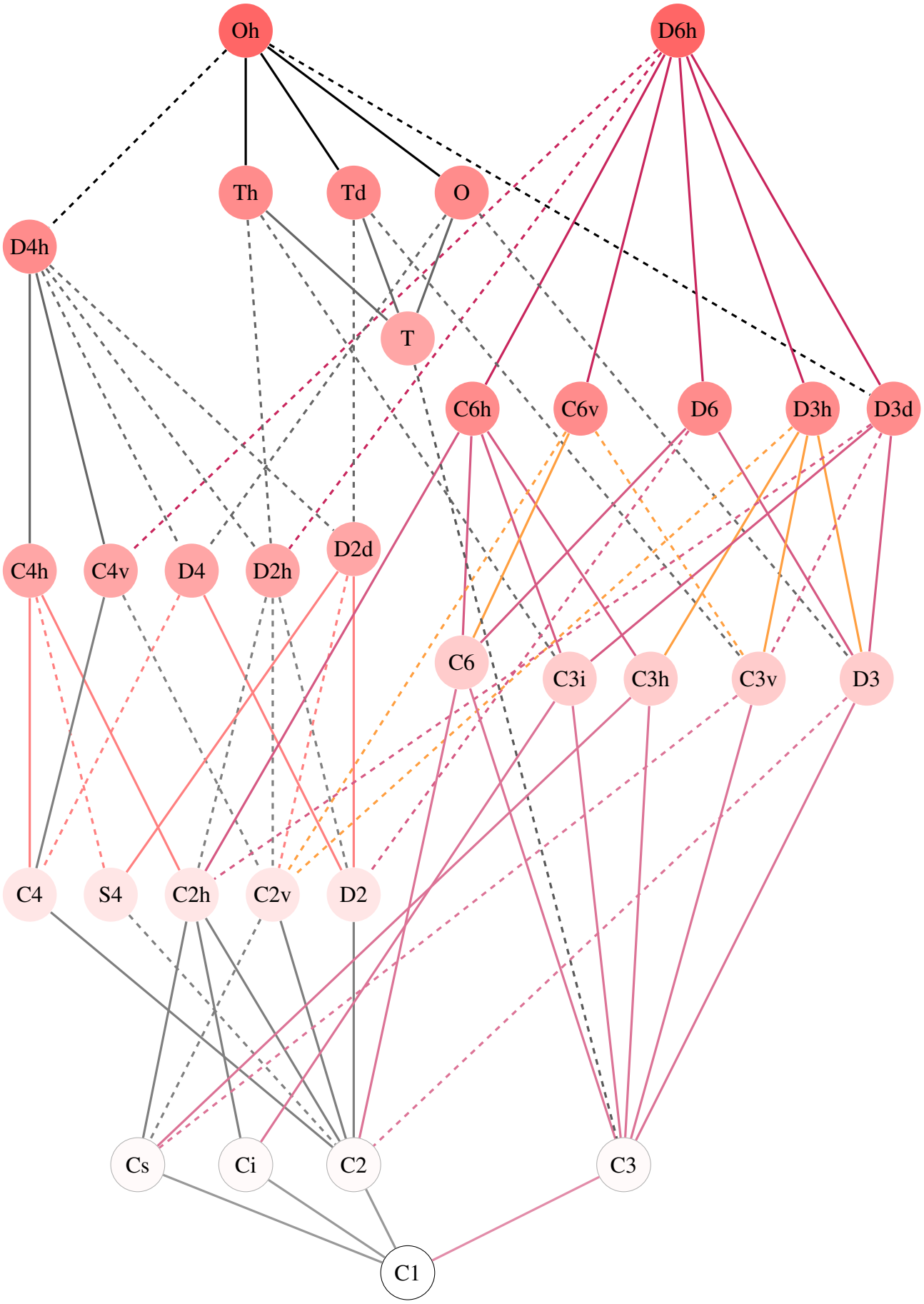


Figure 2.7 Hierarchy of 32 point groups [86]. Groups related by G-sub G relations are connected by solid lines.

Here we take the IRREP X_5^- as an example. Because X_5^- is a 2D IRREP, the order parameter space is of dimension 2. Let $\boldsymbol{\eta} = (\eta_1, \eta_2)$. Now we can begin to construct polynomials that are invariant under the group C_2, then the group C_{2v}. The method is performing symmetry operations of, for instance, C_2 onto the chosen vector $\boldsymbol{\eta}$, then finding combinations of elements in η that are invariant under C_2:

1. $C_2 = \{E, C_{2z}\}$:

$$C_{2z}\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix} = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix} = \begin{pmatrix}\eta_2\\ \eta_1\end{pmatrix}. \tag{2.36}$$

   Hence, the polynomials that are invariant under C_{2z} are $I_{2,1} = \eta_1 + \eta_2, I_{2,2} = \eta_1\eta_2$.

2. $C_{2v} = C_2 + IC_{2y}C_2$:

$$IC_{2y}\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix} = \begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix}. \tag{2.37}$$

   The polynomials now are: $I_{1,1} = \eta_1 + \eta_2, I_{1,2} = \eta_1\eta_2$. Hence, the free energy can be expressed as:

$$\varphi = \varphi_0 + A_1(\eta_1 + \eta_2)^2 + A_2\eta_1\eta_2 + B_1(\eta_1 + \eta_2)^4 + B_2\eta_1^2\eta_2^2 + \cdots. \tag{2.38}$$

It is noteworthy here that $\boldsymbol{\eta}$ is a vector in the order parameter space, but not a basis of the IRREP X_5^-. Basis vectors of X_5^- will pre-assign a direction to $\boldsymbol{\eta}$, hence cannot be applied to studying phases corresponding to $\boldsymbol{\eta}$ along other directions in the order parameter space. More details can be found in Ref. [80].

2.2.2 Reconstructive Phase Transformations

Since no G-sub G relation can be found in reconstructive phase transformations, we cannot write $\delta\rho$ as a linear combination of basis functions of an active representation. Now, the problem becomes searching for a property, such that this property: 1) is a function of some quantities directly related to the phase transformation and 2) contains structure information of the crystal lattice. The atomic displacement is a possible candidate.

Recalling the structure factor applied in the analysis of diffraction patterns that is a sum

of plane waves scattered by atoms in the unit cell[4]; we see that the *lattice function* [37]:

$$\psi(\boldsymbol{r}) = \frac{1}{N}\sum_{i=1}^{N} A_i \exp[\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{R_i})] + \text{c.c}, \tag{2.39}$$

can satisfy these requirements. Here, the sum runs over all atoms in one unit cell, A_i are normalised amplitudes of the plane wave, $\boldsymbol{R}_j$ are positions of atoms in the same unit cell, and $\boldsymbol{k}$ is the wave vector determined by the atomic displacements. For instance, considering a two-atom (of the same kind) unit cell and let the atomic displacement be along the z-direction by $\pm\zeta$. Let atoms locating at $z_1 = c/4$ and $z_2 = -c/4$, respectively, then $\boldsymbol{k}$ here will be chosen as $\boldsymbol{k} = \boldsymbol{b}_3$, where $\boldsymbol{b}_3$ is a reciprocal lattice vector corresponding to the lattice vector along the z-direction. Then Eq. 2.39 is re-written as:

$$\begin{aligned}
\psi(\boldsymbol{r}) &= \frac{A}{2}\Big\{\exp[\mathrm{i}\boldsymbol{b}_3\cdot(z-z_1-\zeta)\boldsymbol{e}_z] + \exp[\mathrm{i}\boldsymbol{b}_3\cdot(z-z_2+\zeta)\boldsymbol{e}_z] + \text{c.c}\Big\},\\
&= A\Big\{\cos[\boldsymbol{b}_3\cdot(z-z_1-\zeta)\boldsymbol{e}_z] + \cos[\boldsymbol{b}_3\cdot(z-z_2+\zeta)\boldsymbol{e}_z]\Big\},\\
&= 2A\cos(\boldsymbol{b}_3\cdot z\boldsymbol{e}_z)\cos[\boldsymbol{b}_3\cdot(\frac{c}{4}+\zeta)\boldsymbol{e}_z],\\
&= 2A\cos(\boldsymbol{b}_3\cdot z\boldsymbol{e}_z)\cos(\frac{\pi}{2}+\boldsymbol{b}_3\cdot\zeta\boldsymbol{e}_z),\\
&= -2A\cos(\boldsymbol{b}_3\cdot z\boldsymbol{e}_z)\sin(\boldsymbol{b}_3\cdot\zeta\boldsymbol{e}_z)
\end{aligned} \tag{2.40}$$

The variation in $\psi(\boldsymbol{r})$ is:

$$\delta\psi(\boldsymbol{r}) = \psi|_{\zeta=0} - \psi|_{\zeta} = 2A\cos(\boldsymbol{b}_3\cdot z\boldsymbol{e}_z)\sin(\boldsymbol{b}_3\cdot\zeta\boldsymbol{e}_z). \tag{2.41}$$

From $\delta\psi(\boldsymbol{r})$, one can define the order parameter $\eta = \eta_0\sin(\boldsymbol{b}_3\cdot\zeta\boldsymbol{e}_z)$ [37]. By tuning $|\zeta|$, we obtain a curve of η that gives different structures occurring during the phase transformation. Consider the phase transformation from an HCP lattice ($P6_3/mmc$) [Fig. 2.8(b)] to a hexagonal lattice ($P6/mmm$) [Fig. 2.8(a,c)]. The original phase has atoms located at:

$$z = \pm 1/4. \tag{2.42}$$

Then, when $\zeta = 0, \pm c/2$, the phase will be in an HCP lattice, and when $\zeta = \pm c/4$, the phase will be in a hexagonal lattice. In addition to atomic displacements, parameters such as variation in the lattice wave amplitude A_i in Eq. 2.39 can also be chosen as order parameters, depending on the phase transformation studied.

[4]Details are in Sec. 3.1.

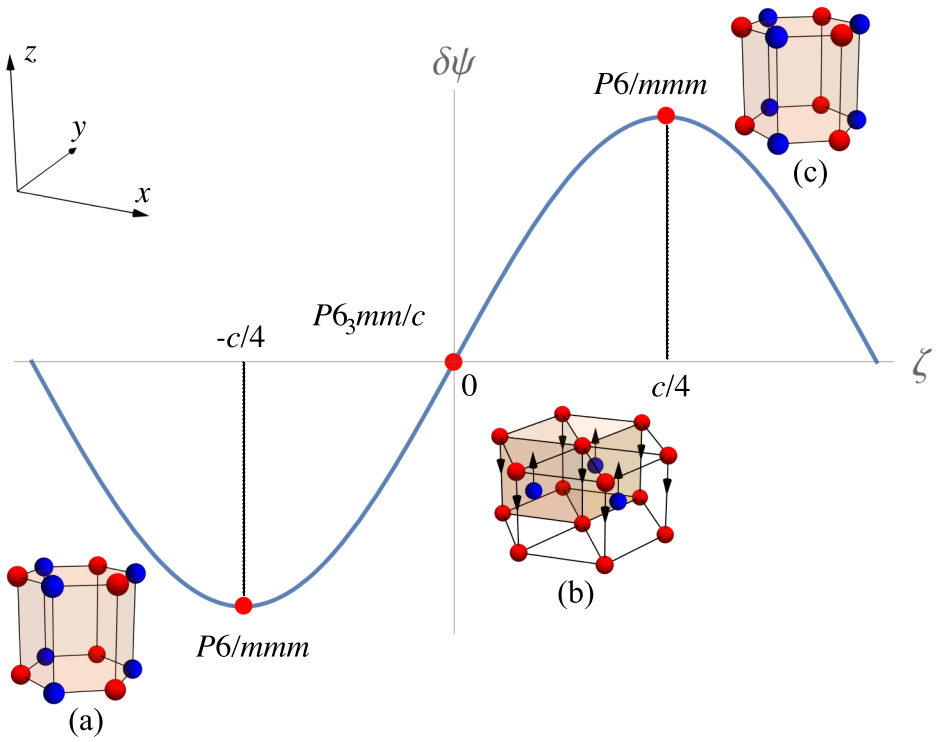


Figure 2.8 Plot of $\delta\psi$ against the displacement ζ. Atoms in (b) the original HCP lattice will displace along directions indicated by the arrows. When $\zeta = c/4$ (resp. $\zeta = -c/4$), the (a) (resp. (b)) hexagonal lattice will occur.

2.2.3 Phonon Softening

Phonon softening is a phenomenon that phonon frequencies decrease substantially around the critical temperature observed principally in structural phase transformations [38]. Recently, it is reported that the phonon softening is also observed in topological phase transformations [87]. Experimentally, it was first observed by Raman and Nedungadi when they studied quartz [88]. Their work inspired theorists studying phase transformations, and contributed to several phase transformations theories, including the one based on Landau's theory proposed by Ginzburg and Levanyuk [61–63], lattice dynamics proposed by Cochran [39], and microscopic anharmonic oscillators proposed by Cowley [64]. Phonon softening can be used to predict the deformation direction of atoms that allows this system to enter a new phase with lower energy. The key to this model is solving the eigenproblem for the *dynamic matrix* $\mathsf{D}(q)$ of this system:

$$\omega_q^2 \boldsymbol{e}(q) = \mathsf{D}(q)\boldsymbol{e}(q). \tag{2.43}$$

where q are phonon modes, ω_q and $\boldsymbol{e}(q)$ are the vibrational frequencies and phonon polarisation vectors corresponding to the mode q. By solving the eigenproblem for the dynamic matrix, we have ω_q being eigenvalues of $\mathsf{D}(q)$ and $\boldsymbol{e}(q)$ being eigenvectors of $\mathsf{D}(q)$. Moreover,

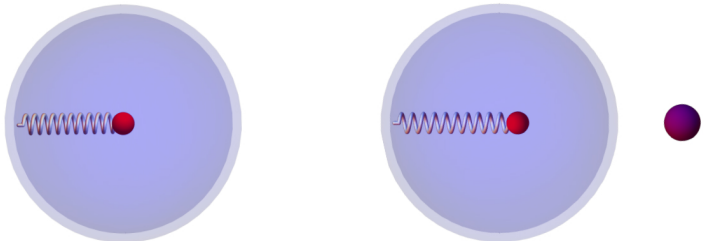
the potential energy can be expressed as:

$$U = \frac{1}{2}\sum_{q} \omega_q^2 |A(q)|^2, \tag{2.44}$$

where $A(q)$ are the vibrational amplitudes corresponding to the mode q. Hence, when there is a vanishing of ω_q, it means that there is a direction e_q such that if the atoms go along this direction, the the system may enter a new lower-energy phase. Moreover, imaginary frequencies with $\omega_q^2 < 0$ are used as an indicator of the occurrence of a new lower-energy phase. More details can be found in Ref. [89, 90]. Cochran [39] based his theory on a harmonic oscillator model and the core-shell model [91]. This core-shell model treats an ion (charge Ze) as a combination of a positive core (charge $Xe > 0$) and a massless negative shell (charge $Ye < 0$), such that:

$$(X+Y) = Z, \qquad X, Y, Z \in \mathbb{Z}. \tag{2.45}$$

The core is coupled to the shell (Fig. 2.9) and behaves as a harmonic oscillator. He considered a cubic diatomic crystal, represented the Coulomb force by a sinusodial polarisation and included the short-range non-Coulomb force such as the coupling between the core and the shell [39]. By solving the dynamic functions of the shell, the core and the neighbour that is a positive ion of one negative iron, he realised that the transverse phonon mode will decrease to zero when the critical temperature is approached [38,39]. Cowley developed his microscopic theory based on an anharmonic oscillator. He introduced a damping term and a sinusoidal driving force and studied the complex susceptibility χ of this system. According to Cowley, there should be a dissipation of χ ($\chi \to \infty$) when the critical temperature is arrived and χ is a function of the frequency ω. Therefore, by looking at the expression of χ in terms of ω, we



(a) (b)

Figure 2.9 Sketch of the core-shell model. (a) Coupling of a core marked by a red sphere to the shell marked by a blue shell. (b) Polarised core and shell along opposite directions under the effect of an ion marked by a purple sphere.

can see at the critical point:

$$\omega^2 = -\omega_0^2 + CT, \qquad \omega_0^2 = \frac{k}{m} > 0, \tag{2.46}$$

where $C, T \in \mathbb{R} > 0$ are physical quantities. By substituting $-\omega_0^2$ by $-CT_c$, we see that the frequency can be expressed as:

$$\omega^2 = C(T - T_c), \tag{2.47}$$

which means ω will vanish when $T = T_c$. More details about phonon softening models mentioned here can be found in Ref. [38, 39, 58, 65, 92, 93]. In real applications, phonon modes are calculated with the help of the density functional theory. One can plot out the phonon frequency along high-symmetry lines in the Brillouin zone, then study the phonon softening by searching for the imaginary frequency, as shown in Fig. 2.10. To be specific, in Fig. 2.10(a), SnS is stable in the *Pnma* space group, and its frequency is always real. However, the *Cmcm* SnS is unstable. Hence, it has imaginary frequencies, which means the system energy can still decrease as atoms are displaced, and the phonon softening is observed at Γ and Y points [94, 95]. Calculations given in Ref. [94] are done at zero temperature. For calculating phonon modes at a finite temperature, effects such as the thermal expansion and the anharmonic effect must be considered. There are several methods proposed to calculate phonon frequencies at a finite temperature, including those based on molecular dynamics that consider atomic vibrations [96–98] and those based on density functional perturbation theory that consider anharmonic phonon-phonon interactions [99]. Finding the equilibrium volume at a given non-zero temperature first and then calculating the phonon frequency at this temperature based on the quasi-harmonic approximation is also a method proposed [100]. More details related to the finite temperature phonon frequency calculation can be found in Ref. [96–99, 101, 102].

2.2.4 Deformation Models

Begin with the $\gamma - \alpha$ iron martensitic phase transformation. Bain [31] proposed a deformation to explain the mechanism behind this FCC$-$BCC lattice deformation, *i.e.*, an elongation along the x- and y-axis by 12% and a compression along the z-axis by 21% [Fig. 2.11(a,b,c)]. The Nishiyama–Wassermann (NW) deformation [103] differs from the Bain deformation in interpreting the transformation as being realised by a $[112]_{\mathrm{FCC}}$ shear rotation [Fig. 2.11(d,e,f)], such that the orthogonal lattice [Fig. 2.11(d)] transforms to a monoclinic lattice [Fig. 2.11(e)] then to an orthogonal lattice [Fig. 2.11(f)]. Similar to the NW deformation [104], the Kurdjumow–

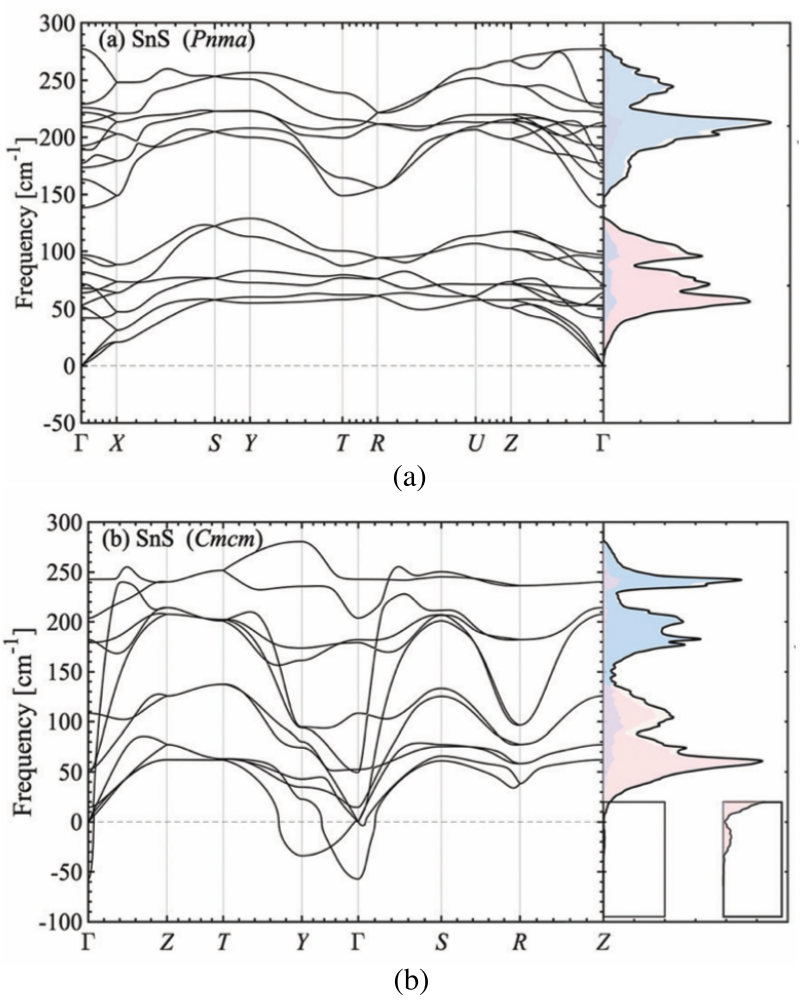


Figure 2.10 Phonon frequency of SnS in (a) *Pnma* space group and (b) *Cmcm* space group. Reproduced with permission [94]. Copyright 2021 Royal Society of Chemistry. Negative frequencies are the imaginary frequencies searched for to indicate the appearance of a new phase. The modulus are chosen for the convenience of plotting. In (b), the vibration modes along Γ and *Y* directions give rise to new phases.

Sachs (KS) deformation realised by a $[112](111)$ shear in the FCC lattice was also proposed. In the Bogers–Burgers (BB) deformation [105], which shares similarities with the Cohen–Olson (CO) [34] mechanism. However, we have the FCC lattice transforming to a BCC lattice by a gliding of atoms in $(111)_{\mathrm{FCC}}$ [Fig. 2.11(g,h,i)].

Burgers [106] also proposed a two-step model to explain the BCC−HCP deformation in zirconium (Fig. 2.12):

1. A $[\bar{1}10]_{\mathrm{BCC}}(110)_{\mathrm{BCC}}$ shear to obtain the HCP stacking sequence;
2. A $[1\bar{1}1]_{\mathrm{BCC}}(\bar{1}12)_{\mathrm{BCC}}$ shear to get the HCP lattice.

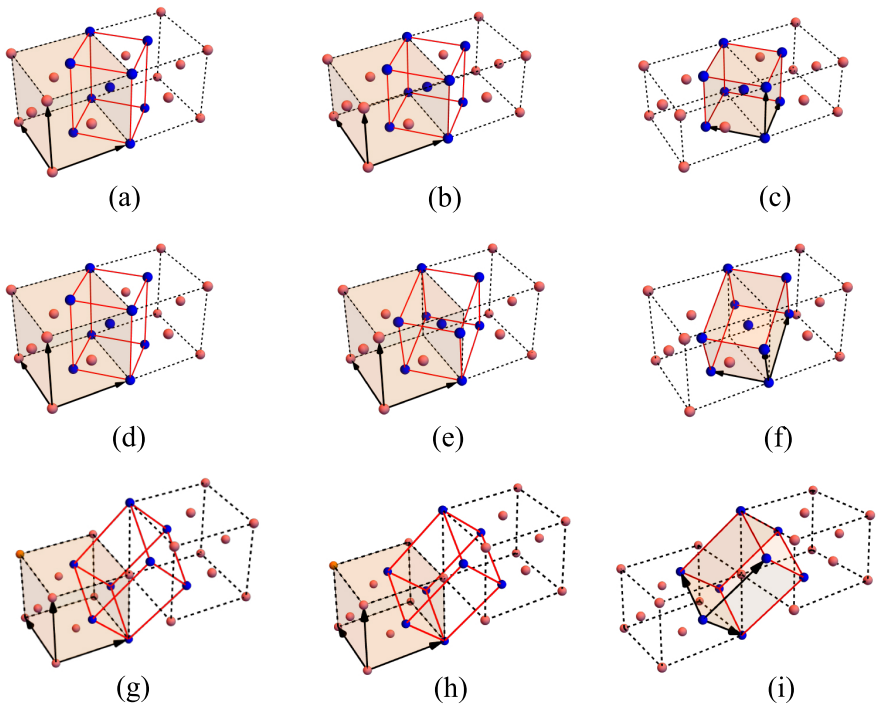


Figure 2.11 Sketch of (a,b,c) the Bain deformation, (d,e,f) the Nishiyama–Wassermann deformation and (g,h,i) the Bogers–Burgers deformation. From (a) an FCC lattice to (c) a BCC lattice. From (d) to (f), we have (d) an FCC lattice that goes to (e) a monoclinic lattice and then to (f) an orthogonal lattice. From (g) to (i), we have (g) an FCC lattice, goes to (h) a monoclinic lattice and then to (i) an orthogonal lattice. Unit cells chosen here are shaded orange, with unit cell basis vectors labelled by black arrows.

Another model describing a BCC−HCP phase transformation proposed by Ojha and Sehitoglu [107] is to study the martensitic phase transformation in Ti-Ta and Ti-Nb alloys. This two-step model involves a $[110](110)$ shear in the BCC lattice by $\sqrt{2}a/6$ and a dilation (resp. compression) along the $\langle 110\rangle_{\mathrm{BCC}}$ (resp. $\langle 001\rangle_{\mathrm{BCC}}$) direction, where a is the lattice parameter of the BCC lattice. The orientation relationships of models mentioned above are given in Table 2.3. Moreover, due to the lack of experimental observations of the intermediate states of transformation paths, it is still being determined whether the models mentioned in this section are actual models describing a real path or not.

Table 2.3 List of orientation relationships

Transformation	Model	Orientation Relationship
FCC−BCC	Bain [31]	$(001)_{\mathrm{FCC}}\|\|(001)_{\mathrm{BCC}}$, $[110]\mathrm{FCC}\|\|[100]_{\mathrm{BCC}}$
	KS [104]	$(111)_{\mathrm{FCC}}\|\|(110)_{\mathrm{BCC}}$, $[\bar{1}10]_{\mathrm{FCC}}\|\|[111]_{\mathrm{BCC}}$
	NW [108]	$(111)_{\mathrm{FCC}}\|\|(110)_{\mathrm{BCC}}$, $[0\bar{1}1]_{\mathrm{FCC}}\|\|[001]_{\mathrm{BCC}}, [\bar{1}\bar{1}\bar{1}]_{\mathrm{FCC}}\|\|[1\bar{1}0]_{\mathrm{BCC}}$
	BB [105]	$(11\bar{1})_{\mathrm{FCC}}\|\|(110)_{\mathrm{BCC}}$, $[\bar{1}10]_{\mathrm{FCC}}\|\|[1\bar{1}1]_{\mathrm{BCC}}$
BCC−HCP	Burgers [106]	$(011)_{\mathrm{BCC}}\|\|(0001)_{\mathrm{HCP}}$, $[\bar{1}11]_{\mathrm{BCC}}\|\|[11\bar{2}0]_{\mathrm{HCP}}$
	Ojiha [107]	$(110)_{\mathrm{BCC}}\|\|(0001)_{\mathrm{HCP}}$, $[\bar{1}11]_{\mathrm{BCC}}\|\|[\bar{2}110]_{\mathrm{HCP}}$

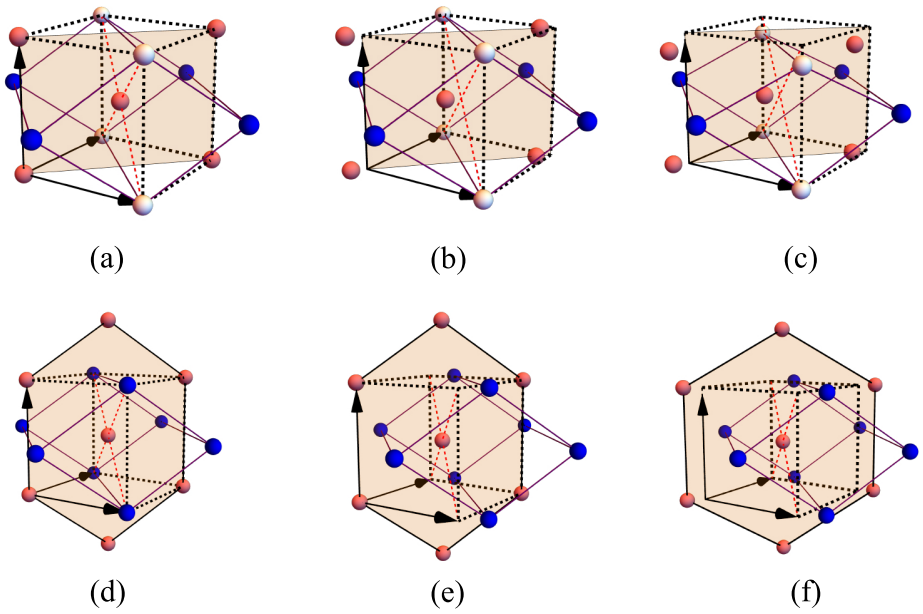


Figure 2.12 Sketch of (a,b,c) the Burgers mechanism and (d,e,f) BCC−HCP deformation model of Ti-Ta/Ti-Tb alloys. (a,d) show the original BCC lattice, (b) the lattice after the first step shear, and (c) the HCP lattice. (e) deformed BCC lattice by a $[110](110)$ shear. (f) the HCP lattice obtained after a compression along the z-axis and a dilation along the other two directions. Blue spheres are atoms at the body centre; pink spheres are atoms aligned on the $[110]_{\mathrm{BCC}}$.

2.2.5 Continuum Theory Based on Energy Minimisation

Bhattacharya [5] proposed a continuum theory for martensitic phase transformations based on the minimisation of free energy, which depends on the deformation gradient and the temperature. The *deformation gradient* $\nabla \boldsymbol{y}$ is defined as [5]:

$$\nabla \boldsymbol{y}(\boldsymbol{x}) = \frac{\partial y_i}{\partial x_j}, \qquad i,j = 1,2,3, \tag{2.48}$$

where $\boldsymbol{x}$ is the original position of a point in the crystal and $\boldsymbol{y}$ is a deformation that maps a point at $\boldsymbol{x}$ to a new position $\boldsymbol{y}(\boldsymbol{x})$ (Fig. 2.13). Since the structural deformation considered here is continuous, the *Cauchy-Born hypothesis*, which is valid when large slip or dislocations compatible with plastic deformations are ignored, should be considered here [5]:

$$\hat{\boldsymbol{e}}'_i(\boldsymbol{x}) = \mathsf{F}(\boldsymbol{x})\hat{\boldsymbol{e}}_i(\boldsymbol{x}), \qquad \mathsf{F}(\boldsymbol{x}) = \nabla \boldsymbol{y}(\boldsymbol{x}) \tag{2.49}$$

where $\{\hat{\boldsymbol{e}}_i(\boldsymbol{x})\}, \{\hat{\boldsymbol{e}}'_i(\boldsymbol{x})\}$ are lattice vectors of the Bravais lattices at the point $\boldsymbol{x}$ before and after the deformation, respectively. Under the Cauchy-Born hypothesis, after a deformation F, the free energy density $\phi(\hat{\boldsymbol{e}}'_i, T)$ stored by the lattice that depends on the lattice vector and the temperature T can be denoted by $\phi(\mathsf{F}, T) = \phi(\mathsf{F}\hat{\boldsymbol{e}}_i, T)$, and is required to satisfy [5]:

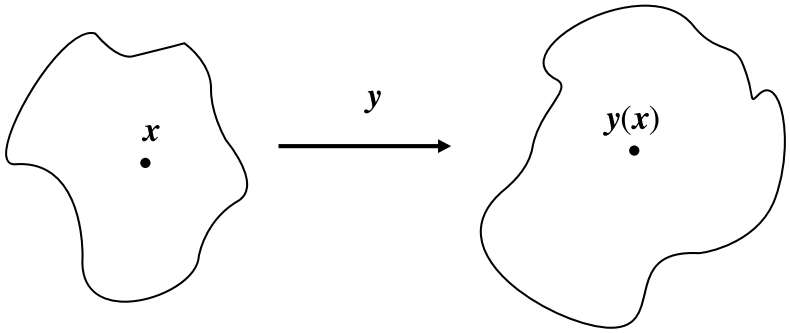


Figure 2.13 Sketch of deformation $\boldsymbol{y}$ that brings a configuration on left to a new configuration to right.

1. Independent of choice of the frame: $\phi(\mathsf{F},T)$ is invariant under all rigid-body rotations.

2. Invariance under point group symmetry: the choice of lattice vectors will not affect $\phi(\mathsf{F},T)$, such that under a symmetry operation $R \in \tilde{G}_0$, where $\tilde{G}_0$ is the point group that maps the original lattice $\{\hat{\boldsymbol{e}}_i(\boldsymbol{x})\}$ back to itself, $\phi(\mathsf{F}R,T) = \phi(\mathsf{F},T)$. This means if we operate a rotation R onto the original lattice $\{\hat{\boldsymbol{e}}_i(\boldsymbol{x})\}$ first then take a deformation gradient F, the free energy density is the same as the one obtained by taking F directly on $\{\hat{\boldsymbol{e}}_i(\boldsymbol{x})\}$.

Now, the total free energy can be expressed as the integral of $\phi(\mathsf{F},T)$ over the whole crystal, such that [5]:

$$\varphi = \int_{\Omega} \phi(\mathsf{F},T)\,\mathrm{d}V, \tag{2.50}$$

where Ω is the total volume of the crystal. The crystal structure is now determined by the minimisation of ϕ [5]. To be specific, let I be the identity matrix, and F_m be the gradient corresponding to a martensitic phase transformation, and T_c be the critical temperature:

$$\begin{aligned} &\phi(\mathsf{I},T) < \phi(\mathsf{F}_m,T), \qquad T > T_c, \\ &\phi(\mathsf{I},T_c) = \phi(\mathsf{F}_m,T_c), \\ &\phi(\mathsf{I},T) > \phi(\mathsf{F}_m,T), \qquad T < T_c. \end{aligned} \tag{2.51}$$

Assume $P_m \subset P_a$, P_m and P_a are the isogonal groups of the martensite and austenite phases, respectively. Because F_m is the deformation gradient describing the martensitic phase transformation, F_m should be invariant under all operations in P_m:

$$R\mathsf{F}_m = \mathsf{F}_m R = \mathsf{F}_m, \qquad R \in P_m. \tag{2.52}$$

For all operations $R \in P_a, R \notin P_m$, since R is also a rigid-body rotation, according to the invariance requirements of $\phi(\mathsf{F}, T)$, we have:

$$\phi(\mathsf{F}_m R, T) = \phi(R\mathsf{F}_m, T) = \phi(\mathsf{F}_m, T), \tag{2.53}$$

which means:

$$\phi(R^{-1}\mathsf{F}_m R, T) = \phi(\mathsf{F}_m, T), \tag{2.54}$$

such that:

$$R^{-1}\mathsf{F_m}R = \mathsf{F}_m^i \neq \mathsf{F}_m, \qquad R \in P_a, R \notin P_m, i = 1, 2, \cdots, |P_a|/|P_m|. \tag{2.55}$$

Also, according to Eq. 2.54, we have:

$$\phi(R^{-1}\mathsf{F}_m R, T) = \phi(\mathsf{F}_m^i, T) = \phi(\mathsf{F}_m, T), \qquad \forall R \in P_a. \tag{2.56}$$

Hence, all variants related by symmetries are of the same energy. Given rotations do not change the state of a crystal, *i.e.*, will not turn a martensite into an austenite, all matrices correspond to the same crystal state are said to form an **energy well** of this state [5]. To be specific, the set:

$$A := \{\mathsf{M} = \mathsf{RI} \,|\mathsf{R} \in \mathrm{SO}(3)\}, \tag{2.57}$$

is said to form an austenite energy well and the set:

$$M_i := \{\mathsf{M} = \mathsf{RF}_m^i \,|\mathsf{R} \in \mathrm{SO}(3)\}, \tag{2.58}$$

is said to form an energy well of the ith variant of the martensite (Fig. 2.14). According to Eq. 2.53, it can be seen that all matrices in the same well are of the same energy. Next, we will gain more insight into the shear, *i.e.*, the shear plane and the shear direction, that leads to a twinning by looking at the invariant plane.

As for the twinning, it can be understood from the perspective of *kinematic compatibility condition* [5]. Begin with the deformation gradient F. The infinitesimal volume element before ($\mathrm{d}V_0$) and after ($\mathrm{d}V$) the phase transformation expressed in Cartesian coordinates are:

$$\mathrm{d}V_0 = (\mathrm{d}\boldsymbol{x} \times \mathrm{d}\boldsymbol{y}) \cdot \mathrm{d}\boldsymbol{z} = \varepsilon^{ijk}\mathrm{d}x_i\mathrm{d}y_j\mathrm{d}z_k, \tag{2.59}$$

$$\mathrm{d}V = (\mathrm{d}\boldsymbol{x}' \times \mathrm{d}\boldsymbol{y}') \cdot \mathrm{d}\boldsymbol{z}' = \varepsilon^{ijk}\mathrm{d}x_i'\mathrm{d}y_j'\mathrm{d}z_k', \tag{2.60}$$

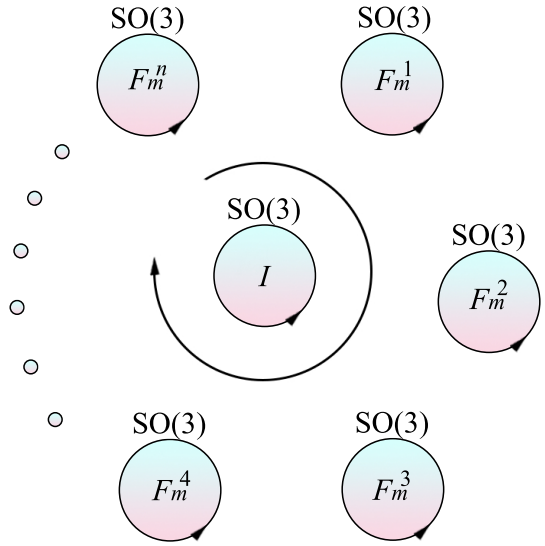


Figure 2.14 Energy wells of $p = |P_a|/|P_m|$ variants and the austenite phase.

with ε_{ijk} being the Levi-Civita symbol. And:

$$\begin{pmatrix} \mathrm{d}\alpha_1' \\ \mathrm{d}\alpha_2' \\ \mathrm{d}\alpha_3' \end{pmatrix} = \begin{pmatrix} \mathsf{F}_{11} & \mathsf{F}_{12} & \mathsf{F}_{13} \\ \mathsf{F}_{21} & \mathsf{F}_{22} & \mathsf{F}_{23} \\ \mathsf{F}_{31} & \mathsf{F}_{32} & \mathsf{F}_{33} \end{pmatrix} \begin{pmatrix} \mathrm{d}\alpha_1 \\ \mathrm{d}\alpha_2 \\ \mathrm{d}\alpha_3 \end{pmatrix}, \qquad \alpha = x, y, z. \tag{2.61}$$

Hence, it can be seen that:

$$\begin{aligned} \mathrm{d}V &= \varepsilon^{ijk}\mathsf{F}_{ih}\mathsf{F}_{jl}\mathsf{F}_{km}\mathrm{d}x_h\mathrm{d}y_l\mathrm{d}z_m, \\ &= (\det\mathsf{F})\mathrm{d}V_0, \end{aligned} \tag{2.62}$$

where $\varepsilon^{ijk}\mathsf{F}_{ih}\mathsf{F}_{jl}\mathsf{F}_{km} = \varepsilon_{hlm}\det\mathsf{F}$. Similarly, the surface area element before ($\mathrm{d}a_0$) and after ($\mathrm{d}a$) the deformation can be expressed using infinitesimal elements $\mathrm{d}\boldsymbol{b}$ and $\mathrm{d}\boldsymbol{c}$ in the reference configuration:

$$(\mathrm{d}a_0)n_i^0 = \varepsilon_i{}^{jk}\mathrm{d}b_j\,\mathrm{d}c_k, \tag{2.63}$$

$$(\mathrm{d}a)n_i = \varepsilon_i{}^{jk}\mathrm{d}b_j'\mathrm{d}c_k' = \varepsilon_i{}^{jk}\mathsf{F}_{jl}\mathsf{F}_{km}\mathrm{d}b_l\,\mathrm{d}c_m, \tag{2.64}$$

where $\boldsymbol{n}^0$ and $\boldsymbol{n}$ are unit normal vectors before and after the deformation. Using equation as following:

$$(\mathsf{F}^{-1}\det\mathsf{F})_{ih} = \mathsf{F}_{ih}^{-1}\det\mathsf{F} \tag{2.65}$$

we can see:

$$(\mathrm{d}a)\boldsymbol{n} = [(\det\mathsf{F})\mathsf{F}^{-1}\mathrm{d}a_0]\boldsymbol{n}^0. \tag{2.66}$$

Considering the deformation gradients $\mathsf{F}_1, \mathsf{F}_2$ on either side of the twinning boundary with $\boldsymbol{n}$ being the unit normal vector at the boundary, according to Eq. 2.66, we have:

$$(\text{cof}\,\mathsf{F}_1)\boldsymbol{n} = (\text{cof}\,\mathsf{F}_2)\boldsymbol{n}, \tag{2.67}$$

where $\text{cof}\,\mathsf{F} = (\det\mathsf{F})\mathsf{F}^{-1}$. Labelling $\mathsf{F}_{12} = \mathsf{F}_1 - \mathsf{F}_2$, then for all vectors $\boldsymbol{r}$ in the space of interest (3D space bounded by the material), we have:

$$\mathsf{F}_{12}\boldsymbol{r}_{||} = 0, \qquad \boldsymbol{r}_{||} = \boldsymbol{r} - (\boldsymbol{n}\cdot\boldsymbol{r})\boldsymbol{n}. \tag{2.68}$$

Since $(\boldsymbol{a}\otimes\boldsymbol{b})\boldsymbol{c} = (\boldsymbol{b}\cdot\boldsymbol{c})\boldsymbol{a}$, we now will arrive at:

$$\mathsf{F}_{12}[\boldsymbol{r} - (\boldsymbol{n}\otimes\boldsymbol{n})\boldsymbol{r}] = 0 \Rightarrow \mathsf{F}_{12} = \mathsf{F}_{12}(\boldsymbol{n}\otimes\boldsymbol{n}), \tag{2.69}$$

which is the *kinematic compatibility condition* [5]. We now set:

$$\mathsf{F}_{12}\boldsymbol{n} = \boldsymbol{a}, \tag{2.70}$$

and let the original lattice vectors to be $\{\boldsymbol{e}_i\}$, $i = 1,2,3$. For neighbouring variants caused by gradients $\mathsf{F}_1, \mathsf{F}_2$ on either side of the boundary, respectively, according to the Cauchy-Born hypothesis, we have:

$$\mathsf{Q}_2\mathsf{F}_2 = \mathsf{Q}_1\mathsf{F}_1, \qquad \mathsf{Q}_1^{-1}\mathsf{Q}_2\mathsf{F}_2 - \mathsf{F}_2 = \boldsymbol{a}\otimes\boldsymbol{n}, \tag{2.71}$$

where $\mathsf{Q}_1, \mathsf{Q}_2 \in$O(3), which is a set of all orthogonal 3×3 matrices with their determinants equal to ± 1 defined on the field of real numbers, describing rigid-body rotations, and:

$$\mathsf{Q}_1^{-1}\mathsf{Q}_2\mathsf{F}_2 = [\mathsf{I} + \boldsymbol{a}\otimes(\mathsf{F}_2^{-1}\boldsymbol{n})]\mathsf{F}_2, \tag{2.72}$$

Since $\det(\mathsf{Q}_1^{-1}\mathsf{Q}_2\mathsf{F}_2) = \det(\mathsf{F}_2)$, we have $\det\mathsf{F}_2 = \det\{[\mathsf{I} + \boldsymbol{a}\otimes(\mathsf{F}_2^{-1}\boldsymbol{n})]\mathsf{F}_2\}$, which means we have:

$$\det\mathsf{F}_2 = [1 + \boldsymbol{a}\cdot(\mathsf{F}_2^{-1}\boldsymbol{n})]\det(\mathsf{F}_2). \tag{2.73}$$

Hence we can see that:

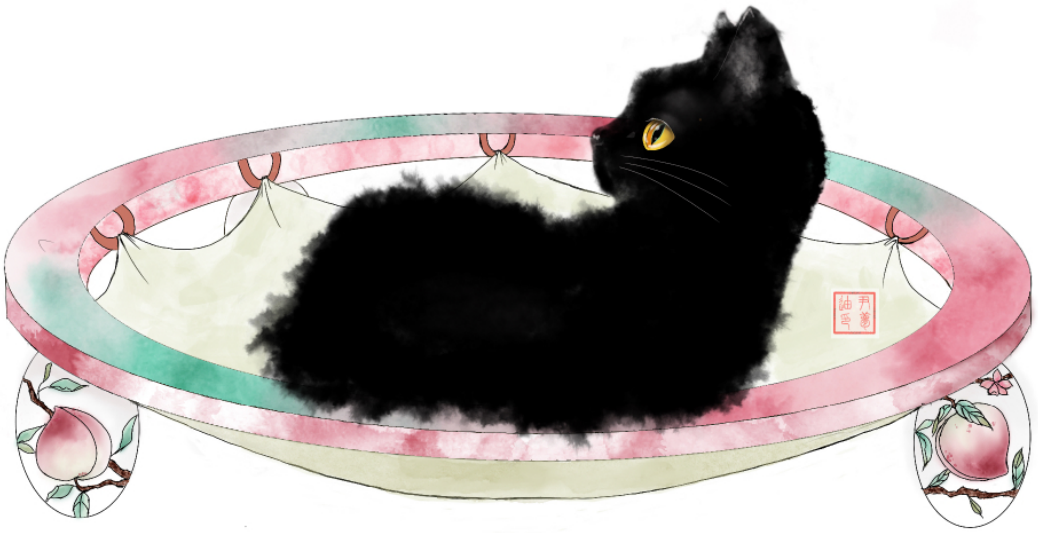
$$\boldsymbol{a}\cdot(\mathsf{F}_2^{-1}\boldsymbol{n}) = 0 \Rightarrow \boldsymbol{a} \perp \mathsf{F}_2^{-1}\boldsymbol{n} \tag{2.74}$$

Noticing that a *simple shear* is defined as **a shear with its deformation gradient of the form** $\mathsf{F} = I + \gamma\boldsymbol{e}_1\otimes\boldsymbol{e}_2$, $\gamma\in\mathbb{R}^+$, $\boldsymbol{e}_1 \perp \boldsymbol{e}_2$ [109], now we can see that F_2 describes a simple shear

that relates phases on either side of the invariant plane to each other. The direction $\boldsymbol{n}_s$ and the modulus S corresponding to this shear are [5]:

$$\boldsymbol{n}_s = \frac{\boldsymbol{a}}{|\boldsymbol{a}|}, \qquad S = |\boldsymbol{a}||\mathsf{F}_2^{-1}\boldsymbol{n}|. \tag{2.75}$$

Thus, we can calculate the direction and modulus of a shear corresponding to a twinning by looking at the deformation gradients at the boundary.



Doves and nursling swallows sleep in a warm spring,
Through the western window, bright sunlight pours in and,
Awakens me from napping.
No issue awaits my tackling,
I enjoy the sunny day by cozily dreaming[5].

—- Su Shi

[5]From *A Spring Day*, translated by Yin Mengdi.

Chapter 3

Triply Periodic Minimal Surfaces

MAUPERTUIS proposed his famous *principal of economy in nature* in 1740s, stating that the total actions required to take to make certain changes in nature should be as small as possible [51]. For instance, the currency prefers a path with the least resistance, people prefer taking a lift to climb stairs, and light prefers to propagate along a route that is most time-saving. The last one was proved by Fermat in 1662 and is called Fermat's principle. Similarly, we can expect a surface, such as a bubble film, to be as "smart" as its peers in nature and thus to form with a minimal surface free energy consumption. Since the surface free energy is proportional to the surface area, we can expect a minimisation of surface free energy as a result of a minimisation of the surface area. Here is how the concept "minimal surface", which is defined as a surface whose (local) area is minimised, comes. When we exert a force on one side of a flat surface, we will observe an increase in the surface area caused by the force, or pressure differences, between two sides. Therefore, the surface area, hence the surface free energy, is proportional to the pressure difference ΔP. According to the Young-Laplace equation [110], $\Delta P \propto M$, where M is the *mean curvature*. Now, we can see that a minimisation of (local) surface area is related to the vanishing of mean curvature everywhere on this surface. And the area will not go to zero due to the existence of a boundary. Mathematically, a surface with its mean curvature vanishing everywhere is a minimal surface. A more detailed explanation will be given in Sec. 3.3.1.

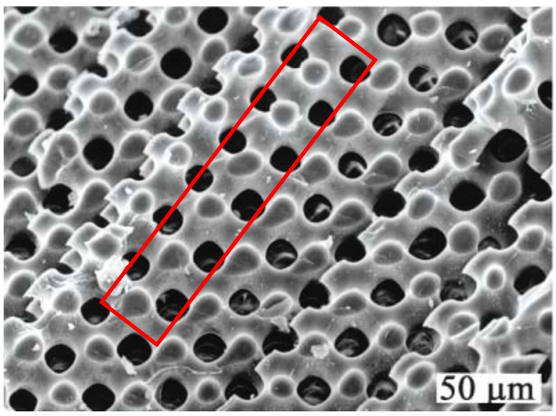
A minimal surface that is triply-periodic, *viz.*, can repeat itself along three independent directions, is said to be a triply-periodic minimal surface (TPMS). Physically, this surface began to draw people's attention when Schnering and Nesper [48–50] discovered that zero equipotential surfaces derived using an electrostatic field resemble TPMS (Sec. 3.1). In addition to crystals, TPMS are also found in various natural systems, such as the amphiphilic membrane [111], the cuticular structures in the butterfly wing scale [112], and the diamond-based structure in beetle scales [113]. It has been observed that the sea urchin skeletal plate re-

sembles the P surface [114], as shown in Fig. 3.1. Based on TPMS, people have also fabricated various materials, including zeolite crystals [115], shellular [111], and self-assembly copolymers [116]. In comparison with traditional materials, these TPMS-based materials are not only superior in strength and stiffness but also much lower in density [111, 115, 117]. The research potential of TPMS is huge as more interesting properties and applications of TPMS are investigated. Our work is using TPMS to study the martensitic phase transformation and the shape memory effect. To help understand the model, we would like to give a mathematical introduction to this concept in this chapter.

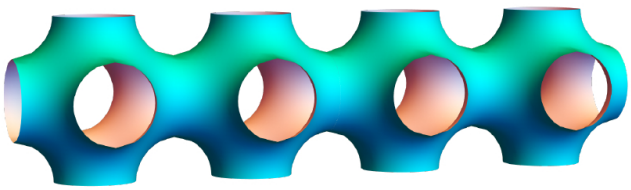
3.1 Zero Equipotential Surfaces

Surfaces of a given charge density in crystals were first studied with the assumption that only the electrostatic field to be considered [48–50, 118–121]. The **electrostatic potential** is expressed as [119]:

$$V_0(\boldsymbol{r}') = \int \frac{\rho(\boldsymbol{r})}{|\boldsymbol{r}-\boldsymbol{r}'|} \mathrm{d}\boldsymbol{r} = \rho(\boldsymbol{r}) * f(\boldsymbol{r}), \qquad f(\boldsymbol{r}) = \frac{1}{|\boldsymbol{r}|}, \tag{3.1}$$



(a)



(b)

Figure 3.1 (a) Cross section through a sea urchin plate. (b) Periodic structure of the P surface. (a) is reproduced with permission [114]. Copyright 2007, Royal Society of Chemistry.

where $\rho(\boldsymbol{r}) * f(\boldsymbol{r})$ stands for the convolution of ρ and f. People employed the *structure factor* that depends on the **arrangement and type of atoms** in a crystal to describe ρ and is defined later in Eq. 3.4. Therefore, for different crystals with different space groups (SGs), we have different expressions of the structure factor, hence different expressions of the *zero equipotential surface* (ZEP), which according to Schnering and Nesper [49], resembles a TPMS. Part of their results are given in Table 3.1 and all TPMS listed in Table 3.1 are given in Fig. 3.2. It is noteworthy that the surfaces given in Fig. 3.2 are not the only possible TPMS. When different symmetries and genus are required, different TPMS can be derived accordingly.

To make the statement clear, we will now provide more details of the structure factor. Consider the diffraction of a plane wave by atoms in a crystal [Fig. 3.3(a)]. The phase difference between waves diffracted by atoms sitting at O and Q will be $\boldsymbol{k}\cdot\boldsymbol{r} - \boldsymbol{k}_0\cdot\boldsymbol{r}$. Since $|\boldsymbol{k}| = |\boldsymbol{k}_0|$, according to Fig. 3.3(b), we have:

$$|\boldsymbol{k} - \boldsymbol{k}_0| = 2|\boldsymbol{k}|\sin\theta = \frac{4\pi\sin\theta}{\lambda}, \tag{3.2}$$

where λ is the wave length. Let us assume that there are n atoms in one unit cell, then the wave diffracted by the whole unit cell is the sum of the wave diffracted by the n atoms. Setting the origin at an arbitrary atom in the unit cell, the phase difference between the ith atom and the origin is $\boldsymbol{k}_i\cdot\boldsymbol{r}_i - \boldsymbol{k}_0\cdot\boldsymbol{r}_i$. Let the diffraction factor of the ith atom be f_i and the diffraction amplitude of an electron be E_e. The diffraction amplitude of a unit cell is now written as [122, 123]:

$$E_{\text{cell}} = E_e\sum_{j=1}^{n} f_j \exp(\mathrm{i}\boldsymbol{s}_j\cdot\boldsymbol{r}_j), \qquad \boldsymbol{s}_j = \boldsymbol{k}_j - \boldsymbol{k}_0. \tag{3.3}$$

Table 3.1 List of ZEP and the TPMS they resemble. X, Y, Z stand for $2\pi x, 2\pi y$ and $2\pi z$, respectively, with $x, y, z \in [0,1) \subset \mathbb{R}$ [49]. The * after each surface is to distinguish the ZEP from the TPMS it resembles. "Gene. Group" means the space group that generates this ZEP.

ZEP	$\{hkl\}$	Gene. Group	Expression	TPMS
P*	$\{100\}$	$Pm\bar{3}m$	$\cos X + \cos Y + \cos Z = 0$	P
D*	$\{111\}$	$Fd\bar{3}m$	$\cos X\cos Y\cos Z + \sin X\sin Y\sin Z = 0$	D
Y**	$\{110\}$	$I4_132$	$\sin X\cos Y + \sin Y\cos Z + \cos X\sin Z = 0$	G
IP$_2$-J*	$\{110\}$	$Im\bar{3}m$	$2(\cos X\cos Y + \cos Y\cos Z + \cos X\cos Z) - (\cos 2X + \cos 2Y + \cos 2Z) = 0$	I-WP
F$_{xx}$-P$_2$F$_z$	$\{111\}$	$Fm\bar{3}m$	$4\cos X\cos Y\cos Z - (\cos 2X\cos 2Y + \cos 2X\cos 2Z + \cos 2Y\cos 2Z) = 0$	F-RD

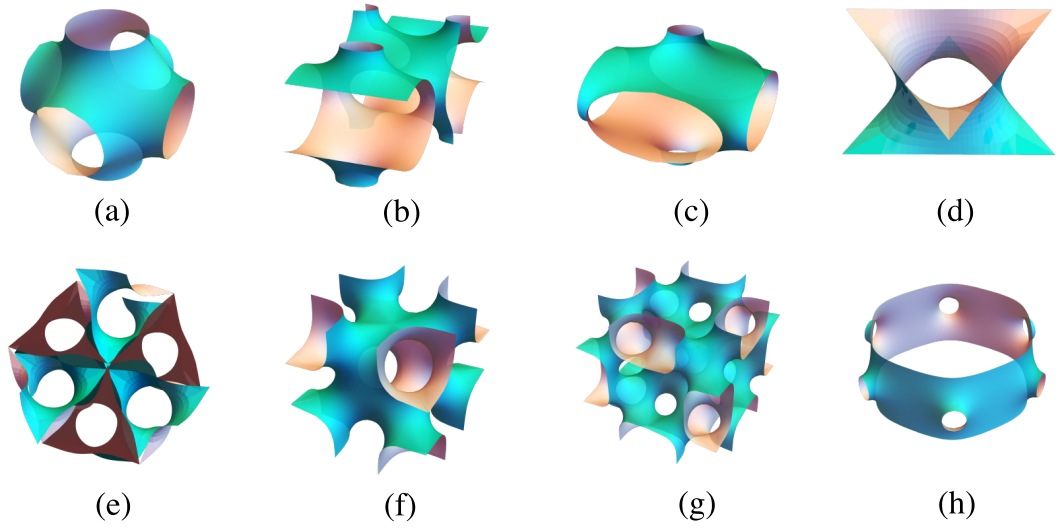


Figure 3.2 Sketch of TPMS: (a) Schwarz P (P), (b) oPb-family (oPb), (c) oPa-family (oPa); (d) Schwarz D (D), (g) Gyroid (G), (f) Schoen's I-WP (I-WP), (g) Schoen's F-RD (F-RD) and (h) Schoen's H′-T (H′-T) surfaces.

The *structure factor* is then defined as [122, 123]:

$$F_{hkl} = \frac{E_{\text{cell}}}{E_e} = \sum_{j=1}^{n} f_j \exp(\mathrm{i}\boldsymbol{r}_j \cdot \boldsymbol{s}_j), \tag{3.4}$$

which is a property including information of the types, numbers and arrangements of atoms in one unit cell. Given Bragg's law:

$$2d_{hkl} \sin\theta = \lambda. \tag{3.5}$$

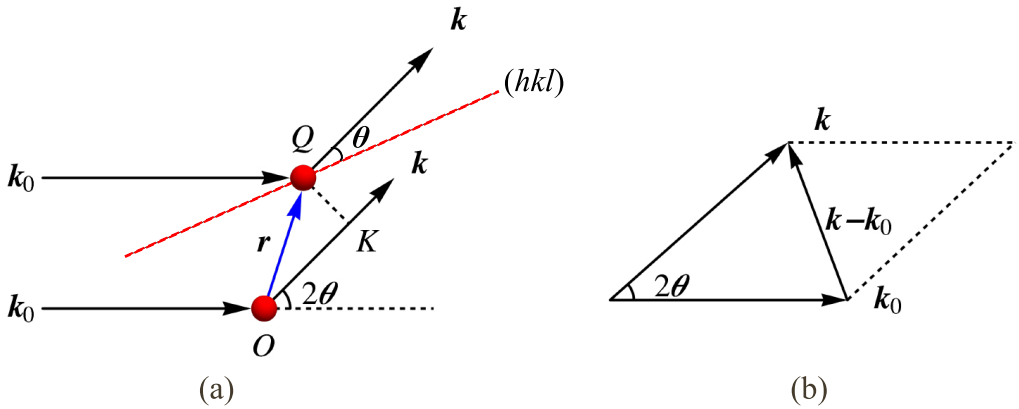


Figure 3.3 Diffraction of a plane wave $\boldsymbol{k}_0$ by atoms labelling by O and Q in a crystal. The wave vector of the original and diffracted wave are $\boldsymbol{k}_0$ and $\boldsymbol{k}$, respectively. In (a), the red dashed line denotes the crystal plane (hkl).

$2\pi/d_{hkl} = |\boldsymbol{K}_{hkl}|$, where $\boldsymbol{K}_{hkl} = h\boldsymbol{b}_1 + k\boldsymbol{b}_2 + l\boldsymbol{b}_3$ is a reciprocal lattice vector and:

$$\boldsymbol{b}_i \cdot \boldsymbol{a}_j = 2\pi\delta_{ij}. \tag{3.6}$$

Here, $\boldsymbol{a}_i$ (resp. $\boldsymbol{b}_i$) are basis vectors of the corresponding unit cell (resp. reciprocal lattice). The phase factor given in Eq. 3.4 is written as:

$$\begin{aligned} \boldsymbol{s}_i \cdot \boldsymbol{r}_i &= \boldsymbol{K}_{hkl} \cdot \boldsymbol{r}_i, \\ &= (h\boldsymbol{b}_1 + k\boldsymbol{b}_2 + l\boldsymbol{b}_3) \cdot (x_i\boldsymbol{a}_1 + y_i\boldsymbol{a}_2 + z_i\boldsymbol{a}_3), \\ &= 2\pi(hx_i + ky_i + lz_i). \end{aligned} \tag{3.7}$$

Therefore, we see $F_{hkl} = \sum_{j=1}^{n} f_j \exp[2\pi\mathrm{i}(hx_i + ky_i + lz_i)]$. Since electrons are distributed continuously in a unit cell, the structure factor can be re-written as:

$$F_{hkl} = V_{\text{cell}} \int_0^1 \int_0^1 \int_0^1 \rho(x,y,z) \exp[2\pi\mathrm{i}(hx+ky+lz)]\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \tag{3.8}$$

where V_{cell} is the volume of the unit cell and F_{hkl} is related to ρ by a Fourier transformation. In a crystal, the triply-periodicity makes it possible for us to expand ρ in a Fourier series:

$$\rho(x,y,z) = \sum_{h,k,l=-\infty}^{\infty} A_{hkl} \exp[-2\pi\mathrm{i}(hx+ky+lz)], \tag{3.9}$$

where A_{hkl} are coefficients. According to Eq. 3.8, we have:

$$\begin{aligned} F_{hkl} = V_{\text{cell}} \int_0^1 \int_0^1 \int_0^1 \sum_{p,q,r=-\infty}^{\infty} A_{pqr} \exp[-2\pi\mathrm{i}(px+qy+rz)] \\ \times \exp[2\pi i(hx+ky+lz)]\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \end{aligned} \tag{3.10}$$

Because when $p \neq h, q \neq k, r \neq l$ the integral will be zero, we then have $F_{hkl} = V_{\text{cell}} A_{hkl}$. Hence:

$$\rho(x,y,z) = \frac{1}{V_{\text{cell}}} \sum_{h,k,l=-\infty}^{\infty} F_{hkl} \exp[-2\pi\mathrm{i}(hx+ky+lz)]. \tag{3.11}$$

Plugging Eq. 3.11 into Eq. 3.1 and using the Fourier convolution theorem, we have:

$$V(\boldsymbol{r}) = \mathscr{F}^{-1}[\mathscr{F}[\rho] \cdot \mathscr{F}[f]] = \frac{1}{\pi V_{\text{cell}}} \sum_{h,k,l=-\infty}^{\infty} \frac{F_{hkl} \exp[-2\pi\mathrm{i}(hx+ky+lz)]}{h^2+k^2+l^2}, \tag{3.12}$$

where $\mathscr{F}[f]$ gives a Fourier transformation of f. Take the space group $Pm\bar{3}m$ formed by one type of atom as an example. Let us consider the $\{100\}$ family of crystal planes, which under

the symmetry of $Pm\bar{3}m$ contains crystal planes:

$$(100),(010),(001),(\bar{1}00),(0\bar{1}0),(00\bar{1}). \tag{3.13}$$

Hence, $h^2+k^2+l^2=1$ and Eq. 3.12 can thus be approximated by Eq. 3.11, which means ZEP now can be approximated by a surface of zero charge density. Therefore, we have Eq. 3.11 written as:

$$\rho \approx \frac{2F_{\{100\}}}{V_{\text{cell}}}[\cos(2\pi x)+\cos(2\pi y)+\cos(2\pi z)]. \tag{3.14}$$

Hence, $\rho=0$ yields:

$$\cos(2\pi x)+\cos(2\pi y)+\cos(2\pi z)=0, \tag{3.15}$$

which resembles the P surface [Fig. 3.2(a)], as shown in Fig. 3.4. In Sec. 4.2.5, we will provide a calculation of this surface to compare it with its corresponding minimal surface.

3.2 Background of Differential Geometry

Let us begin with a *smooth surface* $S\subset\mathbb{R}^3$ that is defined as **a surface that has derivatives of all orders at all points** [124]. Let this surface be parameterised by:

$$\boldsymbol{x}: U\subset\mathbb{R}^2\to S, \qquad S=\boldsymbol{x}(u,v),(u,v)\in U. \tag{3.16}$$

Then at any point $p\in S$, the corresponding normal vector $\boldsymbol{n}_p$ at p is expressed as [47]:

$$\boldsymbol{n}_p=\left.\frac{\partial_u\boldsymbol{x}\times\partial_v\boldsymbol{x}}{|\partial_u\boldsymbol{x}\times\partial_v\boldsymbol{x}|}\right|_p, \tag{3.17}$$

where $\partial_u\boldsymbol{x}$ and $\partial_v\boldsymbol{x}$ are vectors in a plane tangent to S at p. If we intersect S with a plane S', such that $\boldsymbol{n}_p$, $\boldsymbol{v}$ and p are in S', $p\in S$, the curve $\alpha=S\cap S'$ is named the *normal section* of S at p along $\boldsymbol{v}$, where $\boldsymbol{v}$ **is the vector tangent to** α **at** p and is said to be the *direction* of the

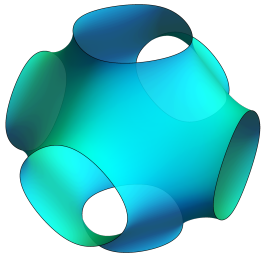


Figure 3.4 Sketch of the P* surface.

normal section (Fig. 3.5). The normal vector has following properties:

1. $\langle \boldsymbol{n}_p, \boldsymbol{n}_p \rangle = 1$, which means if we take the differential, we have:

$$\begin{aligned} \mathrm{d}\langle \boldsymbol{n}_p, \boldsymbol{n}_p \rangle &= \langle \mathrm{d}\boldsymbol{n}_p, \boldsymbol{n}_p \rangle + \langle \boldsymbol{n}_p, \mathrm{d}\boldsymbol{n}_p \rangle \\ &= 2\langle \mathrm{d}\boldsymbol{n}_p, \boldsymbol{n}_p \rangle \\ &= 0 \end{aligned} \tag{3.18}$$

   Hence $\mathrm{d}\boldsymbol{n}_p \perp \boldsymbol{n}_p$, $\mathrm{d}\boldsymbol{n}_p \in T_p(S)$, where $T_p(S)$ is the tangent space formed by all vectors tangent to S at p.

2. $\boldsymbol{n}_s$ is a map $\boldsymbol{n}_s : S \to \mathbb{R}^3$ that gives normal vectors on S. If we restrict $\boldsymbol{n}_s$ on a curve $\alpha(t) \in S, \alpha|_p = 0$ parameterised by t, and $\alpha = \boldsymbol{x}(u(t), v(t))$, then take differential of $\boldsymbol{n}_s(\alpha(t))$ by t, we will see that the differential of $\boldsymbol{n}_s$, $\mathrm{d}\boldsymbol{n}_s$, gives:

$$\begin{aligned} \frac{\mathrm{d}\boldsymbol{n}_s(\alpha(t))}{\mathrm{d}t} &= \mathrm{d}\boldsymbol{n}_s(\alpha'(t)), \\ &= \mathrm{d}\boldsymbol{n}_s(\partial_u \boldsymbol{x}\, u'(t) + \partial_v \boldsymbol{x}\, v'(t)), \\ &= \mathrm{d}\boldsymbol{n}_s(\partial_u \boldsymbol{x})\, u'(t) + \mathrm{d}\boldsymbol{n}_s(\partial_v \boldsymbol{x})\, v'(t), \\ &= N_u u'(t) + N_v v'(t), \end{aligned} \tag{3.19}$$

   Intuitively, $\mathrm{d}\boldsymbol{n}_s$ at a point p measures how fast the normal vector $\boldsymbol{n}_s$ is pulled away

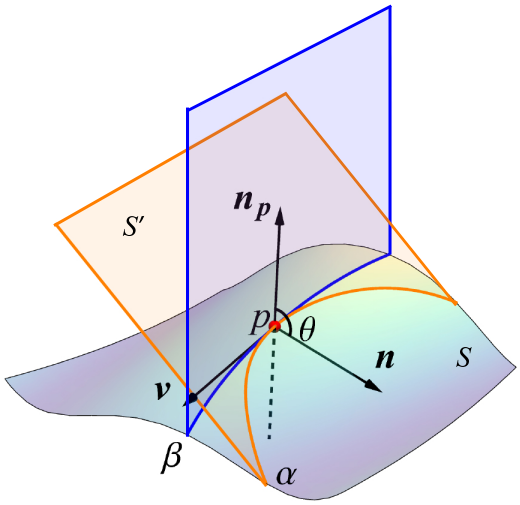


Figure 3.5 Sketch of a normal sections α and β. α (resp. β) is a normal sections along the direction $\boldsymbol{v}$. The normal vector of α at p equal to α'' is $\boldsymbol{n}$. $\boldsymbol{n}_p$ is the normal vector at p of S. The angle between $\boldsymbol{n}_p$ and $\boldsymbol{n}$ is θ.

from p along the curve $\alpha(t)$. Equation 3.19 can also be written as:

$$\begin{aligned}\frac{\mathrm{d}\boldsymbol{n}_s(\alpha(t))}{\mathrm{d}t} &= \frac{\mathrm{d}\boldsymbol{n}_s(u(t),v(t))}{\mathrm{d}t}, \\ &= \partial_u \boldsymbol{n}_s(u,v)u'(t) + \partial_v \boldsymbol{n}_s(u,v)v'(t).\end{aligned} \tag{3.20}$$

Hence we see that $\partial_\sigma \boldsymbol{n}_s(u,v) = \mathrm{d}\boldsymbol{n}_s(\partial_\sigma \boldsymbol{x}) = N_\sigma, \sigma = u,v$.

Consider the curvature of α at p. Projecting α'' onto $\boldsymbol{n}_p$, we see that the *normal curvature* is given by [47]:

$$k_n = \langle \boldsymbol{n}_p, \alpha''(0)\rangle. \tag{3.21}$$

Here, α'' at p is given by:

$$\alpha'' = \left.\frac{\mathrm{d}^2\alpha}{\mathrm{d}s^2}\right|_{s=0}, \tag{3.22}$$

where $\alpha : I \to \mathbb{R}^3$ is a curve parameterised by $t \in I = (a,b)$, $a,b \in \mathbb{R}$, such that $\alpha = \alpha(t)$. Define the *arc length* of α:

$$s = \int_a^t |\boldsymbol{v}(u)|\,\mathrm{d}u, \qquad \boldsymbol{v}(t) = \frac{\mathrm{d}\alpha(t)}{\mathrm{d}t}. \tag{3.23}$$

Parameterising α by s, then we have the expression given by Eq. 3.22. Moreover, at p, we have $s = 0$. The maximum and minimum normal curvatures k_1, k_2 at $p \in S$ are named as the *principal curvatures* at p, whose corresponding vectors tangent to α at p are **orthogonal** to each other [47]. The orthogonality can be understood as follows.

According to Eq. 3.18, we will notice that $\langle \boldsymbol{n}_p, \alpha'\rangle = 0$. Using arc length s for the parameterisation of α, at p we have $s = 0$. Since $\alpha' = \mathrm{d}\alpha/\mathrm{d}s \in T_p(S)$, re-writing Eq. 3.21, we have:

$$\mathrm{d}\langle \boldsymbol{n}_p, \alpha'\rangle = \langle \mathrm{d}\boldsymbol{n}_p, \alpha'\rangle + \langle \boldsymbol{n}_p, \alpha''\rangle = 0, \tag{3.24}$$

which means k_n in Eq. 3.21 can also be expressed as:

$$k_n = -\langle \mathrm{d}\boldsymbol{n}_p, \alpha'\rangle. \tag{3.25}$$

According to Eq. 3.19, $\mathrm{d}\boldsymbol{n}_p$ is a map that maps a vector $\boldsymbol{v} \in T_p(S)$ linearly to another vector in $T_p(S)$. Indeed, Eq. 3.25 means k_n is given by a *quadratic form* Q in $T_p(S)$. More details are contained in Ref. [47]. For a quadratic form, there exists an orthonormal basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ in $T_p(S)$ such that $Q(\boldsymbol{v}) = \lambda_1 x^2 + \lambda_2 y^2$, where $\boldsymbol{v} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2 \in T_p(S)$. This can be proved as follows.

Let V be a 2D linear space and let us restrict the *quadratic form* $Q(x,y) = ax^2 + 2bxy + cy^2$ on a circle $x^2 + y^2 = 1$ such that (x,y) can be parameterised by $(\cos\tau, \sin\tau), \tau \in [0, 2\pi)$, $a,b,c \in \mathbb{R}$. Then Q having a minimum at $(x,y) = (0,1)$ means:

$$\left.\frac{\mathrm{d}Q}{\mathrm{d}\tau}\right|_{\tau=\frac{\pi}{2}} = -2b\cos\pi = 0 \Rightarrow b = 0. \tag{3.26}$$

Let us consider an orthonormal basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2\} \in V$ in which:

$$Q(\boldsymbol{e}_1) = \lambda_1 = a, \qquad Q(\boldsymbol{e}_2) = \lambda_2 = b. \tag{3.27}$$

For a vector $\boldsymbol{v} \in V$ satisfying $|\boldsymbol{v}| = 1$, it can be expressed as $\boldsymbol{v} = x\boldsymbol{e}_1 + y\boldsymbol{e}_2$, with:

$$\boldsymbol{e}_1 = (1,0), \qquad \boldsymbol{e}_2 = (0,1). \tag{3.28}$$

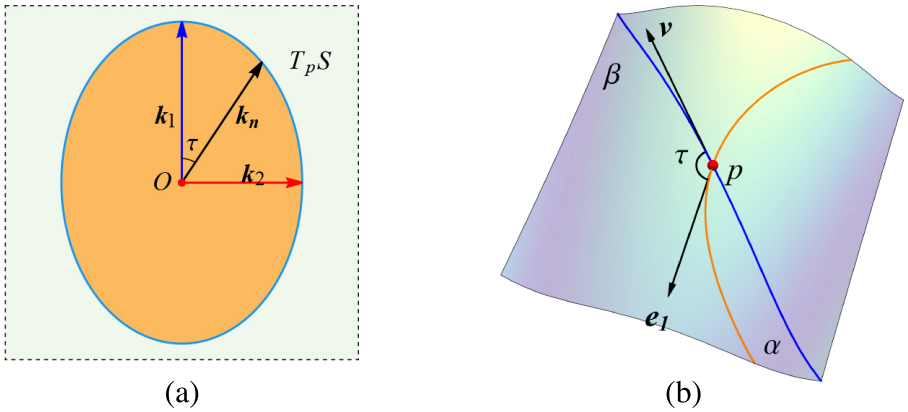
Let Q take a minimum at (0,1), then according to Eq. 3.26 we have:

$$Q(\boldsymbol{v}) = \lambda_1 x^2 + \lambda_2 y^2 \leq \lambda_1 (x^2 + y^2) = \lambda_1 \tag{3.29}$$

Hence λ_1 is a maximum of Q. Furthermore, for $\mathrm{d}\boldsymbol{n}_p$, it means that in the basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$, with the definition of principal curvatures, we have:

$$\mathrm{d}\boldsymbol{n}_p(\boldsymbol{e}_1) = -k_1\boldsymbol{e}_1, \qquad \mathrm{d}\boldsymbol{n}_p(\boldsymbol{e}_2) = -k_2\boldsymbol{e}_2. \tag{3.30}$$

As shown in Fig. 3.6(a), when we rotate the vector $\boldsymbol{v} = \alpha'(0)$ in $T_p(S)$, *viz.*, changing the angle τ, the value of k_n will change accordingly. When $\boldsymbol{v}$ is along $\boldsymbol{e}_1$, then k_n will be k_1,



(a) (b)

Figure 3.6 (a) Sketch of the variation of k_n along a vector $\boldsymbol{v}$ by changing τ in $T_p(S)$. The angle between $\boldsymbol{e}_1$ and $\boldsymbol{v}$ is τ. (b) α and β are principal sections with directions labelled by $\boldsymbol{e}_1$ and $\boldsymbol{v}$, respectively.

similar for $\boldsymbol{e}_2$. Hence, Eq. 3.25 can be expressed as:

$$\begin{aligned} k_n &= -\langle \mathrm{d}\boldsymbol{n}_p(\boldsymbol{e}_1\cos\tau + \boldsymbol{e}_2\sin\tau), \boldsymbol{e}_1\cos\tau + \boldsymbol{e}_2\sin\tau\rangle, \\ &= k_1\cos^2\tau + k_2\sin^2\tau, \end{aligned} \tag{3.31}$$

where τ is the angle between $\boldsymbol{v}$, the direction of the normal section corresponding to k_n and $\boldsymbol{e}_1$, the direction of the normal section corresponding to k_1 [Fig. 3.6(b)]. Equation 3.31 is the *Euler formula* [47]. Moreover, with Euler formula, one can see that the sum of two normal curvatures along two orthonormal directions (not necessarily principal directions) is independent of the choice of these two directions. To be specific, let us consider two normal curvatures k_{n1} and k_{n2} along $\boldsymbol{v}_1, \boldsymbol{v}_2 \in T_p(S)$, $\boldsymbol{v}_1 \perp \boldsymbol{v}_2$:

$$\begin{aligned} k_{n1} + k_{n2} &= (k_1\cos^2\tau + k_2\sin^2\tau) + \left[k_1\cos^2\left(\frac{\pi}{2} - \tau\right) + k_2\sin^2\left(\frac{\pi}{2} - \tau\right)\right], \\ &= k_1 + k_2, \end{aligned} \tag{3.32}$$

This equation will be used later in Sec. 6.3.3.

3.3 Mathematical Description of Minimal Surfaces

3.3.1 Gauss Map, Fundamental Forms and Curvatures

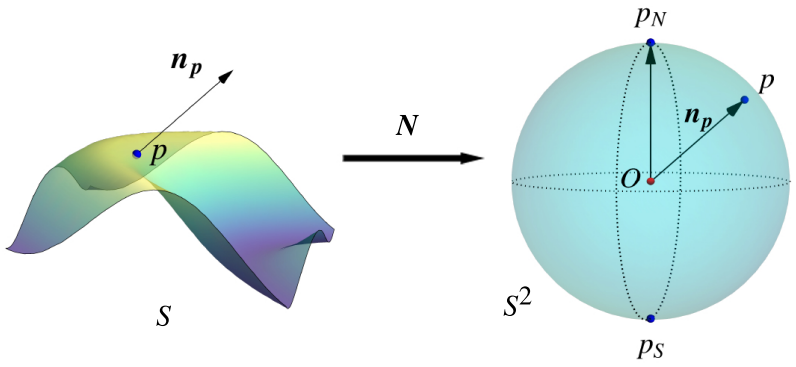
In general, the *Gauss map* $\boldsymbol{N} : S \to S^2$ establishes a correspondence between **points** on an **orientable** surface $S \in \mathbb{R}^3$ with $\boldsymbol{N}$, a differentiable field of unit normal vectors on S, called an orientation of S, and **points** on a **unit sphere**:

$$S^2 := \{(x, y, z) \in \mathbb{R}^3 | x^2 + y^2 + z^2 = 1\}, \tag{3.33}$$

whose unit normal vectors point to the same the direction (Fig. 3.7) [47]. Being *orientable* means that the map defined by Eq. 3.17 is **differentiable** all over this surface, *i.e.*, when described in local coordinates as $\boldsymbol{n} = \boldsymbol{n}(x(u,v), y(u,v), z(u,v))$, functions $x(u,v), y(u,v), z(u,v)$ have derivatives of all orders [47]. A sphere is an orientable surface, while a Möbius strip is not (Fig. 3.8).

In other words, given the coordinate of $p \in S^2$ conforms to the expression of its corresponding unit normal vector, Gauss map maps each point on S to its corresponding normal vector. Now, referring to Eq. 3.19, parameterising u, v by t, the differential of the Gauss map at point $p \in S$ is written as:

$$\mathrm{d}\boldsymbol{N}|_p = \partial_u \boldsymbol{N} u'(0) + \partial_v \boldsymbol{N} v'(0), \tag{3.34}$$



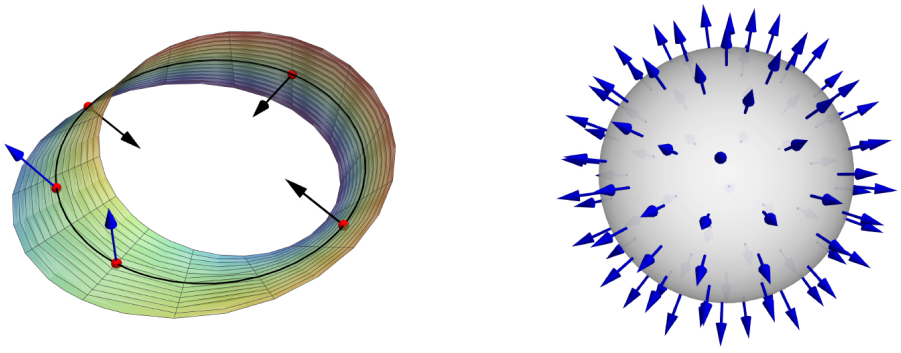
(a) (b)

Figure 3.7 Sketch of Gauss map from a surface S in (a) to the unit sphere S^2 in (b) according to the correspondence between their normal vectors. In (b), O is the sphere centre, p_N and p_S are the North and South poles of S^2, respectively.

with $t = 0$ at p. It is noticeable that $\partial_u \boldsymbol{N}, \partial_v \boldsymbol{N} \in T_p(S)$, and $u'(0), v'(0)$ measure the rate of change of $\boldsymbol{N}$ along $\partial_u \boldsymbol{N}$ and $\partial_v \boldsymbol{N}$. Therefore, $\mathrm{d}\boldsymbol{N}$ gives the rate of $\boldsymbol{N}$ being pulled away from p in the neighbourhood of p.

Thus, the geometric interpretation of $\langle \mathrm{d}\boldsymbol{N}, \boldsymbol{v}/|\boldsymbol{v}| \rangle$ is that it is a projection of the changing rate of $\mathrm{d}\boldsymbol{N}$ along the direction of the curve. Recalling the definition of the arc length s, we see that for a curve $\alpha = \alpha(t), t \in (a,b), \boldsymbol{v} = \mathrm{d}\alpha/\mathrm{d}t$:

$$\frac{\boldsymbol{v}}{|\boldsymbol{v}|} = \frac{\mathrm{d}t}{\mathrm{d}s}\frac{\mathrm{d}\alpha}{\mathrm{d}t} = \frac{\mathrm{d}\alpha}{\mathrm{d}s} = \alpha'. \tag{3.35}$$



(a) (b)

Figure 3.8 Sketch of (a) a Möbius strip and (b) a sphere. Arrows on the surface label unit normal vectors at different points, with blue arrows pointing outward and black arrows pointing inward.

Noticing at p we have $s=0$, which means we can write:

$$\left\langle \mathrm{d}\boldsymbol{N}|_p, \frac{\boldsymbol{v}}{|\boldsymbol{v}|} \right\rangle = \left\langle \mathrm{d}\boldsymbol{N}|_p, \frac{\mathrm{d}\alpha}{\mathrm{d}s}\Big|_{s=0} \right\rangle = -\langle \boldsymbol{N}|_p, \alpha''|_p \rangle = -\langle \boldsymbol{n}_p, \alpha''|_p \rangle, \tag{3.36}$$

where, according to Eq. 3.22, $\alpha'' = \mathrm{d}^2\alpha/\mathrm{d}s^2$. Hence, we have another interpretation of the inner product $\langle \mathrm{d}\boldsymbol{N}, \boldsymbol{v}/|\boldsymbol{v}| \rangle$, which is the projection of the acceleration of the curve α onto the normal vector $\boldsymbol{n}_p$ at p to S. The inner product here is defined as the *second fundamental form* of S at p, *viz.*, [47]:

$$\Pi_2^p(\alpha') = -\langle \mathrm{d}\boldsymbol{N}|_p, \alpha' \rangle, \qquad \alpha' \in T_p(S). \tag{3.37}$$

The matrix form of $\Pi_2^p(\alpha')$ in the basis $\{\partial_u \boldsymbol{x}, \partial_v \boldsymbol{x}\}$ is:

$$\overline{\Pi}_2^p = -\begin{pmatrix} \langle \partial_u \boldsymbol{N}, \partial_u \boldsymbol{x} \rangle & \langle \partial_u \boldsymbol{N}, \partial_v \boldsymbol{x} \rangle \\ \langle \partial_v \boldsymbol{N}, \partial_u \boldsymbol{x} \rangle & \langle \partial_v \boldsymbol{N}, \partial_v \boldsymbol{x} \rangle \end{pmatrix}_p = \begin{pmatrix} \langle \boldsymbol{N}, \partial_{uu} \boldsymbol{x} \rangle & \langle \boldsymbol{N}, \partial_{uv} \boldsymbol{x} \rangle \\ \langle \boldsymbol{N}, \partial_{vu} \boldsymbol{x} \rangle & \langle \boldsymbol{N}, \partial_{vv} \boldsymbol{x} \rangle_p \end{pmatrix} \tag{3.38}$$

Recalling the expression of k_n, then we see that the principal curvatures are **eigenvalues** of $\Pi_2^p(\alpha')$ in an orthonormal basis chosen in $T_p(S)$. Additionally, there exists a nonsingular matrix C that transforms an orthonormal basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ to the original basis:

$$\{\partial_u \boldsymbol{x}, \partial_v \boldsymbol{x}\} = \{\boldsymbol{x}_1, \boldsymbol{x}_2\}, \tag{3.39}$$

with (Appendix B.1):

$$C^{\mathrm{T}} \widetilde{\Pi}_2^p C = \overline{\Pi}_2^p, \tag{3.40}$$

where $\widetilde{\Pi}_2^p$ and $\overline{\Pi}_2^p$ are the matrix forms of Π_2^p in the orthonormal and original bases, respectively. Since the vector length is invariant under coordinate transformations, considering (Einstein summation convention is used here) $\boldsymbol{v} = x^i \boldsymbol{x}_i, i = 1, 2$, we have:

$$\begin{aligned} \langle \boldsymbol{v}, \boldsymbol{v} \rangle &= x^i x^j \langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle \\ &= x^i x^j \langle \boldsymbol{e}_k C_i{}^k, \boldsymbol{e}_l C_j{}^l \rangle \\ &= x^i x^j C_i^{\mathrm{T}\,k} C_j{}^l \langle \boldsymbol{e}_k, \boldsymbol{e}_l \rangle \\ &= x^i x^j C_i^{\mathrm{T}\,k} C_j{}^l \delta_{kl} \\ &= x^i x^j C_i^{\mathrm{T}\,k} C_{jk} \\ &= x^i x^j (C^{\mathrm{T}} C)_{ij}. \end{aligned} \tag{3.41}$$

The meaning of Eq. 3.41 will be clear after the introduction to the first fundamental form. The *first fundamental form* of a surface S parameterised by $\boldsymbol{x}(u,v),(u,v)\in U\subset\mathbb{R}^2$ at $p\in S$ is defined as:

$$\Pi_1^p=\begin{pmatrix}\langle\partial_u\boldsymbol{x},\partial_u\boldsymbol{x}\rangle,\langle\partial_u\boldsymbol{x},\partial_v\boldsymbol{x}\rangle\\ \langle\partial_v\boldsymbol{x},\partial_u\boldsymbol{x}\rangle,\langle\partial_v\boldsymbol{x},\partial_v\boldsymbol{x}\rangle\end{pmatrix}_p \tag{3.42}$$

If we look further at the tensor algebra, this definition precisely conforms to the definition of a *metric* [51]. Hence the first fundamental form is also named the metric, which allows us to measure, for instance, the surface area and the curve length on the surface [47]. More details about metric and metric space will be discussed later in Appendix B.2.1. Now let us go back to Eq. 3.41, according to the definition given in Appendix B.2.1, it can be seen that $C^{\mathrm{T}}C$ is a metric, hence first fundamental form, at $p\in S$. Substitute C^{T} and C into $\det|\widetilde{\Pi}_2^p-\lambda E|=0$ and $\widetilde{\Pi}_2^p\boldsymbol{v}_\lambda=\lambda\boldsymbol{v}_\lambda$:

$$\det|C^{\mathrm{T}}\widetilde{\Pi}_2^pC-\lambda C^{\mathrm{T}}C|=0,\qquad (C^{\mathrm{T}}\widetilde{\Pi}_2^pC-\lambda C^{\mathrm{T}}C)C^{-1}\boldsymbol{v}_\lambda=0, \tag{3.43}$$

which means:

$$(\overline{\Pi}_2^p-\lambda\Pi_1^p)C^{-1}\boldsymbol{v}_\lambda=0\Rightarrow[\overline{\Pi}_2^p(\Pi_1^p)^{-1}-\lambda E]\Pi_1^pC^{-1}\boldsymbol{v}_\lambda=0. \tag{3.44}$$

Therefore, the principal curvatures are eigenvalues of $\overline{\Pi}_2^p(\Pi_1^p)^{-1}$, from which we can define the *mean curvature* M and the *Gaussian curvature* K:

$$M=\frac{1}{2}\mathrm{Tr}[\overline{\Pi}_2^p(\Pi_1^p)^{-1}]=\frac{k_1+k_2}{2},\qquad K=\det[\overline{\Pi}_2^p(\Pi_1^p)^{-1}]=k_1k_2. \tag{3.45}$$

Geometrically, the Gaussian curvature reveals how a surface bends locally, as depicted in Fig. 3.9, while the mean curvature is related to minimising the surface area. In the next two sections we will explain this statement more specifically.

3.3.2 Geometric Interpretation of Gaussian Curvature

If we take a Gauss map $\boldsymbol{N}$ of a plane S (area $A(S)$), since the normal vectors of a plane are pointing in the same direction [Fig. 3.10(a)], all points on a plane will be mapped to the same point on a sphere. Therefore, the ratio of the area of $\boldsymbol{N}(S)$ to $A(S)$ is zero. Then, let us deform this plane such that it conforms to a piece of a unit sphere [purple area in Fig. 3.10(b), area $A'(S)$]. Taking a Gauss map of this purple area, we see that its image is the pink area in S^2, as shown in Fig. 3.10(b). Now, we have the ratio of the pink area to $A'(S)$ becoming one. The deformation here changes the Gaussian curvature, and the ratio reveals the extent to which the

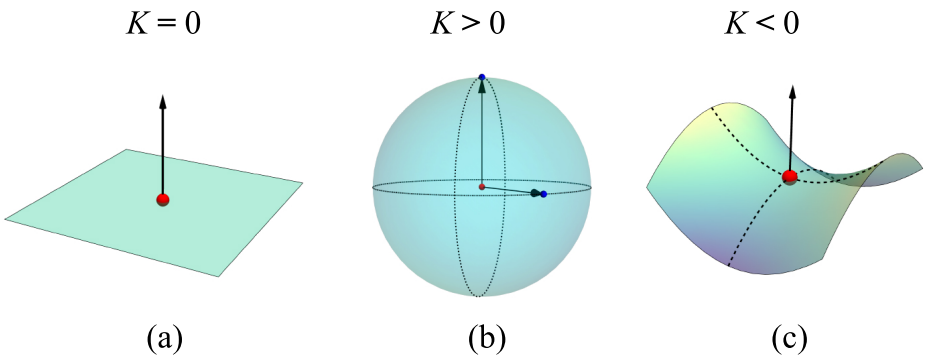


Figure 3.9 Sketch to illustrate different Gaussian curvatures. (a) $K = 0$, if a surface has $K = 0$ everywhere then it is said to be *flat*, a plane is a flat surface. (b) $K > 0$, k_1 and k_2 have the same sign, (c) $K < 0$, k_1 and k_2 have different signs.

surface is curved. Intuitively, this ratio can be understood as the Gaussian curvature, when the surface S, hence the area $A(S)$, becomes very small, as given in Eq. 3.46. A mathematical proof for this statement can be found in Appendix D.3, and now we see that:

$$\lim_{U\to p}\frac{A'(U)}{A(U)} = \lim_{U\to p}\frac{\int_U K|\boldsymbol{x}_u\times\boldsymbol{x}_v|\mathrm{d}u\,\mathrm{d}v}{\int_U|\boldsymbol{x}_u\times\boldsymbol{x}_v|\mathrm{d}u\,\mathrm{d}v} = K|_p, \tag{3.46}$$

where $A(U)$ and $A'(U)$ are the surface area before and after a Gauss map, $U\subset\mathbb{R}^2$ and is mapped to a surface $S\subset\mathbb{R}^3$ by $\boldsymbol{x}: U\to S$. Based on Eq. 3.46, moving the term $\lim_{U\to p}A(U)$ to the right-hand side and taking the integral over the region U, we have:

$$\lim_{U\to p}\int_U A'(U) = \lim_{U\to p}\int_U A(U)K \Rightarrow \int_T \mathrm{d}T = \int_T K\mathrm{d}S, \tag{3.47}$$

where T is a **spherical triangle** that is created by the intersection of three great circles on the unit sphere (Fig. 3.11). Equation 3.47 will be used for the derivation of the Gauss-Bonnet

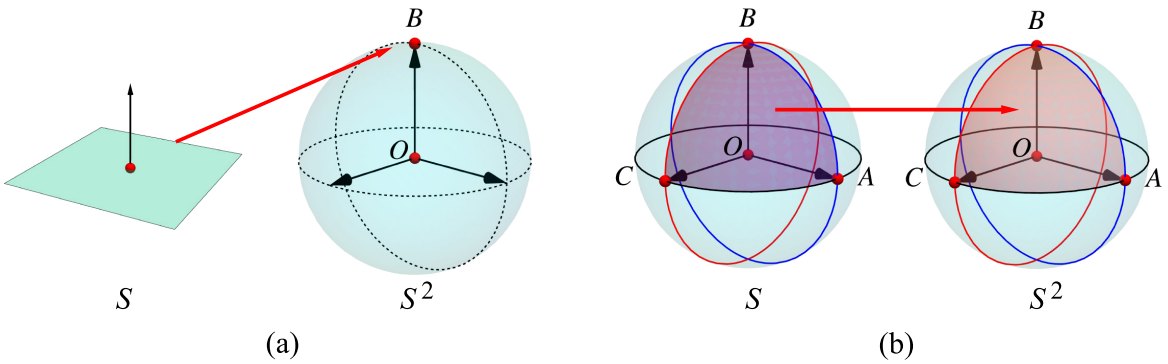


Figure 3.10 Gauss map of (a) a plane and (b) a unit sphere. In (a), all points in the plane will be mapped to the same point B since their normal vectors are pointing in the same direction. In (b), taking a region ABC coloured purple as an example, its image under a Gauss map will be the same as the preimage.

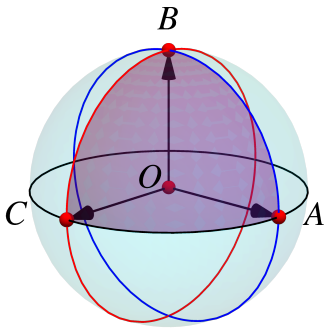


Figure 3.11 Sketch of a spherical triangle ABC coloured in purple on a sphere. O is the centre of the sphere, arrows pointing from the centre to points on the sphere. Black, red and blue circles are great circles on this sphere.

theorem later in Appendix D.7.

3.3.3 Geometric Interpretation of Mean Curvature and Minimal Surface

Considering an orientable surface $S = \boldsymbol{x}(u,v), (u,v) \in U \subset \mathbb{R}^2$, then the surface area is calculated as:

$$A = \int_S \mathrm{d}S, \qquad \mathrm{d}S = \sqrt{g}\,\mathrm{d}u\,\mathrm{d}v, \tag{3.48}$$

where $g = \det \Pi_1^p$ is the determinant of the metric and we label $\Pi_1^p = \{g_{ij}\}, i,j = 1,2$. This can be seen from the fact that the area element will pick up a factor of the Jacobian under a coordinate transformation. For example, under a coordinate transformation from the Cartesian coordinates (x,y) to the polar coordinates (r,θ), we have:

$$|\mathrm{d}x\ \boldsymbol{e}_1 \times \mathrm{d}y\ \boldsymbol{e}_2| = |J|\ |\mathrm{d}r\ \boldsymbol{e}_r \times \mathrm{d}\theta\ \boldsymbol{e}_\theta|, \qquad |J| = \left|\frac{\partial(x,y)}{\partial(r,\theta)}\right|, \tag{3.49}$$

where $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ and $(\boldsymbol{e}_r, \boldsymbol{e}_\theta)$ are basis vectors in the Cartesian coordinates and the polar coordinates, respectively. Since $\Pi_1^p = J^{\mathrm{T}}J$ (using Eq. E.7), we have:

$$|J| = \sqrt{g}. \tag{3.50}$$

Thus, we see the surface element given in Eq. 3.48. More details can be found in Ref. [125]. Add another dimension along $\boldsymbol{n}_p$, suppose that S evolves along $\boldsymbol{n}_p, \forall p \in S$. Parameterising this evoluion by $t \in \mathbb{R}$, then the area is a function of t. No evolution takes place when $t = 0$. Let S be a *closed* manifold, which means it has no boundary ($\partial S = \emptyset$). A 3D *manifold* is a **metric space** S such that for all $s \in S$, there exists a neighbourhood $U \subset S$ of s, U is

homeomorphic to $\mathbb{R}^3$ [126]. More details are in Sec. B.2.1. Take the partial differential of A and let $\partial_t \boldsymbol{x}_t(u,v,t) = \boldsymbol{F}$ (details are in Appendix B.2.2):

$$\partial_t A(t)\Big|_{t=0} = \int_S \partial_t \mathrm{d}S(t)\Big|_{t=0} = \int_S -2M\langle \boldsymbol{F}, \boldsymbol{N}\rangle \mathrm{d}u\,\mathrm{d}v \tag{3.51}$$

Hence, the area will be minimised when $M \equiv 0$ everywhere on the surface. A surface with $M \equiv 0$ all over it is said to be a *minimal surface*. This definition can be understood more intuitively. The concept of "minimal surface" was first proposed to study the formation of a bubble film. If we dip a circle wire into the soap water and then take it out, a flat soap film, whose surface area is minimised [Fig. 3.12(a)], will form on it. However, if we blow that film up from one side, the film will transform into a hemisphere [Fig. 3.12(b)], and the surface area will increase accordingly. The action of "blowing up" also increases the surface tension and, hence, the surface free energy. Therefore, we see that the minimisation of surface area, which is ensured by the vanishing of the mean curvature, will give rise to the minimisation of surface energy. This is where the name "minimal surface" comes from. If we **impose triple-periodicity on minimal surfaces**, these surfaces are named as *triply periodic minimal surfaces* (TPMS), as introduced in Sec. 1. In the next section, we will show how to describe TPMS mathematically.

3.3.4 Projection of TPMS onto $\mathbb{C}^*$ and The Weierstrass Representation

When it comes to the expression of minimal surfaces, there are three most frequently adopted ways:

1. Implicit functions;
2. Parametric functions;
3. Weierstrass representations.



(a) (b)

Figure 3.12 Sketch of (a) a bubble film, (b) a blown-up bubble film, on a circle wire.

As for TPMS, we use the Weierstrass representation, which, roughly speaking, is a method that maps the expression of a surface S defined in $\mathbb{C}$ back into the $\mathbb{R}^3$ space. Before a detailed explanation, we will first introduce the *stereographic projection* $\pi_N : S^2 \to \mathbb{C}$, which is defined as **a projection of points on a unit sphere S^2 from the North/South pole onto the equatorial plane $\mathbb{C}$, such that the whole sphere (except North/South pole) can be parameterised by the plane**. As shown in Fig. 3.13, $p \in S^2$ is projected onto p' in the equatorial plane, which is the intersection of $p_N p$ and the plane. If a minimal surface is triply periodic, we can then label the three axes along which shows the periodicity by $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$. The role a stereographic projection plays here can be understood as projecting a TPMS along one of $\{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}\}$. For example, considering projecting 3D objects onto a 2D plane, as shown in Fig. 3.14, the periodicity along other two directions (or equivalently, the translational symmetry) is preserved.

The periodicity in the direction along which the stereographic projection is taken will not be destroyed because: 1) this projection does not alter the repetition along this direction and 2) this projection does not “collapse” the 3D figure directly [Fig. 3.15(a)] but "expand" the figure into a 2D version that keeps information of positions of points before the projection [Fig. 3.15(b)]. To be specific, in Fig. 3.15(b), we first project this flower onto a sphere (blue region in the middle panel) by a Gauss map (according to the normal vector) then onto $\mathbb{C}$ by a stereographic projection. Taking the point P in the blue region as an example, we see that it will be mapped to $Q \in \mathbb{C}$ under a stereographic projection. Hence, after a stereographic projection of the blue region on the sphere, we will obtain a purple region in the right panel lying in $\mathbb{C}$. The height of P can be expressed as $r \cdot OP/OQ$, where $r = 1$ is the radius of this unit sphere. OP/OQ equal to the ratio of the orthogonal projections of OP and OQ onto $\mathbb{C}$.

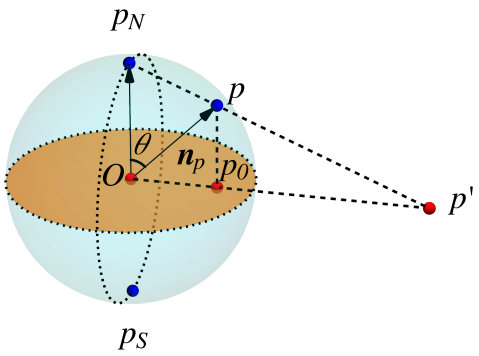


Figure 3.13 Sketch of the stereographic projection. p_N, p_S are the North and South poles of the unit sphere, respectively. p is a point on this sphere, p' is the intersection of the line $p_S P$ and the equatorial plane, pp_0 is perpendicular to the equatorial plane and θ is the angle between pO and $p_N O$.

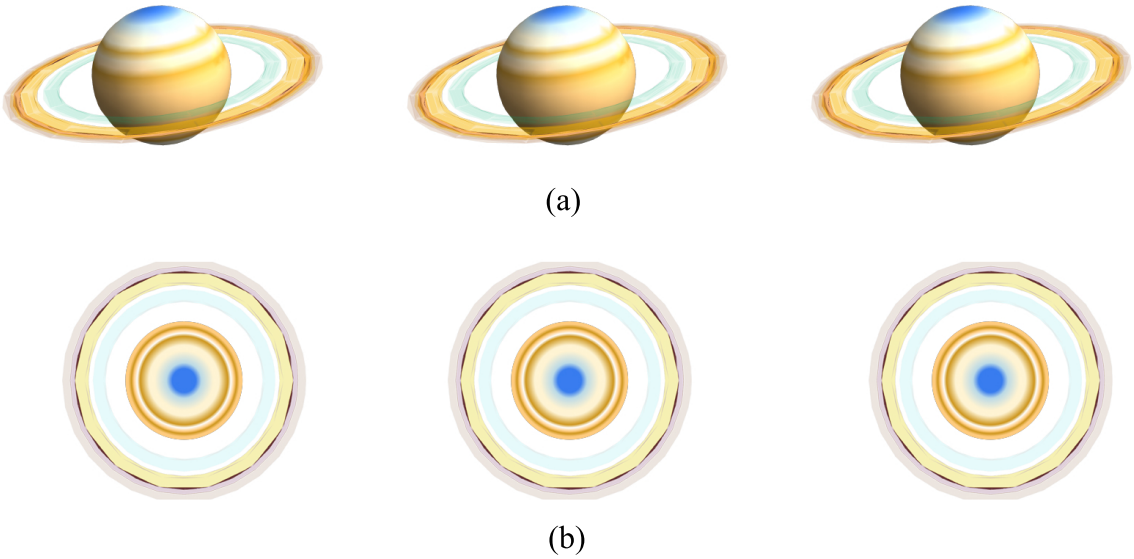


Figure 3.14 Sketch of projecting (a) a line of 3D objects onto a 2D plane (b). It can be seen that the translational symmetry is preserved.

Thus, the height information of P can be told by the length of the orthogonal projection of OP onto $\mathbb{C}$. Hence, after a stereographic projection, we have the image of a triply-periodic object on $\mathbb{C}$ that is not only still possessing a double periodicity, but also preserving the

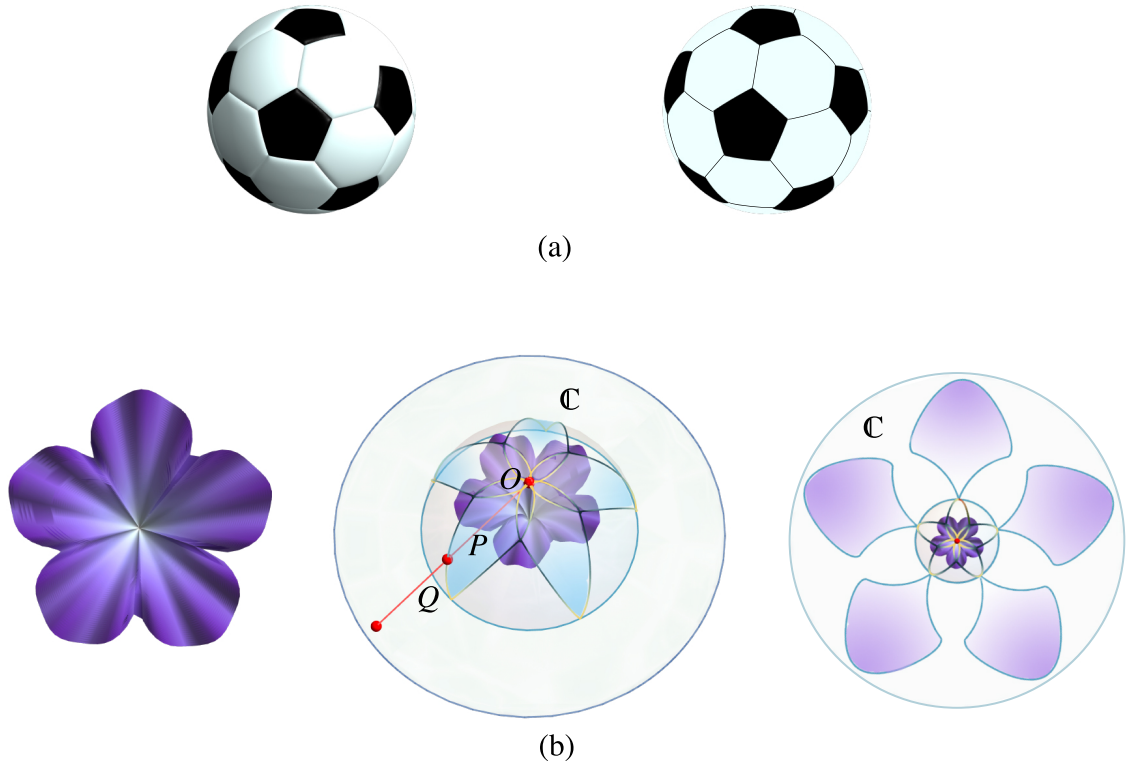


Figure 3.15 Sketch of projections that (a) smashes a soccer into a pie and (b) expands a flower onto the $\mathbb{C}$.

height information along the projection direction. Given that such a projection will neither alter the repetition order nor loose the height information, we can expect the periodicity along the third direction to be preserved. This is important because after the projection, we only need to consider the remaining double periodicity in $\mathbb{C}$, which simplifies the study of a TPMS.

Since a stereographic projection is taken on a sphere, before we operate it onto a TPMS, we need to map this TPMS onto a sphere using the Gauss map (Sec. 3.3.1). Hence, we see that to describe a TPMS mathematically, we need to 1) take a Gauss map that maps a TPMS onto a sphere and 2) take a stereographic projection that maps the sphere onto $\mathbb{C}$. Moreover, since we want the North/South pole to be mapped to ∞, the image of the sphere under a stereographic projection is extended to $\mathbb{C}^* = \mathbb{C} \cup \{\infty\}$, which is identified as the *Riemann sphere*.

Now we can consider the problem in $\mathbb{C}^*$. Let the periodicity be represented by $w_1, w_2 \in \mathbb{C}$ with w_1, w_2 being linearly independent over $\mathbb{R}$, such that the function g_s defined in $\mathbb{C}$ satisfies:

$$g_s(z+w_1) = g_s(z+w_2) = g_s(z), \qquad \forall z \in \mathbb{C}, \qquad \frac{w_2}{w_1} \notin \mathbb{R}, \tag{3.52}$$

and is said to be *doubly-periodic*. Define $\tau = w_2/w_1$, the function $G_s(z) := g_z(w_1 z)$ should have period 1 and τ, which means:

$$G_s(z+n+m\tau) = G_s(z), \qquad n, m \in \mathbb{Z}. \tag{3.53}$$

Similar to crystals exhibiting triple periodicity by the repetition of unit cells, we can also define G_s on a "lattice" Λ in $\mathbb{C}$ generated by $(1, \tau)$ [127]:

$$\Lambda := \{n + m\tau | n, m \in \mathbb{Z}\}. \tag{3.54}$$

The "unit cell" associated with Λ is defined as the *fundamental parallelogram* P_0, which is **the quotient of** $\mathbb{C}$ **and** Λ (Fig. 3.16), expressed as [127]:

$$P_0 = \{z \in \mathbb{C} | z = a + b\tau,\ a, b \in [0, 1)\} = \mathbb{C}/\Lambda. \tag{3.55}$$

As depicted in Fig. 3.16, two edges of P_0 are pairwise identified, which means P_0 behaves like a torus. Recall that we want to describe a TPMS projected onto the $\mathbb{C}^*$ plane. To make things easier, we set the image of this TPMS on $\mathbb{C}^*$ being doubly-periodic. Thus, the function G_s describing the TPMS after a Gauss map defined on $\mathbb{C}^*$ is also doubly-periodic. Mathematically, we use θ **functions** to construct a doubly-periodic function defined on P_0,

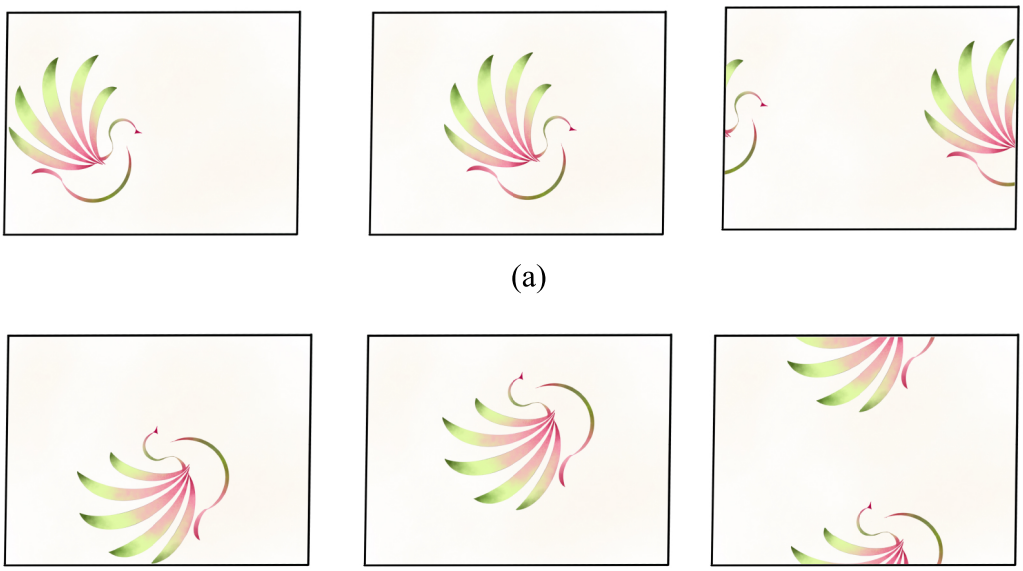


Figure 3.16 Sketch of a parallelogram whose opposite edges are pairwise identified. When we parallel transport a pattern (a) horizontally or (b) vertically, the pattern will behave as moving on a circle. In other words, if we fold this parallelogram such that pairwise edges are connected, we will have a torus.

with [127]:

$$G_s(z) = \prod_{l=1}^{n} \frac{\theta(z-p_l,\tau)}{\theta(z-q_l,\tau)}, \tag{3.56}$$

where p_l and q_l are zeros and poles of $G_s(z)$, which, geometrically, are vertices of a polygon that is the image of the TPMS mapped on $\mathbb{C}^*$ [128]. The θ function is given by [127]:

$$\theta(z,\tau) = \sum_{j\in\mathbb{Z}} \exp[\mathrm{i}\pi(j+\tfrac{1}{2})^2\tau + 2\pi\mathrm{i}(j+\tfrac{1}{2})(z+\tfrac{1}{2})], \tag{3.57}$$

because the periodicity condition requires:

$$G_s(z+\tau) = \exp[2\pi\mathrm{i}\sum_l (p_l - q_l)]G_s(z), \tag{3.58}$$

which means $\sum_j (p_j - q_j) \in \mathbb{Z}$. After the expression of a TPMS on $\mathbb{C}^*$ is clarified, to describe the TPMS in $\mathbb{R}^3$, we need to map these functions defined in a 2D space back to the 3D space. Therefore, we not only need G_s, but also need a function describing the "height" along the direction the stereographic projection takes. Such a height function is named as the *height differential* $\mathrm{d}h$, and **the pair** $(G_s, \mathrm{d}h)$ is called the *Weierstrass representation* of this TPMS [51, 129].

Hence by taking an integral on $\mathbb{C}^*$, we obtain the expression of a TPMS in Cartesian coordinates:

$$\begin{aligned} x &= \int \frac{\mathrm{d}h}{2}\left(\frac{1}{G_s} - G_s\right)\mathrm{d}w, \\ y &= \int \frac{\mathrm{i}\,\mathrm{d}h}{2}\left(\frac{1}{G_s} + G_s\right)\mathrm{d}w, \\ z &= \int \mathrm{d}h\,\mathrm{d}w. \end{aligned} \tag{3.59}$$

Derivation of Eq. 3.59 can be found in Appendix B.2.3. Moreover, the Gaussian curvature on a minimal surface in the Weierstrass representation is expressed as (details in Appendix B.2.4):

$$K = -\frac{4|G_s|^2}{|\mathrm{d}h|^2(1+|G_s|^2)^4}\left|\frac{\mathrm{d}G_s}{\mathrm{d}w}\right|^2. \tag{3.60}$$

3.3.5 Expression of Triply Periodic Minimal Surfaces

Triply-periodic minimal surfaces (TPMS) are minimal surfaces having triple-periodicity. According to previous section, we can define a TPMS on a parallelogram that is doubly-periodic after a projection onto $\mathbb{C}^*$. Hence, we are able to use the θ function defined on the parallelogram to express this surface. Then by taking this expression back to $\mathbb{R}^3$, we obtain the expression of the TPMS in $\mathbb{R}^3$. Noticing that in Eq. 3.59 we have the $z-$component expressed as an integral of the height differential $\mathrm{d}h$. For simplicity, we can take the projection onto $\mathbb{C}^*$ along the z direction, hence $\mathrm{d}h$ can be set to 1 (or some phase factors denoting a twisting), and the Gaussian map G_s is written in a combination of θ functions, as given in Eq. 3.56. Weierstrass representations of TPMS mentioned in this work are listed below [129, 130]. A summary of these Weierstrass representations are given in Table 3.2 and their corresponding

Table 3.2 Weierstrass representations $(G_s, \mathrm{d}h)$ of different TPMS in $\mathbb{R}^3$.

Surface	G_s	$\mathrm{d}h$
Schwarz P	Eq. (3.61) [129, 130]	1
oPb Family	Eq. (3.63) [129, 130]	1
oPa Family	Eq. (3.61) [130]	1
Schwarz D	Eq. (3.61) [44, 45, 131]	$\exp(\frac{1}{2}\mathrm{i}\pi)$
Gyroid	Eq. (3.61) [45, 132]	$\exp(\mathrm{i}\theta), \theta \approx 38.01^\circ$
Schoen's I-WP	Eq. (3.64) [132]	1
Schoen's F-RD	Eq. (3.66) [132]	1
Schoen's H′-T	Eq. (3.68) [132]	1

surfaces are given in Fig. 3.2.

$$G_s(z) \equiv \sqrt{\frac{\theta_1(z-b_s,\tau)\theta_1(z+b_s,\tau)\theta_1(z-c_s,\tau)\theta_1(z+c_s,\tau)}{\theta_1(z-a_s,\tau)\theta_1(z+a_s,\tau)\theta_1(z-d_s,\tau)\theta_1(z+d_s,\tau)}}, \tag{3.61}$$

where $\theta_1(z,\tau)$ is the Jacobi's theta function [127], and:

$$b_s = \tfrac{1}{2} - a_s\,, \qquad c_s = a_s + \tfrac{1}{2}\tau\,, \qquad d_s = b_s - \tfrac{1}{2}\tau\,, \tag{3.62}$$

where $a_s \in \mathbb{R}$ and τ is an imaginary number [129, 133]. This expression can be used to describe the P surface, the gyroid, the D surface and surfaces from the oPa family. oPa family is a set of TPMS obtained by tuning one parameter (a_s) such that the P surface is stretched/-compressed along three mutually orthogonal directions, as shown in Fig. 3.17(a,b).

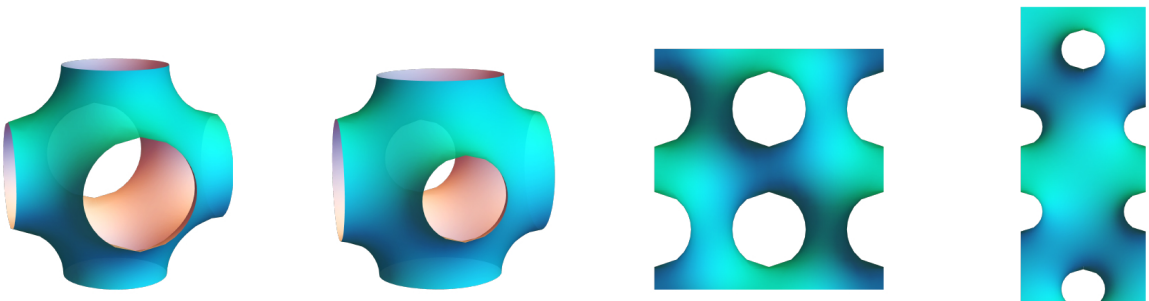
To describe surfaces belonging to the oPb family that connects the P surface to the H surface [Fig. 3.17(c,d)], we can tune two parameters a_s and b_s in the following expression [130]:

$$G_s(z) \equiv \sqrt{\frac{\theta_1(z-a_s,\tau)\theta_1(z+a_s,\tau)\theta_1(z-c_s,\tau)\theta_1(z+c_s,\tau)}{\theta_1(z-b_s,\tau)\theta_1(z+b_s,\tau)\theta_1(z-d_s,\tau)\theta_1(z+d_s,\tau)}}, \tag{3.63}$$

where the parameters should satisfy $c_s = -a_s + (1-\tau)/2$ and $d_s = -b_s + (1+\tau)/2$, $a_s, b_s \in \mathbb{R}$, and τ is determined by a_s, b_s [133].

To describe the I-WP surface, we have:

$$G_s(z) \equiv \exp(\mathrm{i}\tfrac{\pi}{2}) \sqrt{\frac{\theta_1(z-a_s,\tau)^{\frac{3}{2}}\theta_1(z+b_s+\frac{\tau}{2},\tau)^{\frac{3}{2}}}{\theta_1(z+a_s,\tau)^{\frac{3}{2}}\theta_1(z-b_s+\frac{\tau}{2},\tau)^{\frac{3}{2}}}}, \tag{3.64}$$



(a) (b) (c) (d)

Figure 3.17 Sketch of (a) the P surface, (b) a surface from the oPa family, (c) the P surface and (d) a surface from the oPb family.

where:

$$a_s = \frac{1}{6}, \qquad b_s = \frac{1}{3}. \tag{3.65}$$

To describe a F-RD surface, we have:

$$G_s(z) \equiv \sqrt{\frac{\theta_1(z+a_s,\tau)\theta_1(z+b_s,\tau)^{\frac{3}{2}}\theta_1(z-c_s-\frac{\tau}{2},\tau)^{\frac{3}{2}}\theta_1(z-d_s-\frac{\tau}{2},\tau)}{\theta_1(z-a_s,\tau)\theta_1(z-b_s,\tau)^{\frac{3}{2}}\theta_1(z+c_s-\frac{\tau}{2},\tau)^{\frac{3}{2}}\theta_1(z+d_s-\frac{\tau}{2},\tau)^{\frac{3}{2}}}}, \tag{3.66}$$

where:

$$a_s + d_s = \frac{1}{2}, \qquad c_s + b_s = \frac{1}{2}, \qquad 2a_s = 1 - 3b_s. \tag{3.67}$$

To describe the H′-T surface, we have:

$$G_s(z) \equiv \sqrt{\frac{\theta_1(z-a_s,\tau)}{\theta_1(z+a_s,\tau)}}, \tag{3.68}$$

where:

$$a_s = \frac{1}{6}. \tag{3.69}$$



Mind's so bright as a mirror,
That nothing can obstruct its insight.
Hence the whole universe's transparent,
Both inside and outside,
When reflected on by a bright mind [1].

—- Xuan Jue (Master Yong Jia)

[1] Selected from *Song of Enlightenment by Master Yong Jia*, translated by Yin Mengdi.

Chapter 4

Visualisation of Minimal Surfaces Corresponding to Crystals

Density functional theory (DFT) is a method used to solve the Schrödinger equation approximately when considering the many-electron interactions. Such a theory makes it possible to "predict" physical properties and the structure of a material, hence contributes remarkably to the material science. This theory is based on the *Hohenberg–Kohn theorem* [134], saying that **the system ground state energy is determined uniquely by the ground state charge density.** Therefore, instead of calculating the many-body wave function, to study a system, we only need to focus on the charge density ρ and the system energy E can be expressed as a functional of ρ. Kohn and Sham [135] simplified the N-body problem further by absorbing the many-body terms (exchange-correlation potentials) into the "effective potential" then treat the N-body system as N Schrödinger-like noninteracting Kohn–Sham orbitals. However, precise expressions of exchange-correlation (XC) potentials are unknown. Thus, approximations of XC potentials need to be taken in DFT calculations.

One of the most general approximations is the *local density approximation* (LDA) with the XC functionals depending solely on ρ, based on the assumption that the electron gas is locally homogeneous. If we include further the gradient of ρ into the functional, we will obtain a *generalised-gradient approximation* (GGA). GGA corrects the overbinding problem of LDA when dealing with hydrogen-bonded solids [136]. Other XC functionals such as meta-GGA including the Laplacian of ρ [137] and hybrid functionals including empirical Hartree-Fock exchanges [136] are also developed. LDA and GGA are regarded to have a satisfactory balance between the accuracy in calculating material properties and computing costs, hence are adopted most frequently. Other functionals, although can be more accurate in some calculations, are much more time-consuming, hence are less preferred in comparison with LDA and GGA. Table 4.1 gives a comparison between different XC functionals (XCFs).

Table 4.1 Comparison between XCFs, arranged according to "Jacob's ladder" proposed by J. P. Perdew [144, 145].

Central Quantity	XCFs	Details
ρ	LDA	1 Assumes a homogeneous electron gas [144, 145]. 2 Works for materials such as alkali [144, 145]. 3 Overestimates cohesive energies and shows unsatisfactory performance when applied to weakly bonded systems, *e.g.*, H-bonded systems [144].
$\nabla\rho$	GGA	1 Semi-local functionals, consider rapidly varying charge densities [144, 145]. 2 Eclipses LDA when dealing with covalent bonds and weak-bonded systems [144, 145]. 3 Needs correction for the Coulomb self-correlation energy of an electron. Inaccurate when dealing with long-ranged van der Waals (vdW) interactions [144, 145].
$\nabla^2\rho$	meta-GGA	1 Includes $\nabla^2\rho$ and the kinetic energy density [144]. 2 Performs well in calculating surface energies and describing weak-bonded systems [144, 145]. 3 Higher computational cost but does not improve much over GGA functionals [144].

When XCFs are chosen, one can then do simulations to do structural optimisations, calculate energy bands, optical and electron-energy loss spectra, *et al.* [136]. To do DFT calculations, software such as the Vienna *Ab initio* Simulation Package (VASP) [138–141], Quantum Espresso [142] and WIEN2k [143] have been developed.

In this chapter, we will use VASP [138–141] to obtain ground state charge density distribu-

tions of Na, Cu, Al, NiTi and Zr in different phases. XCFs adopted in this work are LDA and GGA functionals. Then we will calculate the surface S of a given ρ. We find that S will converge to a triply-periodic minimal surface (TPMS), such that the point group of the crystal is a subgroup of this TPMS. Recall that the point group of the crystal should be a subgroup of its corresponding Bravais lattice. The group-subgroup relation between the point groups of the crystal and their corresponding TPMS makes it possible for us to make an analogy between the Bravais lattice and the TPMS. That is to say, the TPMS gives the basic structure and periodicity information of a crystal lattice, similar to the Bravais lattice. Moreover, the genus of the TPMS corresponding to two end phases of a martensitic phase transformation is conserved. Also, the skeletal graph of a TPMS resembles the links in the corresponding crystal connecting one lattice point to its nearest neighbours. These two discoveries will be discussed later in Chapter 5.

4.1 Introduction to Density Functional Theory

The main idea of DFT is studying material systems with the help of the ground state charge density distribution ρ. The system energy thus is a functional of ρ. To be specific, let us begin with the system Hamiltonian [146]:

$$\hat{H} = \hat{T}_e + \hat{V}_{\text{ext}} + \hat{V}_{ee}, \tag{4.1}$$

where $\hat{V}_{ee}$ is for interactions between electrons, $\hat{V}_{\text{ext}}$ is the external field and $\hat{T}_e$ is the kinetic energy of N electrons in the system. $\hat{V}_{ee}$, $\hat{V}_{\text{ext}}$ and $\hat{T}_e$ are operators. Let the N-electron wave function be denoted by Ψ_N. We then have:

$$\hat{H}|\Psi_N\rangle = E_N|\Psi_N\rangle. \tag{4.2}$$

Additionally, the wave function yielding the lowest energy E_0 is called the ground state and is labelled as $|\Psi_0\rangle$, *i.e.*,

$$E_0 = \min\{E\} = \min\{\langle\Psi_N|\hat{H}|\Psi_N\rangle\} = \langle\Psi_0|\hat{H}|\Psi_0\rangle. \tag{4.3}$$

The electron density distribution can be expressed as [147]:

$$\rho(\boldsymbol{r}_1) = N\int |\Psi_N(\boldsymbol{r}_1, \boldsymbol{r}_2, \cdots, \boldsymbol{r}_N)|^2 \mathrm{d}\boldsymbol{r}_2 \cdots \mathrm{d}\boldsymbol{r}_N. \tag{4.4}$$

Referring to Eq. 4.3, then in terms of electron density, the ground state energy E_0 is written as:

$$\begin{aligned} E_0 &= \min\{\langle\Psi_N|\hat{T}_e + \hat{V}_{ee}|\Psi_N\rangle + \langle\Psi_N|\hat{V}_{\text{ext}}|\Psi_N\rangle\}, \\ &= \min\{F[\rho] + \int \rho(\boldsymbol{r}) v_{\text{ext}} \mathrm{d}\boldsymbol{r}\}, \\ &= \min\{E[\rho]\}, \\ &= E[\rho_0], \end{aligned} \tag{4.5}$$

where v_{ext} is the external potential. It is noticeable that we do not have an exact expression for the electron-electron energy here, because it not only contains a classical Coulomb term, but also the exchange (caused by indistinguishability of electrons) and correlation (caused by many-body electron-electron interactions) energy, which is unknown. According to the *Hohenberg–Kohn theorem*, **there is a one-to-one correspondence between the ground state charge density and the external field** [134]. Because the external field fixes $\hat{H}$, in turn we will have the ground state energy being a unique functional of the ground state charge density [134]. Hence, instead of specifying the ground state wave function, we can study a system by simply looking at its ground state charge density. As introduced before, Kohn and Sham simplified this problem further by introducing the *Kohn–Sham equation* [135,146,147]:

$$(\frac{-\hbar^2}{2m_e}\nabla^2 + \hat{V}_H + \hat{V}_{\text{XC}} + \hat{V}_{\text{ext}})|\Psi_N\rangle = E_N|\Psi_N\rangle. \tag{4.6}$$

where m_e is the electron mass, $\hat{V}_H$ is an operator corresponding to the the Hartree potential energy E_H that stands for the classic electron-electron repulsions:

$$V_H(\boldsymbol{r}_1) = \int \frac{\rho(\boldsymbol{r}_2)}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|}\mathrm{d}\boldsymbol{r}_2, \qquad E_H = \frac{1}{2}\int V_H(\boldsymbol{r}_1)\rho(\boldsymbol{r}_1)\mathrm{d}\boldsymbol{r}_1, \tag{4.7}$$

where V_H is the Hartree potential. $\hat{V}_{\text{XC}}$ is an operator corresponding to the XC potential energy, whose precise expression is unknown. Therefore, in DFT calculations, an approximation must be made to get a numerical value for the XC energy. Comparisons between different XCFs are given in Table 4.1. Defining the effective potential operator as:

$$\hat{V}_{\text{eff}} = \hat{V}_{\text{H}} + \hat{V}_{\text{XC}} + \hat{V}_{\text{ext}}, \tag{4.8}$$

we will arrive at a Schrödinger equation for noninteracting particles [148]:

$$\left(\frac{-\hbar^2}{2m_e}\nabla^2 + \hat{V}_{\text{eff}}\right)|\phi_i\rangle = \varepsilon_i|\phi_i\rangle, \qquad i = 1,2,\cdots,N. \tag{4.9}$$

Using the *Kohn–Sham orbitals* ϕ_i, taking the N-particle wave function as a product of these N orbitals given ϕ_i are treated as noninteracting particles, we now see [145, 148]:

$$\rho(\boldsymbol{r}) = \sum_{i=1}^{N} f_i |\phi_i|^2, \tag{4.10}$$

where f_i are occupation numbers that follow Fermi-Dirac occupancies at non-zero temperature and are noninteger at zero temperature when the orbitals are degenerate at the Fermi level [148]. The ϕ_i are named as *Kohn–Sham orbitals*, and solution to Eq. 4.9 should yield a **ground state charge density**. The *self-consistent field* (SCF) calculation is designed based on this requirement. To be specific, after the XC functional is chosen, one can use an initial guess of ρ to build the system Hamiltonian. Plugging this Hamiltonian in to Eq. 4.9, the Kohn–Sham equation, one can solve a set of Kohn–Sham orbitals with which one can obtain a charge density ρ'. If ρ' does not converge to ρ, a new charge density will be chosen using orbital optimisation methods such as damping [149], direct inversion in the iterative subspace [150, 151], level shifting [152], *etc.*, and the aforementioned procedure will be repeated. When the calculated charge density and the input charge density at the given interaction agree with each other within a threshold, *i.e.*, achieve a "self-consistency", we can yield an output. The flowchart is shown in Fig. 4.1.

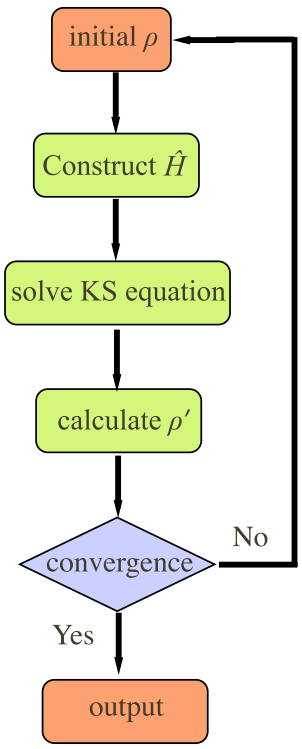


Figure 4.1 Flow chart of how a SCF calculation is realised in DFT calculations.

In real calculations, the Coulomb potential felt by valence electrons is usually replaced by a *pseudopotential* for simplicity [153]. In real systems, the valence-electron wavefunctions are required to be orthogonal to the core-electron states. Hence, in the core region, the valence-electron wavefunctions exhibit a rapid oscillation, and many Fourier terms are needed to describe them. Thus, the all-electron methods require a fine mesh to calculate the valence-electron states, which complicates the calculation. The pseudopotential method treats the system as a collection of valence electrons and ionic cores consisting of nuclei and core electrons strongly bound to the nuclei. The ionic cores are assumed not to participate in chemical bonding. Hence, they will contribute nothing to the structural modification. There are three types of pseudopotentials that are adopted frequently in calculations [153]: 1) norm-conserving pseudopotentials (NCPPs), 2) ultrasoft pseudopotentials (USPPs) and 3) the projector-augmented wave (PAW) method. Details of these potentials can be found in Ref. [153, 154]. Roughly speaking, NCPPs are highly accurate and transferable [153]. This method is based on the requirement that the atomic wavefunctions (eigenvalues) and the pseudo wavefunctions (eigenvalues) are identical beyond some cut-off radius. Thus, this method needs a large basis for the calculation [153, 155]. Hence, NCPPs are accurate but of low efficiency and are suitable for small systems such as molecules. USPPs, however, relax the norm conservation by allowing augmentation charges to remove core electron orbitals that lead to high cut-off energy [153]. Hence, USPPs need fewer basis vectors for the calculation and are, although less accurate, more suitable for large systems such as those involving heavy metals. The PAW method predefines a set of projector functions that are localised in the augmentation region and can map pseudo wavefunctions to all-electron wavefunctions to capture the all-electron properties. Given that solving all-electron wavefunctions in the core region using projector functions is easier than solving them directly, the PAW method is both accurate and computationally efficient [153, 154, 156]. Here, the PAW method is adopted for the following calculations.

Wave functions in crystals should obey the *Bloch's theorem*: [154]:

$$\phi_{\boldsymbol{k}}(\boldsymbol{r}) = \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r})u_{\boldsymbol{k}}(\boldsymbol{r}), \qquad \phi_{\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{R}) = \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R})\phi_{\boldsymbol{k}}(\boldsymbol{r}), \qquad u_{\boldsymbol{k}}(\boldsymbol{r}+\boldsymbol{R}) = u_{\boldsymbol{k}}(\boldsymbol{r}), \tag{4.11}$$

where $\boldsymbol{k}$ is a wavevector in the first Brillouin zone and $\boldsymbol{R} = \sum_{i=1}^{3} m_i \boldsymbol{a}_i$, $m_i \in \mathbb{Z}$. $\boldsymbol{a}_i$ are unit cell basis vectors. Hence, using the periodicity condition, we will see:

$$\exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{a}_i N_i) = 1 \Rightarrow \boldsymbol{k}\cdot\boldsymbol{a}_i = \frac{2\pi n_i}{N_i}, \qquad n_i \in \mathbb{Z}, \tag{4.12}$$

where N_i are numbers of unit cells along $\boldsymbol{a}_i$ in the crystal and we have $\boldsymbol{k} = \sum_i \tau_i \boldsymbol{b}_i$, where $\boldsymbol{b}_i$ are reciprocal lattice basis vectors and $\tau_i = n_i/N_i$, $n_i \in [-N_i/2, N_i/2) \in \mathbb{Z}$. Thus, there will

be one value of $\boldsymbol{k}$ for each unit cell and for calculating properties such as the total energy per unit cell, one only need to sum over $\boldsymbol{k}$ states then take the average, such that [154]:

$$\overline{g} = \frac{1}{n_{\boldsymbol{k}}} \sum_{\boldsymbol{k}} g(\boldsymbol{k}) = \frac{1}{n_{\boldsymbol{k}}} \int_{\mathrm{BZ}} g(\boldsymbol{k}) \mathrm{d}\boldsymbol{k} = \frac{V_{\mathrm{cell}}}{(2\pi)^3} \int_{\mathrm{BZ}} g(\boldsymbol{k}) \mathrm{d}\boldsymbol{k}, \tag{4.13}$$

where $g(\boldsymbol{k})$ is a function of $\boldsymbol{k}$, $n_{\boldsymbol{k}}$ is the number of wave vectors and the integral is taken over the Brillouin zone. V_{cell} is the volume of the unit cell. Moreover, this integrating over the whole Brillouin zone is too computational expensive and can be simplified. Baldereschi [157] proposed that there should be a special point $\boldsymbol{k}_0$ in the Brillouin zone such that $g(\boldsymbol{k}_0) = \overline{g}$. Chadi and Cohen [158] extended this special point to several points then based on this, Monkhorst and Pack [159, 160] developed a $\boldsymbol{k}$-point sampling method such that one can expand $\overline{g}$ in terms of plane waves at the sampled $\boldsymbol{k}$-points. The Monkhorst–Pack method samples **an even number of** $\boldsymbol{k}$ points along three reciprocal lattice vectors within the *irreducible Brillouin zone*, defined as **a portion of the first Brillouin zone that can fill the whole first Brillouin zone with no approximation by taking point group symmetry operations on it** [154], such that all sampled points should have the point group symmetry of this crystal. As shown in Fig. 4.2(a), a square lattice can satisfy the symmetry condition. However, for example, a hexagonal lattice will generate more $\boldsymbol{k}$ points under a C_3 rotational symmetry in the Monkhorst–Pack sampling [Fig. 4.2(b,c)]. Hence, for a hexagonal lattice, the Gamma-centred sampling [Fig. 4.2(d)] that includes the Γ point will be more suitable. According to Eq. 4.10 and Eq. 4.13, the charge density per unit cell now can be calculated as:

$$\rho(\boldsymbol{r}) = \frac{1}{n_{\boldsymbol{k}}} \sum_{\boldsymbol{k},i} f_{i,\boldsymbol{k}} \rho_{\boldsymbol{k},i}, \tag{4.14}$$

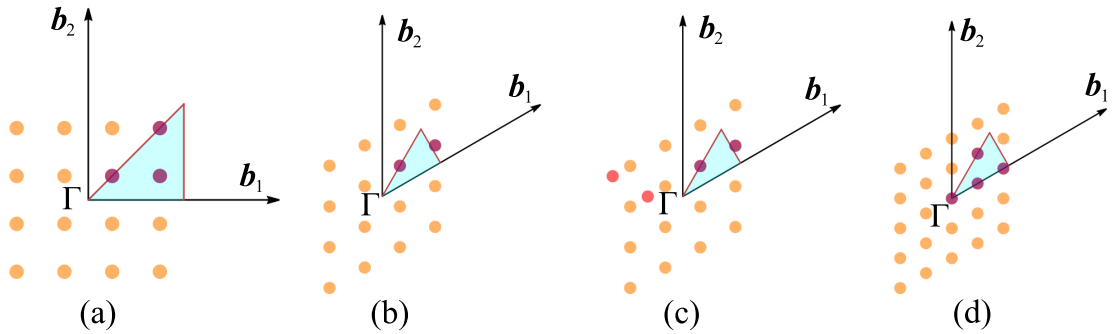


Figure 4.2 Sketch of (a) Monkhorst–Pack sampling by 4×4 of a square lattice in the Brillouin zone and (b) Monkhorst–Pack sampling by 4×4 of a hexagonal lattice in the Brillouin zone. (c) Points in the irreducible Brillouin zone under a 3-fold rotational symmetry of (b). (d) Gamma-centred sampling by 5×5 of a hexagonal lattice. Cyan area bonded by red lines are the irreducible Brillouin zones. Orange points are for the sampled $\boldsymbol{k}$ points and purple points are points in the irreducible Brillouin zone. Red points are additional $\boldsymbol{k}$ points under the C_3 rotation.

where $\rho_{\boldsymbol{k},i} = |\phi_{\boldsymbol{k},i}(\boldsymbol{r})|^2$, $\phi_{\boldsymbol{k},i}(\boldsymbol{r})$ are the Kohn–Sham orbitals at $\boldsymbol{k}$ state, i is for the number of particles in the system. According to Bloch's theorem, using Eq. 4.11, we have [153]:

$$\phi_{\boldsymbol{k},i}(\boldsymbol{r}) = \frac{1}{\sqrt{V_{\text{crystal}}}} \sum_m c_{m,i}(\boldsymbol{k}) \exp[\mathrm{i}(\boldsymbol{k}+\boldsymbol{G}_m)\cdot\boldsymbol{r}] = \frac{1}{\sqrt{n_{\text{cell}}}} \exp(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}) u_{\boldsymbol{k},i}(\boldsymbol{r}), \tag{4.15}$$

where $c_{\boldsymbol{k},i}$ are expansion coefficients, $V_{\text{crystal}} = V_{\text{cell}}\, n_{\text{cell}}$, n_{cell} is the number of unit cells in the crystal. $\boldsymbol{G}_m = m_i \boldsymbol{b}_i$, $m_i \in \mathbb{Z}$, $i = 1,2,3$, and $\boldsymbol{b}_i$ are reciprocal lattice basis vectors and m runs from one to n_{G}. n_{G} is the number of reciprocal lattice vectors needed to describe the density. Since we require Kohn–Sham orbitals to be orthonormal, we see:

$$\int_{\text{cell}} u^*_{\boldsymbol{k},i}(\boldsymbol{r}) u_{\boldsymbol{k},i'}(\boldsymbol{r}) \mathrm{d}\boldsymbol{r} = \frac{1}{V_{\text{cell}}} \sum_m c_{m,i}(\boldsymbol{k}) c_{m,i'}(\boldsymbol{k}) = \delta_{i,i'}, \tag{4.16}$$

$c_{m,i}(\boldsymbol{k})$ can be solved in $\boldsymbol{k}$-space (details are in Appendix C.1). Also, in $\boldsymbol{k}$-space, the Bloch state is expressed as:

$$u_{\boldsymbol{k},i}(\boldsymbol{G}_m) = \frac{1}{\sqrt{V_{\text{cell}}}} c_{m,i}(\boldsymbol{k}). \tag{4.17}$$

In real space, with Eq. 4.15, the charge density $\rho_{\boldsymbol{k},i}(\boldsymbol{r})$ is expressed as:

$$\rho_{\boldsymbol{k},i}(\boldsymbol{r}) = \frac{1}{n_{\text{cell}}} |u_{\boldsymbol{k},i}(\boldsymbol{r})|^2, \tag{4.18}$$

Therefore, to calculate the charge density, we can first obtain the Bloch states in $\boldsymbol{k}$-space using Eq. 4.17, then transform $u_{\boldsymbol{k},i}(\boldsymbol{G}_m)$ back to real space and next use Eq. 4.18. Noticing that n_{G} reciprocal lattice vectors in $\boldsymbol{k}$-space means dividing the real space in to a n_{G}-point grid, we thus can get the numerical value of $\rho(\boldsymbol{r}_m)$ at the mth point on this grid. The number n_{G} is used as a variant in our later calculations.

4.2 Method and Results

Results obtained by ignoring electron-electron interactions are widely applied in modelling minimal surface-based materials [117]. However, in real systems, electron-electron interactions should always be included and will play an important role in calculating the electronic properties of crystals. Therefore, we take a step further by including the exchange-correlation energy when calculating the surface S_ρ of a given ρ in crystals and discovered that S_ρ converges to TPMS. Crystals we calculated in this work are sodium (Na), copper (Cu), aluminium (Al), zirconium (Zr) and nitinol (NiTi), with details presented in the following sections and general results given in Table 4.2.

We first take a structure relaxation to optimise the structure of the crystal, then take a SCF

Table 4.2 Symmetry groups of materials studied in this work with their corresponding TPMS. SG and PG stand for space group and point group, respectively.

Material	SG (Material)	PG (Material)	TPMS	SG (TPMS)	PG (TPMS)
Na	$Im\bar{3}m$	O_h	I-WP	$Im\bar{3}m$	O_h
Na	$P6_3/mmc$	D_{6h}	H′-T	$P6/mmm$	D_{6h}
Al	$Fm\bar{3}m$	O_h	F-RD	$Fm\bar{3}m$	O_h
Al	$Im\bar{3}m$	O_h	I-WP	$Im\bar{3}m$	O_h
Cu	$Fm\bar{3}m$	O_h	F-RD	$Fm\bar{3}m$	O_h
Cu	$Im\bar{3}m$	O_h	I-WP	$Im\bar{3}m$	O_h
NiTi	$Pm\bar{3}m$	O_h	P	$Pm\bar{3}m$	O_h
Zr	$Im\bar{3}m$	O_h	I-WP	$Im\bar{3}m$	O_h
Zr	$P6_3/mmc$	D_{6h}	H′-T	$P6/mmm$	D_{6h}
Zr	$P6/mmm$	D_{6h}	H′-T	$P6/mmm$	D_{6h}

calculation to obtain the ground state charge density. All calculations were carried out using VASP [138–141], with the energy cut-off being set to be 520 eV and the $\boldsymbol{k}$ spacing being set to be 0.02Å^{-1}. As mentioned before, we estimated the extent to which S_ρ converges to its corresponding TPMS S using root-mean-square (RMS). To calculate RMS between S_ρ and S, we first take a grid of 100×100 mesh points on each elementary piece of TPMS. The mesh grid is taken in $\mathbb{C}$ because the TPMS is expressed using a Weierstrass representation that maps points in $\mathbb{C}$ to $\mathbb{R}^3$ (see Sec. B.2.3 and Sec. 3.3.5). By taking symmetry operations on these elementary pieces, we can obtain the "unit cell" of their corresponding TPMS. Therefore, we have 8×10^4 points on the P surface, 1.6×10^5 points on the I-WP surface and 3.2×10^5 points on the F-RD surface. Then we extract the position of points on S_ρ by selecting points falling in an interval $[\rho_a, \rho_b]$ (the number of points falling in this interval is labelled by n_{ab}), hence ρ corresponding to S_ρ is approximated by $(\rho_a + \rho_b)/2$. Moreover, we have $\rho_{\min} < \rho_a < \rho_b < \rho_{\max}$, where $\rho_{\min}$ and $\rho_{\max}$ are the minimum and the maximum charge density value in the VASP calculation. Next, by scaling TPMS to make it larger or smaller, we can find the smallest distance d_i between each point on S_ρ and S. After that, we can calculate the corresponding RMS from the data set of these smallest distances using the equation:

$$\text{RMS} = \sqrt{\frac{\sum_i^{n_{ab}} d_i^2}{n_{ab}}}, \tag{4.19}$$

and plot the RMS against ρ to see the convergence of S_ρ to S. Following are the explicit steps of how the RMS analysis is conducted:

1. Calculating the charge density distribution of a given phase;

2. Plotting surfaces of a given ρ using the charge density obtained in previous step;

3. Comparing this surface with all TPMS and then selecting the one that fits S_ρ best;

4. Calculating the shortest distance between each point on S_ρ and S;

5. Calculating the RMS of the data set of shortest distances;

6. Scaling S up and down, to find the smallest RMS.

4.2.1 Sodium

Sodium (Na) is an alkali metal in the first column of the periodic table with atomic number 11. Its electronic configuration is $1s^2\,2s^2\,2p^6\,3s^1$. Theoretically, Na can undergo a martensitic phase transformation from its austenite state (BCC, $Im\bar{3}m$) to its martensite state (HCP, $P6_3mm/c$) when cooled below 36K [161, 162]. In our calculation, BCC Na is sampled on a Monkhorst–Pack scheme and HCP sodium is sampled on a Gamma-centred scheme. The XCF chosen here is LDA. Figure 4.3(d,h) plots the RMS against the charge density value for Na in the BCC and HCP lattice for several values of the mesh density $(n_G)^{1/3}$. For Na in the BCC lattice, the minimum decreases significantly with increasing mesh density, and a sharp minimum can be observed for all meshes. With increasing mesh density, the minimum

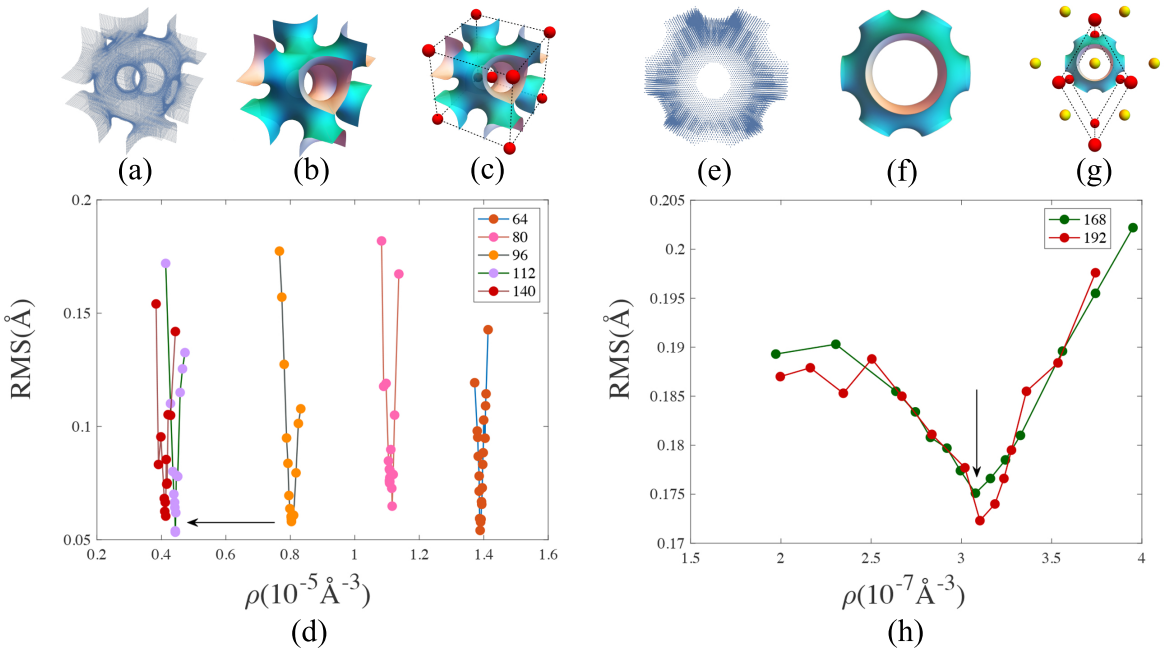


Figure 4.3 (a) Surface of a given ρ of Na in the BCC lattice. (b) The I-WP surface. (c) Plots of the BCC lattice against the I-WP TPMS. Red spheres in (c) represent atoms. (d) RMS against the value of ρ of Na in the BCC lattice for several values of the mesh density $(n_G)^{1/3}$ given in the legend. (e) Surface of given ρ in HCP Na. (f) The H′-T surface. (g) Plots of the HCP lattice against the H′-T TPMS. Red spheres in (g) represent atoms. (h) RMS against the value of ρ of Na in the HCP lattice for several values of the mesh density $(n_G)^{1/3}$ given in the legend.

converges to $\rho \approx 4.3 \times 10^{-6}$Å^{-3}. The converged surface shown in Fig. 4.3(a) resembles the I-WP surface corresponding to the BCC lattice of Na [Fig 4.3(b)]. Positions of the Na atoms on the TPMS are given in Fig. 4.3(c). Thus, for BCC sodium with the LDA, we see convergence of the surfaces of constant charge density to the I-WP minimal surface. Convergence of surface of a given ρ in HCP Na towards the H′-T surface is given Fig. 4.3(e,f,g,h). In this case the minimum converges to $\rho \approx 3.1 \times 10^{-7}$Å^{-3}.

4.2.2 Aluminium

Aluminium is a combustible metal belonging to the Boron group with the atomic number 13. It has three valence electrons per atom $3s^2\,3p$, with the third valence electron partially occupying the second and third p-bands, giving rise to hybridisation between some of its $3p$-bands with some admixture of its $3s$-bands. Aluminium has the FCC lattice ($Fm\bar{3}m$) at standard temperature and pressure and can transfer to a BCC lattice ($Im\bar{3}m$) under 363 GPa at 920K [163]. A Gamma-centred scheme is adopted to sample its reciprocal lattice. Due to the hybridisation, ρ in Al cannot be treated as a homogeneous gas. Therefore the GGA-PBE XC potential is chosen here. Figure 4.4(d) shows the RMS against ρ for Al in the FCC lattice as a function of the indicated mesh density $(n_G)^{1/3}$. The arrow indicates the

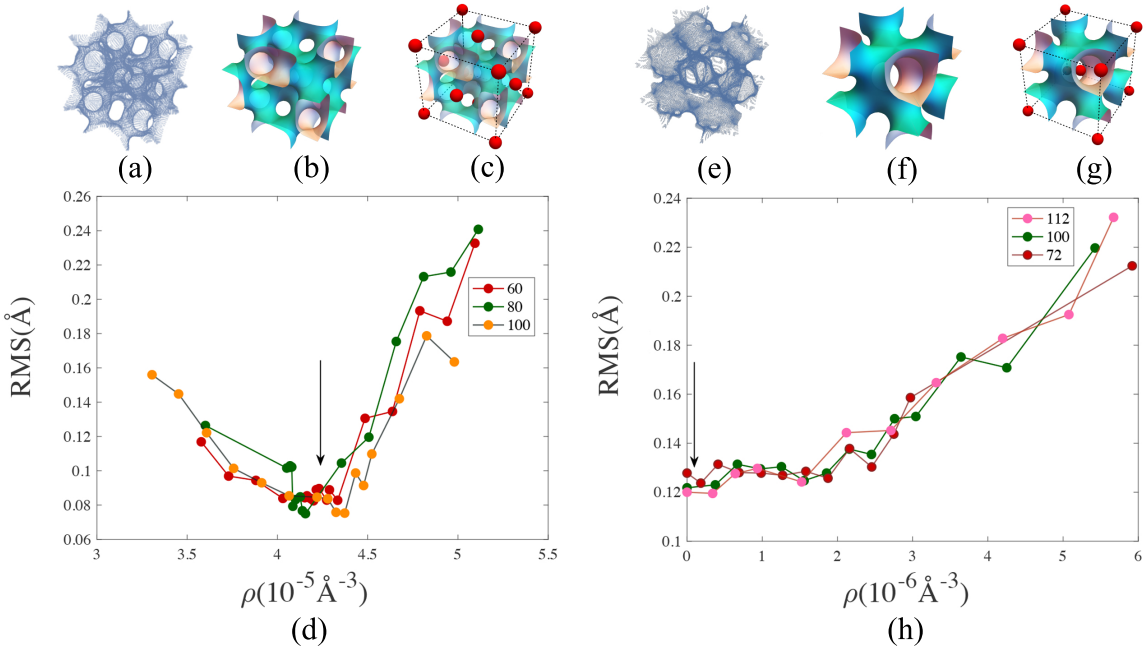


Figure 4.4 Surfaces of a constant ρ in Al in (a) an FCC lattice and (e) a BCC lattice. (b) (resp. (f)) the F-RD (resp. I-WP) surface. (c) and (g) are plots of the FCC and BCC lattice against the F-RD and I-WP surface, respectively. Red spheres in (c) and (g) represent atoms. (d) and (h) show RMS against charge density ρ of Al in the FCC and BCC lattice. Mesh densities $(n_G)^{1/3}$ are shown in the legend. Arrows show where the minima are in the two small panels.

minimum value of the RMS. For all meshes, the RMS shows sharp minima that converge around $\rho \approx 4.2 \times 10^{-5}$ Å^{-3}. The converged surface shown in Fig. 4.4(a) resembles the F-RD surface that corresponds to the FCC lattice of Al [Fig. 4.4(b)]. Positions of the Al atoms with respect to the TPMS are given in Fig. 4.4(c). Similarly, we see for the BCC Al, the minima converge around $\rho \approx 1 \times 10^{-7}$ Å^{-3}, and the surface of a given ρ resembles the I-WP surface that corresponds to the BCC lattice of Al. Thus, we see a convergence of the surfaces of a constant charge density to the F-RD surface for the FCC Al and the I-WP for the BCC Al.

4.2.3 Copper

Copper (Cu) is a transition element having 11 valence electrons $3d^{10}4s^{1}$ with atomic number 29. The $3d$ electrons are hybridised with the delocalised $4s$ electrons. Cu has an FCC lattice ($Fm\bar{3}m$) at room temperature and standard pressure and transforms to a BCC lattice ($Im\bar{3}m$) under an internal strain [164, 165] or through a shock [166]. A Gamma-centred scheme is applied to sample its reciprocal lattice, and similar to Al, GGA is chosen here to calculate the exchange-correlation energy. Figure 4.5(d) (resp. (h)) shows the RMS against the charge density of Cu in an FCC (resp. BCC) lattice as a function of the indicated mesh density $(n_G)^{/3}$. For all meshes, the RMS shows sharp minima in the same region of constant density

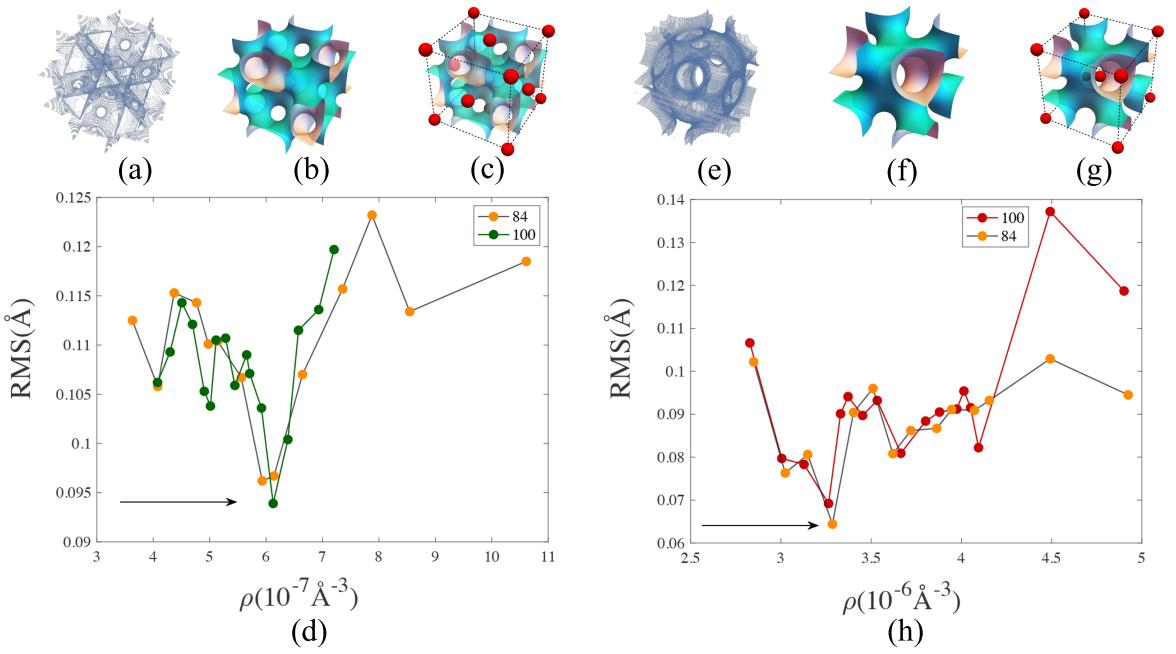


Figure 4.5 Charge density distributions of constant ρ of Cu in (a) the FCC lattice and (e) the BCC lattice. (b) and (f) are F-RD and I-WP surfaces, respectively. (c) and (g) are plots of the FCC and BCC lattice against the F-RD and I-WP surface, respectively. Red spheres in (c) and (g) represent atoms. (d) and (h) show RMS against charge density ρ of Cu in an FCC and a BCC lattice. Mesh densities $(n_G)^{/3}$ are shown in the legend. Arrows show where the minima are in the two small panels.

$\rho \approx 6.1 \times 10^{-7}$Å^{-3} (resp. $\rho \approx 3.3 \times 10^{-6}$Å^{-3}), with convergence increasing slightly with the increasing mesh density. The converged surface resembles the F-RD surface for the FCC lattice and the I-WP surface for the BCC lattice. Positions of the Cu atoms with respect to the TPMS are shown in Fig. 4.5(c,g). Thus, for FCC (resp. BCC) Cu with the GGA-PBE XC potential, we see convergence of the surfaces of constant charge density to the F-RD (resp. I-WP) surface.

4.2.4 Zirconium

Zirconium (Zr) is also a transition metal, and its atomic number is 40, having the valence electrons $4d^2\,5s^2$. Zr is in an HCP lattice ($P6_3mm/c$) at standard pressure and room temperature and can undergo martensitic phase transformations. It has two martensite phases: 1) the β phase in a BCC lattice ($Im\bar{3}m$) at high temperature and 2) the ω phase in a hexagonal lattice ($P6/mmm$) under high pressure [167]. A Gamma-centred scheme is adopted to sample its reciprocal lattice. The LDA XC is used here to calculate the exchange-correlation energy of Zr in BCC and hexagonal lattices, and the GGA-PBE XC is adopted here for Zr in the HCP lattice. Figure 4.6(a,b,c) plots the converged surface of a given ρ, the converged surface against the H′-T surface, and atom positions against the H′-T surface, respectively. Figure

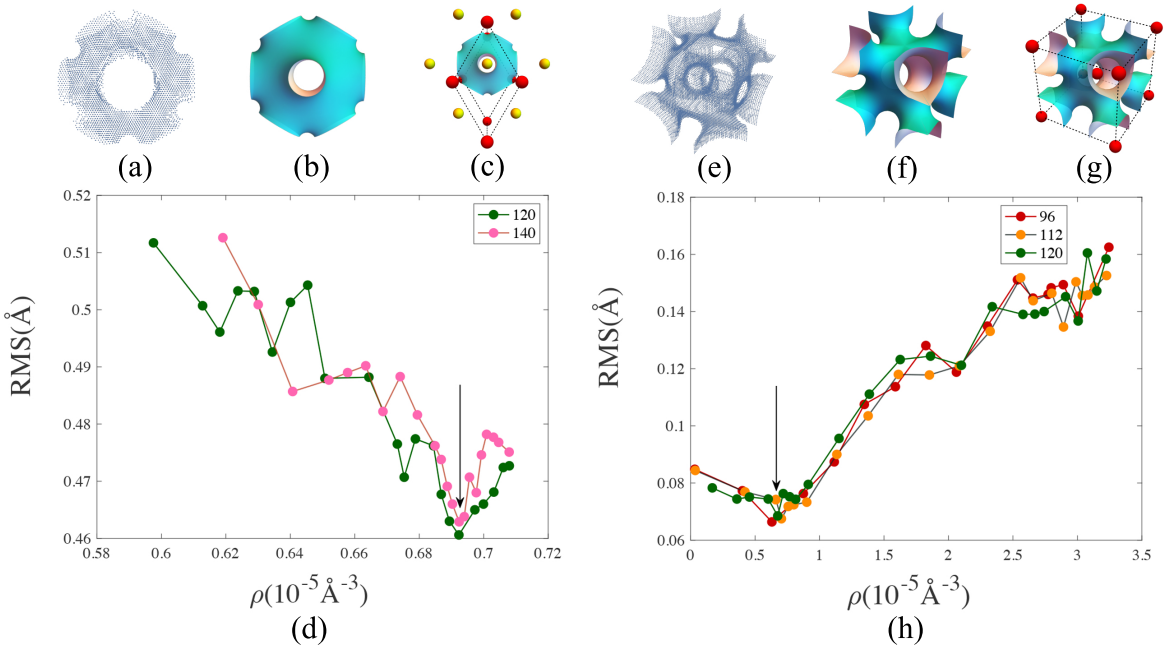


Figure 4.6 Surfaces of a given ρ in (a) (resp. (b)) Zr in HCP (resp. BCC) lattice in $P6_3/mmc$ (resp. $Im\bar{3}m$) SG. Plot of ρ against (b) the H′-T surface for the HCP Zr and (f) the I-WP surface for the BCC Zr. (c,g) Sketches of atoms in the HCP and BCC unit cells against the H′-T and I-WP surface. Red and yellow spheres represent Zr atoms. (d,h) RMS against charge density ρ of Zr in the HCP and BCC lattice. Mesh densities are shown in the legend. Arrows show where the minima are in the two small panels.

4.6(d) plots the RMS against the charge density for HCP Zr with different mesh densities, and it can be observed that the minimum converges to $\rho \approx 6.9 \times 10^{-6}$Å^{-3}. Figure 4.6(e) shows the surface converges to the I-WP TPMS [Fig. 4.6(f)]. Atom positions of the BCC zirconium with respect to the I-WP surface are shown in Fig. 4.6(g). Figure 4.6(h) shows the RMS against the charge density for Zr in the BCC lattice as a function of the indicated mesh density. For all meshes, the RMS decreases towards their minima, with convergence to the I-WP TPMS appearing near 6.8×10^{-6}Å^{-3}. In Fig. 4.7(a,b,c,d), convergence of a surface of a given ρ in the hexagonal Zr calculated using LDA XC is shown. Similar to the HCP Zr, the surface of given ρ here converges to the H′-T surface, with the RMS minima appearing around $\rho \approx 4.47 \times 10^{-8}$Å^{-3}.

4.2.5 Nitinol

NiTi has attracted attention since its first discovery by Buehler [168] for being a shape memory material. Austenite NiTi is in the $Pm\bar{3}m$ space group and called by material scientists a BCC lattice (which actually is a simple cubic lattice). It has been discovered experimentally that its martensite phase can have $P2_1/m$, $P3$, $P\bar{3}$ $P3m1$ space groups [32,43]. All these martensite

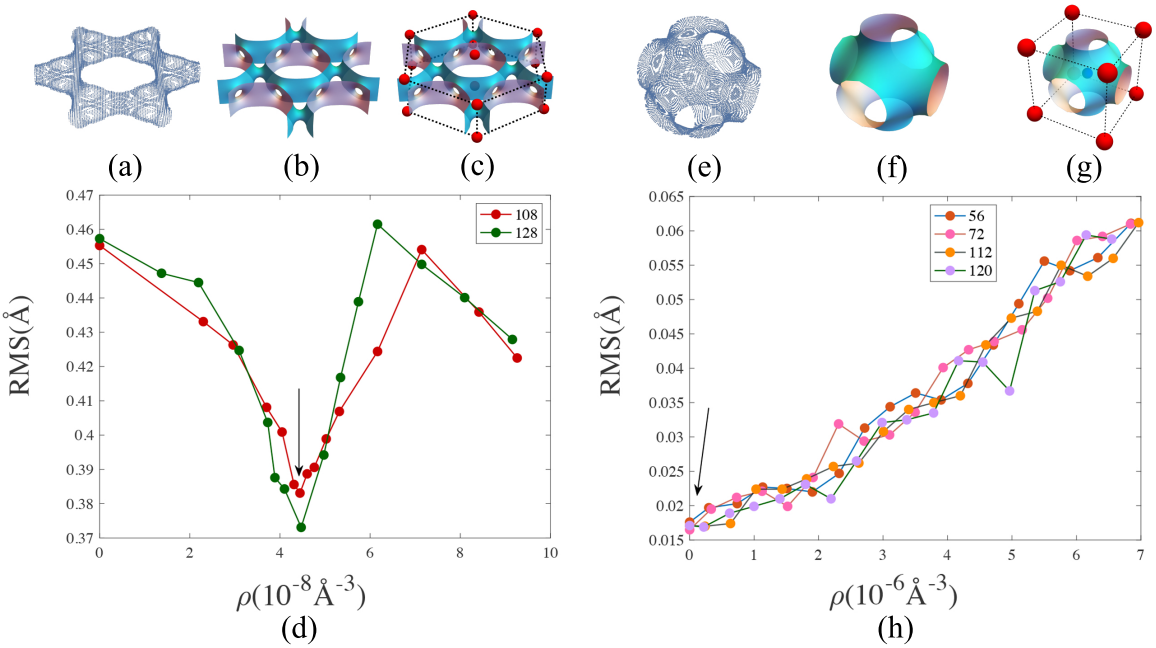


Figure 4.7 (a) Surfaces of a given charge density plotted against (b) the H′−T surface of the hexagonal Zr. (c) Atomic positions of the hexagonal Zr against the H′-T surface. (d) RMS against ρ of Zr in a hexagonal lattice as a function of mesh density given in the legend. Surface of given ρ of (e) NiTi in BCC lattice ($Pm\bar{3}m$). (f) Plot of the P surface. (g) Plot of atoms in the BCC NiTi against the P surface. Red balls and blue ball in (g) represent real Ni and Ti atoms. (h) RMS shown against charge density ρ of BCC NiTi as a function of mesh density $(n_G)^{1/3}$ given in the legend. Arrows show where the minima are in the two small panels.

phases are proposed to be able to connect to the austenite phase by a path with no energy barrier [32, 43, 169–172]. A Gamma-centred scheme is adopted to sample its reciprocal lattice, and the LDA XCF is adopted here. The RMS against the charge density for NiTi in the BCC lattice as a function of the indicated mesh density is shown in Fig. 4.7(h). The RMS converge uniformly towards zero, and it can be observed that the converged surface resembles a P surface. Figure 4.7(e) shows the converged surface, this surface against the P surface [Fig. 4.7(f)], and atom positions in the BCC NiTi in the P TPMS [Fig. 4.7(g)].

4.3 Conclusion

In this chapter, we have determined the surfaces of a given ρ by introducing exchange and correlation with the help of DFT rather than considering the electrostatic field only, which is the subject of the work done by Schnering and Nesper [49, 50]. To find the best fit to TPMS associated with various materials, *i.e.*, Na, Al, Cu, Zr and NiTi, we employed RMS and show that their surfaces of a given ρ converge to TPMS. We discovered that the point group of the crystal is a subgroup (either proper or improper) of the point group of its corresponding TPMS. As given in Table 4.3, we can see that the DFT results converge more to TPMS than ZEP (Fig 4.8). The reason why DFT results show better convergence than ZEP is that DFT calculations are closer to the real system. Although DFT can not describe real systems precisely, DFT calculations include exchange-correlations that are important in describing real systems, while ZEPs are obtained by considering only electrostatic potentials. Moreover,

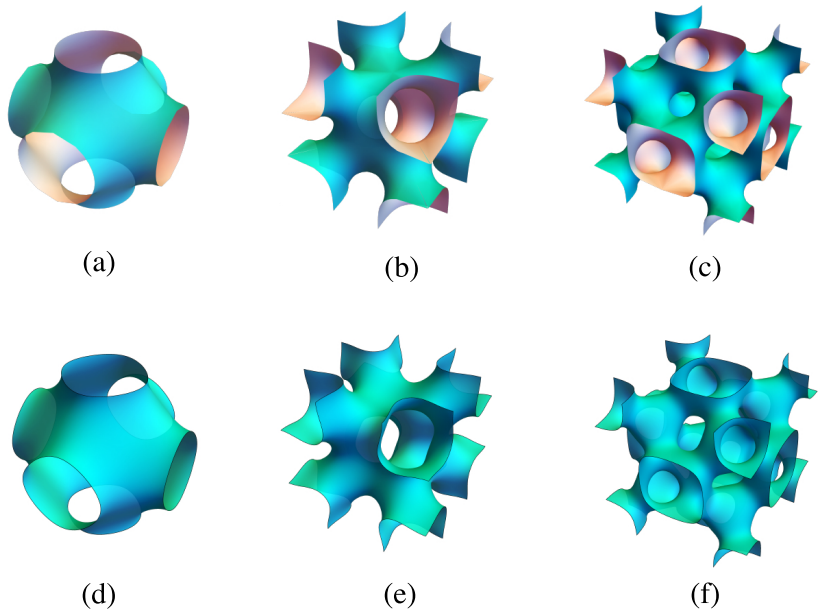


Figure 4.8 Plots of TPMS and their corresponding ZEP. From (a) to (c) are, the P surface, the I-WP surface and the F-RD surface. From (d) to (e) are, the P* surface, the IP$_2$-J* surface and the F$_{xx}$-P$_2$F$_z$ surface.

Table 4.3 Comparison of convergence of surfaces of a given charge density calculated using DFT and their corresponding ZEP to TPMS using RMS. The unit of these RMS is Å.

TPMS	RMS (DFT)	ZEP	RMS (ZEP)
P	0.0171	P^*	0.0402
I-WP	0.0568	IP_2-J^*	0.1252
F-RD	0.0820	F_{xx}-P_2F_z	0.1537

both DFT results and ZEP show a convergence when the charge density value goes towards zero. Therefore, we can infer that a surface of zero charge density value in a crystal is a TPMS. A detailed discussion is given in Sec. 6.3.3, and by then, we will see clearer why DFT results show better convergence to TPMS than ZEP. Moreover, our results show that martensitic phase transformations should have their two end phases of the same genus, while other phase transformations do not. More discussions will be presented in the next chapter. RMS data sets are given in Appendix C.2.



The moon's beyond the reach of man,
It follows us where'er it can.
Like mirror bright o'er palace wall,
When clouds disappear, it's seen by all.
We see the ancient moon,
But it has shone on men of yore.
Like flowing stream, they passed away;
They saw the moon like we do today[1].

—- Li Bai

[1] Selected from *Reflections on The Moon While Drinking*, translated by Xu Yuanchong [173].

Chapter 5

Reversibility and Continuous Deformations

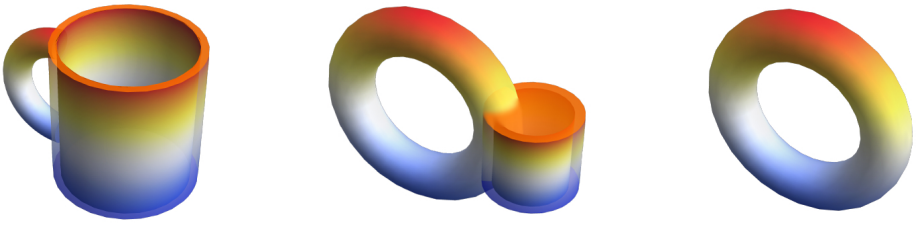
TOPOLOGY, as a subject considering relative positions between points, is the study of continuity. The word "topology" was first mentioned by Listing [174] in 1847. In 1854, Riemann introduced his ideas on manifolds in his habilitation lecture for the first time [174]. According to Riemann, points on a manifold can take transformations from one to the other either discretely or continuously, with the latter laying foundation for the development of topology. Cantor also did work crucial to the development of topology [175]. Cantor proposed point sets, a fundamental concept upon which general (or equivalently, point set) topology was built [174]. With all basic concepts prepared, in 1895, Poincaré's publication *Analysis Situs* inaugurated the emergence of a new mathematical field, topology.

Topology was first applied to physics problems of the electromagnetic field. It was discovered that the integral of the magnetic flux density around a loop in a magnetic field generated by constant currents is proportional to the linking number of two curves [174]. The formula yielded by the electromagnetic field is the same as that obtained using knot theory. After that, Poincaré solved the three-body problem from the perspective of differential geometry and derived periodic solutions to the three-body problem in a topological framework [174]. In 1931, Dirac's monopole was proposed, and one can find that the integral of its Maxwell field tensor on a unit sphere is proportional to $2\pi m, m \in \mathbb{Z}$ from the perspective of topology. In 1959, the Aharonov–Bohm effect stated that even if the magnetic field and the electric field are zero, a charged particle will still be affected by an electromagnetic potential [174, 176]. The non-triviality of the electromagnetic potential is proved to be topological. Topology was also adopted to study other physical topics, such as gravitational collapse and topological materials [174].

Though tightly connected with physical research, topology is seldom applied in studying phase transformations. In this chapter, we will study the martensitic phase transformation using results obtained in previous chapters. Recalling that our density functional theory (DFT) calculations show that for crystals undergoing a martensitic phase transformation, the end phases should be described by triply periodic minimal surfaces (TPMS) that are topologically equivalent. Now we will explain this phenomenon from the perspective of topology, then take a step further to study the shape memory effect.

5.1 A Glance at Topology and Manifold

Intuitively, *topology* is a subject about continuous deformations. Therefore, although two objects are in different shapes, they may still be "the same" from the topological perspective. For instance, a torus can *continuously*, *i.e.*, **without tearing, piecing or penetrating**, deform to a mug (Fig. 5.1), while a sphere cannot. Indeed, two objects are *topologically equivalent* if they are associated by a continuous deformation. Such a topological equivalence can be specified by a property called the *genus* that can be roughly understood as **the total number of holes on a surface**. As we can tell from Fig. 5.1, both the torus and a mug have genus one; hence, they can deform into each other continuously. Topology is designed to capture the basic structure of a space, such as the closeness, openness, neighbourhood, *etc.*. In a metric space, it is easy to define the closeness using distance. For example, a closed ball in $\mathbb{R}^3$ can be defined as $\overline{B} := \{(x,y,z) \in \mathbb{R}^3 | x^2+y^2+z^2 \leq 1\}$. However, in a non-metrisable space such that distance cannot be defined, it will be hard to define the closeness, unless using open sets[1]. Hence, with open set we can define all basic properties of a space and topology is defined using **open sets**. Mathematically, the object on which a "topology" can be defined is a set S, and a *topological space* is **a pair of the set S and its topology**, *i.e.*, the collection of all open subsets of S satisfying three axioms defining the topology of S (details are in



(a) (b) (c)

Figure 5.1 Sketch of a continuous deformation from a mug to a torus.

[1] **Open set**: an open set can be understood as a set with no boundary. For instance, an open interval $(a,b) \in \mathbb{R}$ is an open set.

Appendix D.1) [124]. Usually, we can use S to refer to the topological space for convenience.

5.2 Gauss-Bonnet Theorem

The main idea of the *Gauss-Bonnet theorem* is linking the intrinsic curvature of a surface to its topology. As introduced in Sec. 3.3.1, Gaussian curvature describes how a surface curves locally and indeed is an *intrinsic* property of a surface (mean curvature is extrinsic). This is according to *Gauss' Theorema Egregium* [177], which states that **Gaussian curvature can be described only using the first fundamental form that depends solely on the in-surface properties** (proof can be found in Ref. [47]). The word "intrinsic" means this curvature is **determined by properties defined within this surface**. Hence, an intrinsic curvature (Gaussian curvature) should be detectable to creatures living both inside and outside this surface, while an extrinsic curvature (mean curvature) is only detectable to outside observers.

To clarify this point, we will now provide an intuitive example. Let us consider a cat living in a 2D Euclidean space, *viz.*, a paper, as shown in Fig. 5.2(a), the cat can go along a straight line to reach the apple. Also, to an observer living in the 3D Euclidean space $\mathbb{E}^3$ the cat is following a straight line, too. Now, let us fold this paper to make a cylinder such that the edges AB and DC meet, with A touching D and B touching C. Then to an observer in $\mathbb{E}^3$, the cat is walking along a circle. However, to the cat itself, it is still walking along a straight line and has no idea of the circle detected by the observer. The "straight line" here is intrinsic, since it is detectable by both the cat and the observer (the observer knows that to the cat it is walking along a straight line). The "circle" here, however, is extrinsic, since it can only be



Figure 5.2 Sketch of a 2D paper in which a cat lives. The cat can walk along a straight line to the apple. Even we fold this paper to make AB and DC pairwise, to the cat itself it is still walking along a straight line.

detected by the observer.

Mathematically, being "intrinsic" means the Gaussian curvature can be obtained without referring to its *embedding* in the Euclidean space. An *embedding* is defined as a **differentiable mapping** φ between differentiable manifolds S and W ($\varphi : S \to W$) if 1) $\mathrm{d}\varphi_p : T_p(S) \to T_{\varphi(p)}(W)$ is **one-to-one** for all $p \in S$; 2) φ is a **homeomorphism** [47]. Here, $T_p(S)$ is the tangent space at p to a surface S. More details are in Sec. D.2. In our cat example, considering the cylinder case, the fact that "to the 2D cat it is 'walking along a straight line" can be detected by both the cat and the observer, hence can be sensed without knowing that the plane is folded. Recalling topology focuses on the property of a surface invariant under a continuous deformation, *i.e.*, being independent of the embedding of a surface in $\mathbb{E}^3$, now it can be understood why the Gauss-Bonnet theorem is based on the Gaussian (intrinsic) curvature.

The Gauss-Bonnet theorem gives that [47]:

$$2 - 2g_N = \frac{1}{2\pi} \int \int_S K \mathrm{d}S. \tag{5.1}$$

where g_N is the genus that can be roughly understood as the total number of holes on S[2]. The integral on the right-hand side is taken over the whole surface S, whose characteristics are:

1. **compact**: which means we can take a triangulation of this surface using finite number of triangles.

2. **connected**: which means no disjoint part can be found on this surface. A surface S is said to be connected if it cannot be written as $S = S_1 \cup S_2$, $S_1 \cap S_2 = \emptyset$ [124].

3. **closed**: such as a sphere and a torus that no boundary can be found on this compact surface. Boundaries will contribute extra terms on the right-hand side. More details can be found in Ref. [47].

The derivation of the Gauss-Bonnet theorem is given in Appendix D.7. Roughly speaking, the Gaussian curvature K is a property that can be obtained without referring to the embedding of this surface. Taking an integral of K over a closed and connected surface S, we can thus obtain a property that is independent of the embedding of S in a higher dimensional space (such as $\mathbb{E}^3$), hence, being invariant under a continuous deformation of S.

Now we can look back on triply-periodic minimal surfaces (TPMS) and consider how many genus can a TPMS have. The key here is triple-periodicity, which requires a pattern to fill a

[2]Strictly this statement requires S to be a closed surface.

3D space by repeating itself along three independent directions. Let us take a hollow cube as an example. As shown in Fig. 5.3(a), triple-periodicity requires the surface $ABCD$ to be equivalent to the surface $EFGH$, $EDAH$ to be equivalent to $FCBG$ and $EDCF$ to be equivalent to $HABG$. Similar to a torus that is formed by glueing opposite sides of a rectangle, we now glue equivalent faces of this cube. For instance, glueing $ABCD$ and $EFGH$, we will obtain Fig. 5.3(b). Repeating this process to $EDAH$ and $FCBG$ [Fig. 5.3(c)], $EDCF$ and $HABG$ [Fig. 5.3(d)], we will end up with a 3-torus [Fig. 5.3(e)], which has three genus. Hence, to ensure triple-periodicity, a TPMS should have at least three genus.

5.3 Conservation of Genus and Martensitic Phase Transformation

In the previous chapter, we discussed the convergence of the surface of a given ρ to a triply-periodic minimal surface (TPMS). By comparing their genus (Table 5.1), it can be seen that for martensitic phase transformations, *viz.*, $\alpha - \beta$ phase transformation and $\alpha - \omega$ phase transformation in Zr, and BCC−HCP phase transformation in Na, TPMS describing the two end phases have the same genus, while for other phase transformations, their genera are not required to be the same.

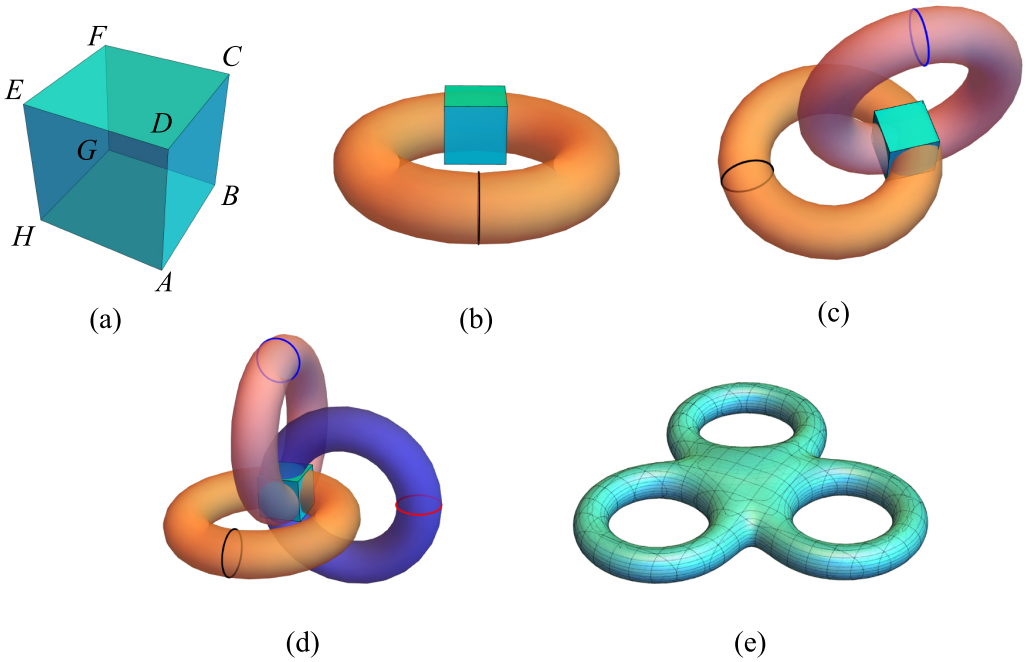


Figure 5.3 Sketch of deforming a cube to form a three-torus. (a) A cube. (b) A cube after faces $ABCD$ and $EFGH$ are connected to each other. (c) Panel (b) after $EDAH$ and $FCBG$ are connected to each other. (d) Panel (c) after $EDCF$ and $HABG$ are connected to each other. (e) A three-torus.

Table 5.1 Genus of TPMS and their corresponding materials [178, 179].

TPMS	Genus	Material	Structure
P	3	NiTi	BCC
I-WP	4	Na	BCC
		Zr	BCC (β)
		Cu	BCC
		Al	BCC
H′-T	4	Na	HCP
		Zr	HCP (α)
		Zr	Hexagonal (ω)
F-RD	6	Cu	FCC
		Al	FCC

The reason why martensitic phase transformations require a conservation of g_N is that they are "diffusionless". A diffusionless phase transformation expects small atomic displacements, which means there should be no breaking in the connection between *lattice points*. We will clarify this statement later and will introduce the definition of lattice points first. A lattice point **refers to a union of atoms such that all these unions have the same chemical and structural environment [180, 181]**. A lattice point is a mathematical point. Take the graphene as an example. As shown in Fig. 5.4(a), atom A and atom B are in different structural environments because B needs to be parallel transported along $\boldsymbol{y}$ direction by $2a$ to coincide with another atom but for A the distance is a. a is the lattice parameter. Now let us consider the union of A and B, as shown in Fig. 5.4(b), it can be seen that the distances between two adjacent unions along the same direction are the same. Thus, in graphene a lattice point represents a union of one atom A and one atom B.

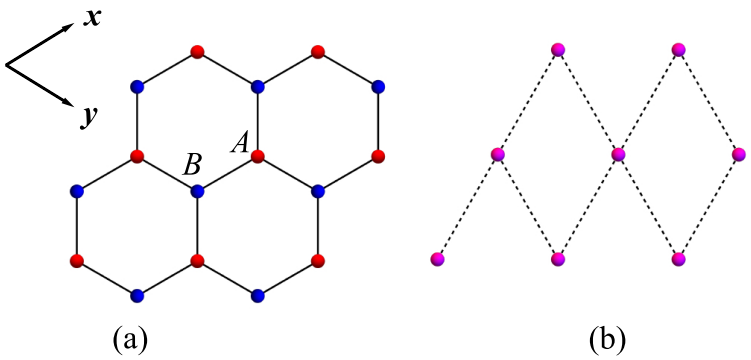


Figure 5.4 Sketch of (a) graphene and (b) its corresponding lattice formed by lattice points. Red (labelled by A) and blue (labelled by B) spheres are carbon atoms. Magenta spheres are lattice points representing the union of one atom A and one atom B.

With lattice points, a Bravais lattice can thus be defined. Therefore, for NiTi, its lattice point is a union of one Ni atom and one Ti atom in the neighbourhood of this Ni atom [Fig. 5.5(a)]. Hence, the "BCC" NiTi has a simple cubic Bravais lattice. While for Al, Cu, BCC/hexagonal Zr and BCC Na, lattice points can be chosen to be atoms or other positions that preserve the lattice structure [Fig. 5.5(b,c,d)]. For HCP Zr/Na, the lattice point is a union of one atom at the vertex and one sitting in the centre of the primitive cell [Fig. 5.5(e)]; thus, their lattice points will form a hexagonal lattice. It is noteworthy that all phase transformations we consider in this work involve no change in magnetism. Magnetism will lead to a change of symmetry and a change in the formation of lattice points. Take a BCC lattice formed by one type of atoms as an example. When we exclude the magnetism, we have its space group being $Im\bar{3}m$

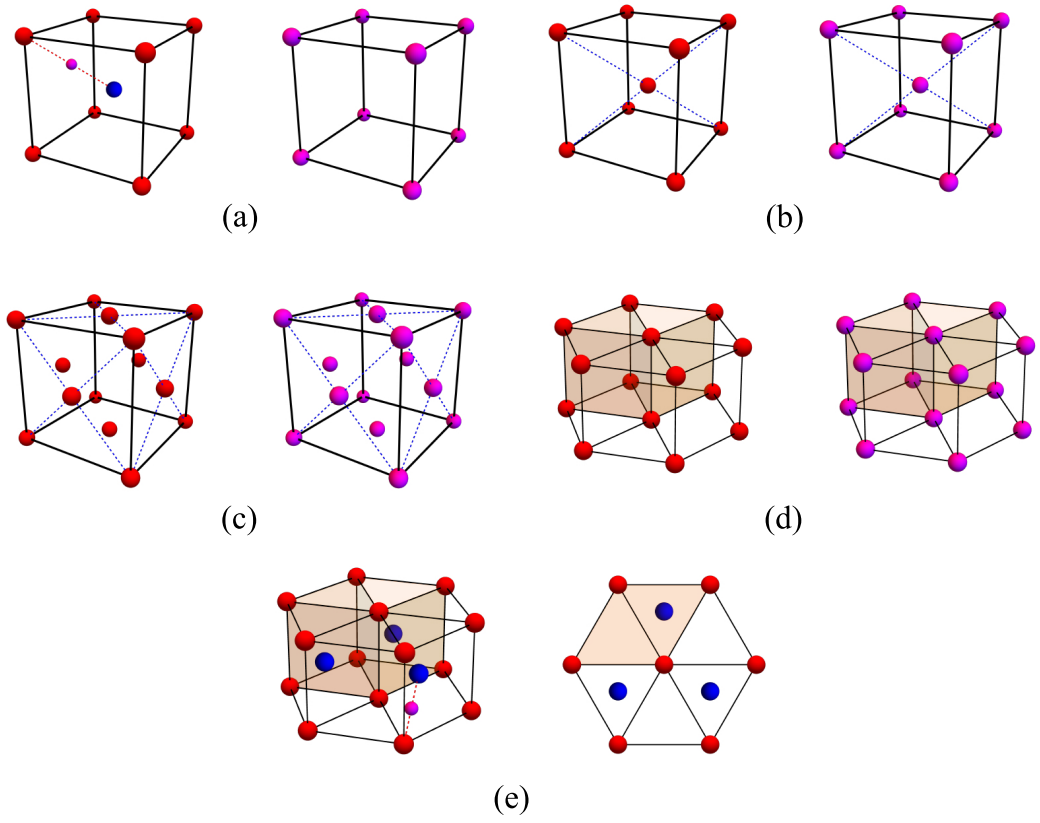


Figure 5.5 Sketch of lattices and their corresponding structures formed by lattice points. (a) NiTi lattice (left) and a simple cubic lattice formed by its lattice points (right). Red and blue spheres represent Ni and Ti atoms, respectively. Magenta spheres represent lattice points. (b) The BCC lattice of Cu, Al and Na (left) and its corresponding lattice point structure. Since, in this case, there is only one type of atom, one lattice point is formed by one atom. (d) Hexagonal Zr (left) and its corresponding lattice point structure, which is also hexagonal (right). The primitive cell is shaded orange. (e) The HCP lattice of Na or Zr. Blue spheres represent atoms sitting in the centre of the primitive cell shaded orange. On the right-hand side is a top view of this lattice.

[Fig. 5.6(a)]. However, the ferromagnetism will lead to a vanishing of the mirror reflection symmetry about the (100), (010) and (001) plane and make the space group going to $I4/mmm$ [Fig. 5.6(b,c)]. If it is anti-ferromagnetic, then we see the lattice point now should be formed by two atoms, for instance, A and B in Fig. 5.6(d).

The no-breaking in connections between lattice points during a martensitic phase transformation can be understood as follows. In diffusionless phase transformations, the relative arrangement of atoms is unchanged. We can take a simple cubic monatomic crystal (hence lattice points can be chosen as atoms) as an example. In Fig. 5.7(a), we can see that each lattice point is connected to other 6 points. If the phase transformation is not diffusionless, as shown in Fig. 5.7(b), the neighbours of a lattice point may change, hence the arms connecting this point to its neighbours. If this phase transformation is diffusionless, then the neighbours are invariant and the arms connecting one lattice point to its neighbours are unbroken [Fig. 5.7(c)]. Therefore, we see the invariance of number of arms connected to each lattice point in a martensitic phase transformation. For a crystal lattice formed by more than one type of atoms, similarly, diffusion will cause a breaking of arms connecting lattice points.

If we look at a lattice point in BCC Al/Cu, we see that it has 8 nearest neighbours [Fig. 5.8(a)] while in the FCC Al/Cu, it has 12 nearest neighbours [Fig. 5.8(b)]. And the FCC−BCC phase transformation in both Al and Cu are not martensitic. While for phase transformations like HCP−BCC in Zr, HCP−hexagonal in Zr and BCC−HCP in Na that are all martensitic, the number of neighbours connected to each lattice point is always 8 [Fig. 5.8(c)]. Now let us come back to the TPMS and form *skeletal graphs* [3]by **shrinking and expanding TPMS to obtain graphs formed by edges and vertices** (Fig. 5.9) [132, 182]. Because skeletal graphs are obtained by continuously deforming TPMS, according to the definition of topological

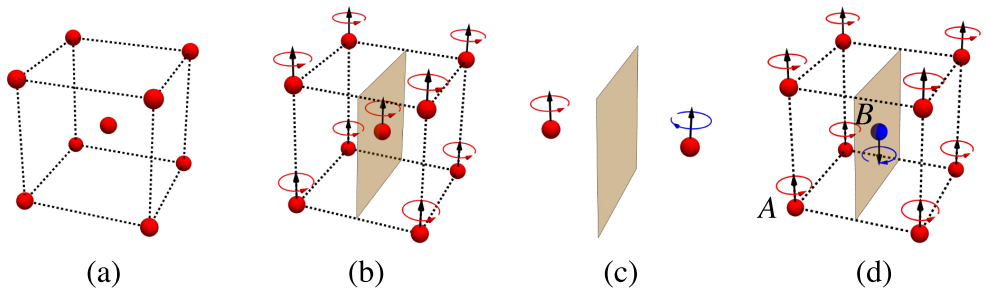


Figure 5.6 Sketch of (a) BCC iron. (b) the ferromagnetic iron in a BCC lattice. (c) Behaviour of the spin change under a mirror reflection operation. (d) Sketch of an anti-ferromagnetic iron in a BCC lattice.

[3]**Skeletal graphs**: graphs obtained by continuously shrinking and stretching a TPMS.

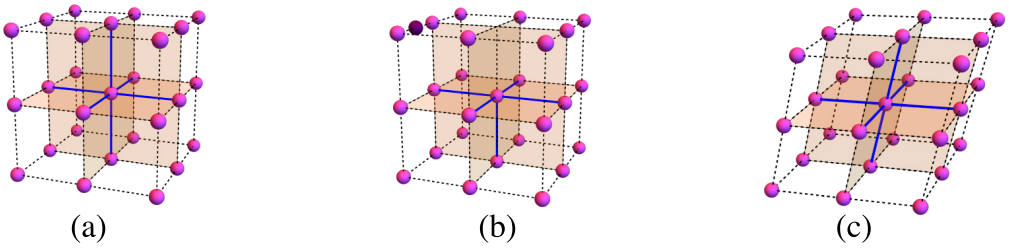


Figure 5.7 Sketch of (a) a simple cubic lattice. Magenta spheres are for lattice points and blue solid lines are for the arms that connect one lattice point to its nearest neighbours. (b) Deformed simple cubic lattice under a diffusion. The purple sphere is for the displaced lattice point. (c) Deformed simple cubic lattice under a shear that is diffusionless and can preserve the atom arrangement.

continuity (Sec. 5.1), skeletal graphs and their corresponding TPMS are of the same topology. From Fig. 5.10, it can be seen that the skeletal graphs show the connections between a lattice point and its neighbours in the corresponding crystal. Noticing that the "connection" we mention here is restricted to nearest neighbours of each lattice point. For instance, setting a lattice point in the simple cubic lattice at the origin, then neighbours connected to it are at $\pm\boldsymbol{a}_i, i = 1,2,3$, $\boldsymbol{a}_i$ are basis vectors. We will elaborate this using Al, Cu, Zr, Na and NiTi lattices in later sections. Here, we would like to make the "continuous deformation" of skeletal graphs given in Fig. 5.10 more clear.

Two skeletal graphs are said to be topologically equivalent to each other if they can deform into each other continuously without breaking or adding edges or vertices. Since the skeletal graph is obtained by continuously shrinking or extending a TPMS (Fig. 5.9), if two TPMS are topologically equivalent to each other, their corresponding skeletal graphs are topologically equivalent to each other, and *vice versa*. To be specific, take the P surface as an example. The skeletal graph of a P surface has each of its vertices (represented by balls in Fig. 5.10) connected by six arms [Fig. 5.10(a)]. Hence, the simplest case is deforming it into another

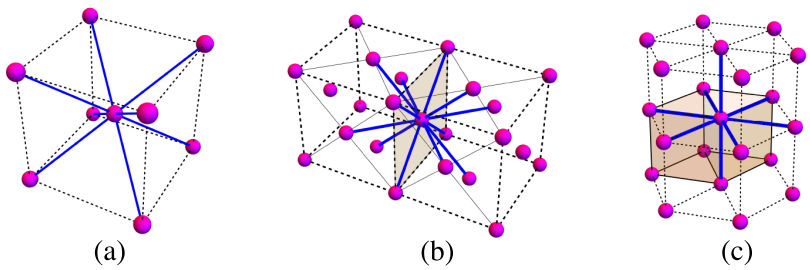


Figure 5.8 Arms connecting one lattice point to its neighbours in (a) a BCC lattice, (b) an FCC lattice and (c) a hexagonal/HCP lattice. Magenta spheres are lattice points and blue lines are the arms.

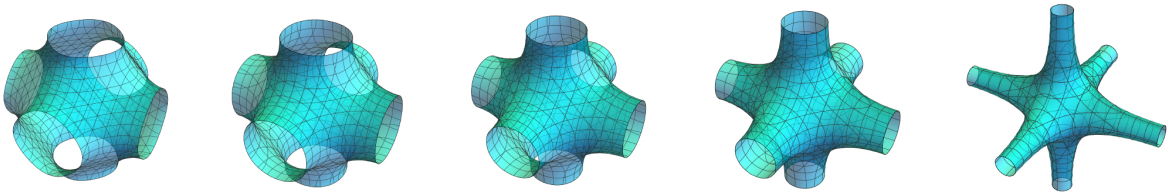


Figure 5.9 Example of obtaining a skeletal graph of a TPMS by a continuous shrinking and expanding. (a) P surface, goes from (b) to (c) to (d) then obtain the skeletal graph (e).

skeletal graph that also has six arms connected to each end (or equivalently, vertex). Then, if we stretch the end in the middle, we can either obtain Fig. 5.11(b) or Fig. 5.11(h), which corresponds to a Schwarz D surface and a Schwarz H surface, respectively. If we further stretch one end coloured yellow in Fig. 5.11(b) to make each end connected to 3 arms, we will now see the skeletal graph of a gyroid. Additionally, we see there being no way for the skeletal graph of the I-WP surface to deform into the one of the F-RD surface. Since all ends in the skeletal graph of the I-WP surface have eight arms, we cannot stretch Fig. 5.10(b) in a way that two connected ends have different numbers of arms, such as the one given in Fig. 5.11(h), to obtain the one given in Fig. 5.10(c). All operations shown in Fig. 5.11 are continuous hence do not change the genus. However, if we stretch each end in Fig. 5.10(c) with two connected ends having the same number of arms [similar to the one given in Fig. 5.10(b)], we can only obtain a graph with ends having arms less or equal to 7. Nevertheless,we see the skeletal graphs given in Fig. 5.10(b) and the surface given in Fig. 5.10(d) can deform into each other continuously. In a phase transformation, if it involves a stretching of lattice points, then we will have all lattice points stretched in the same way.

The Bain deformation, although it is physically continuous, from the perspective of topology it is not. To be specific, as shown in Fig. 5.12(a), in a FCC lattice the length of the purple line l_{purple} equals the length of the green line l_{green} given they connecting the nearest neighbours of the same lattice point. Expressing their length using the lattice parameters of the FCC

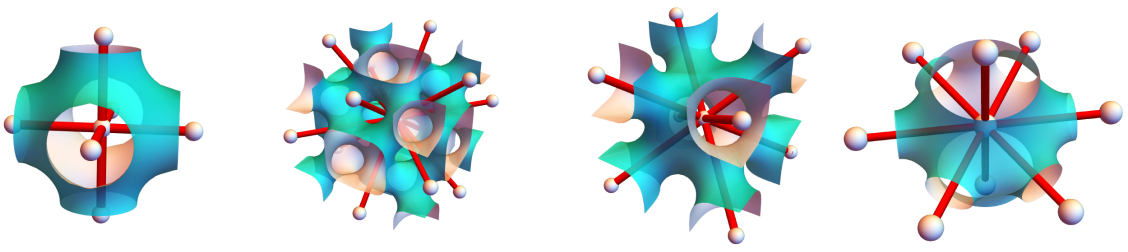


Figure 5.10 Sketch of skeletal graphs in (a) P surface, (b) I-WP surface, (c) F-RD surface, (d) H′-T surface. Each end in these graphs has 6, 8, 12 and 6 arms, respectively.

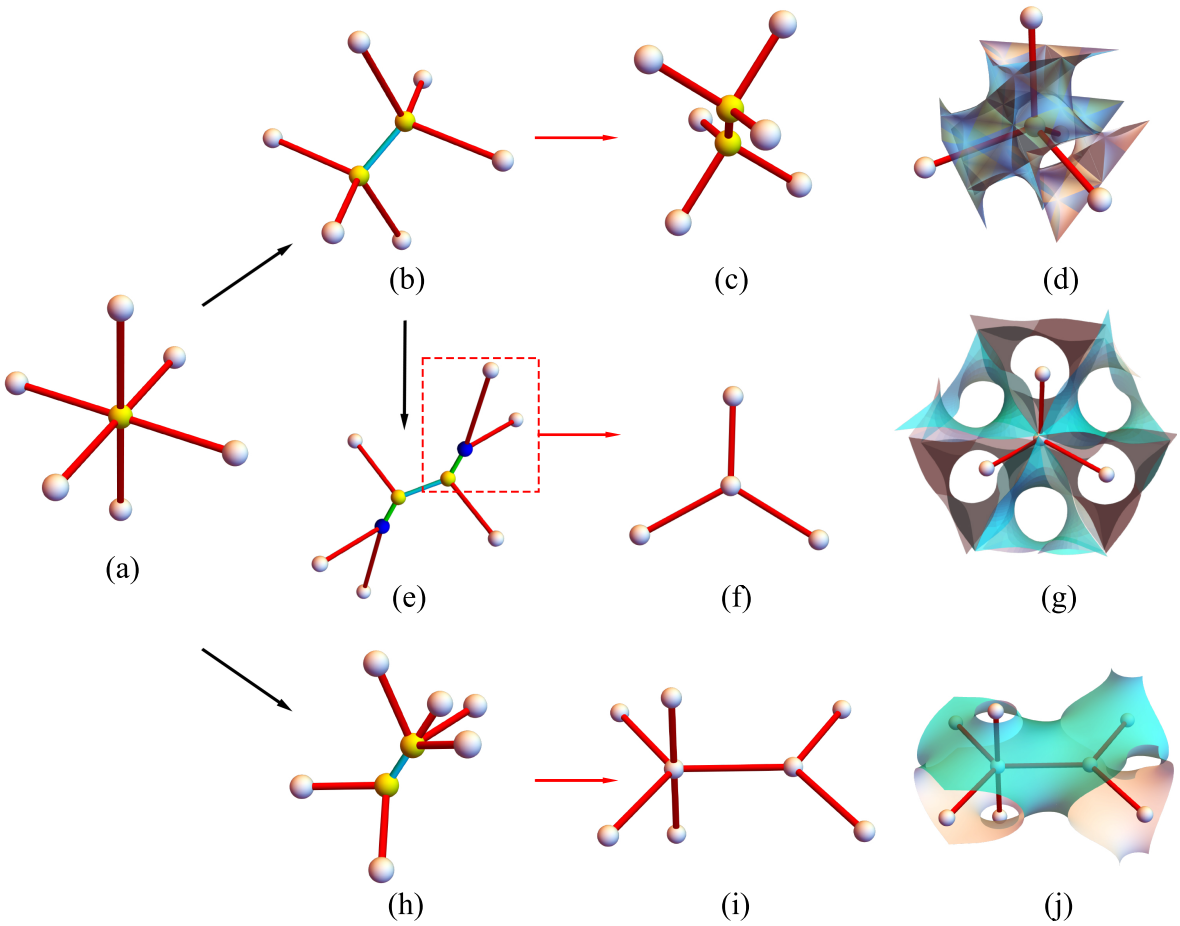


Figure 5.11 Sketch of continuous deformations of the skeletal graph of a P surface. (b, c, e, f, h, i) are deformed skeletal graphs. (d, g, j) are the Schwarz D, the gyroid and the Schwarz H surfaces, respectively.

lattice, we see:

$$l^2_{\text{green}} = \frac{a^2+b^2}{4}, \qquad l^2_{\text{purple}} = \frac{a^2+c^2}{4}. \tag{5.2}$$

Thus, when $b > c$, the centre point A will only have 8 nearest neighbours instead of 12 nearest

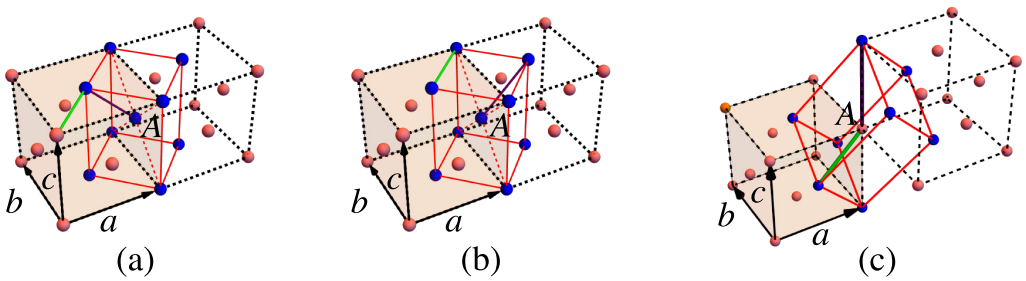


Figure 5.12 Sketch of initial states in an FCC lattice of (a) Bain deformation, (b) Nishyama–Wassermann deformation and (c) Bogers–Burgers deformation. Blue spheres label atoms that form the BCC lattice sketched by red lines. Both blue and pink spheres are atoms in the FCC lattice.

neighbours required by the skeletal graph of the F-RD surface. Hence, the sudden change of the arms connecting to one lattice point from 8 (corresponding to the I-WP surface) to 12 (corresponding to the F-RD surface) makes the Bain deformation topologically discontinuous. It is noteworthy here that links between nearest neighbours are not required to have exactly the same length. Indeed, these neighbours are chosen along the basis vectors of primitive cells. Taking a 2D lattice as an example, for the central point (coloured magenta) in Fig. 5.13(a,b), both of them have 4 nearest neighbours (coloured blue). Similar argument can be made for the Nishiyama–Wassermann deformation [Fig. 5.12(b)]. For the Bogers–Burgers deformation [Fig. 5.12(c)], we see the green and purple arms having their length expressed as:

$$l^2_{\text{green}} = \frac{a^2 + c^2}{4}, \qquad l^2_{\text{purple}} = c^2 . \tag{5.3}$$

In an FCC lattice, $l^2_{\text{purple}} > l^2_{\text{green}}$ while in a BCC lattice we have $l^2_{\text{purple}} = l^2_{\text{green}}$. Thus, when $a = \sqrt{3}c$, the green arm and the purple one will have the same length, then the topological discontinuity will happen when $a \neq \sqrt{3}c$ because the number of nearest neighbours of the central point changes. In the following sections, we will gain more insight into the arguments made above by looking at the deformation models introduced in Sec. 2.2.4 and will discuss why these models can or cannot describe a martensitic phase transformation.

5.3.1 FCC−BCC Phase Transformation

We have seen that the Bain deformation is topologically discontinuous. Looking at the skeletal graphs, since we are talking about crystals formed by only one type of atom, we can choose the lattice points to be the atoms. According to Fig. 5.14(a,b), the skeletal graph of the I-WP surface conforms to the one obtained by connecting one lattice point to its nearest neighbours in the BCC lattice. For the FCC crystal lattice, as shown in Fig. 5.14(c), we need to consider

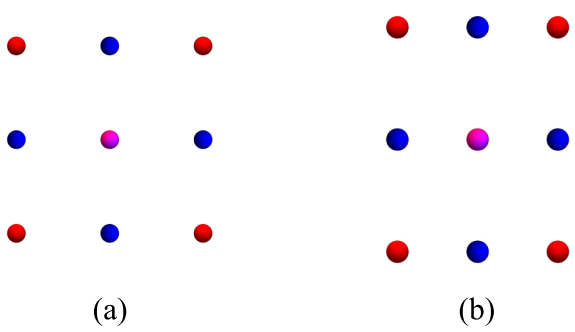


Figure 5.13 Sketch of nearest neighbours of a central lattice point. Spheres in (a) and (b) are for lattice points, blue spheres are the nearest neighbours of the central lattice point, coloured magenta.

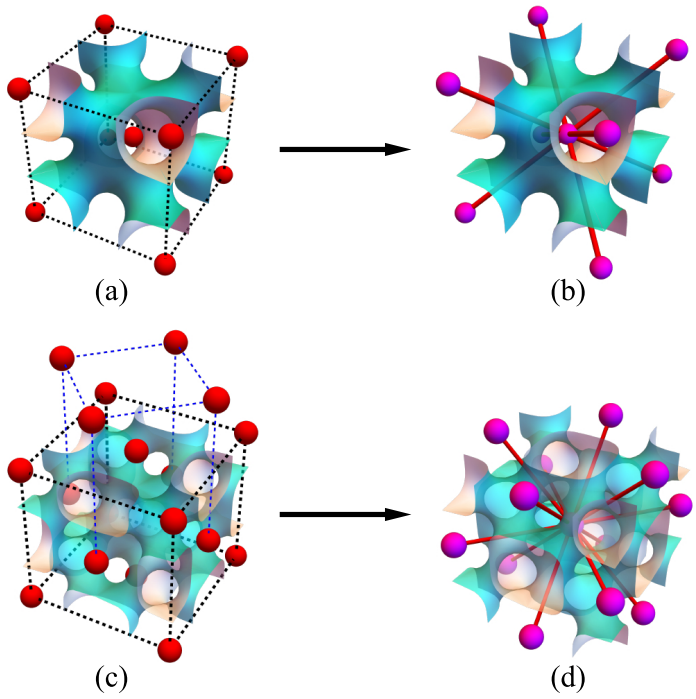


Figure 5.14 Sketch of the Bravais lattices of the BCC and the FCC crystal lattices. (a) Sketch of the BCC crystal lattice against the I-WP surface. (b) The skeletal graph of the I-WP surface that is formed by connecting lattice points. Edges are labelled by red solid lines. (c) Sketch of the FCC crystal lattice against the F-RD surface. (d) Plot of the skeletal graph of the F-RD surface against the surface. Red spheres represent atoms, and magenta spheres are lattice points.

the atoms one layer above, which are face-centred atoms here. Then, looking at the F-RD surface, including this extra layer of atoms, we see that the skeletal graph is the same as the one obtained by linking lattice points [Fig. 5.14(d)]. Hence, according to our previous discussion, these two lattices are not topologically equivalent. Therefore, the BCC−FCC phase transformation in Cu and Al are not martensitic.

5.3.2 HCP−BCC and HCP−Hexagonal Phase Transformations

Zr in the HCP lattice, as stated above, has a hexagonal lattice. Therefore, to understand the HCP−BCC phase transformation, we should look at the transformation between the hexagonal and the BCC lattices. The hexagonal−BCC phase transformation can be realised by displacing atoms along $[\overline{1}11]_{\text{BCC}}$ and $[111]_{\text{BCC}}$ by $a/6$, where a is the lattice parameter of the original BCC lattice, such that we have $[111]_{\text{BCC}} \parallel [0001]_{\text{hexagonal}}$ (Fig. 5.15). No breaking of arms connecting lattice points will happen in such a deformation. The skeletal graphs of a hexagonal and an HCP lattice are given in Fig. 5.16. The HCP lattice has each of its lattice points formed by two atoms from adjacent layers [Fig. 5.16(a)]. If we let the lattice point coincide with the atoms coloured yellow in Fig. 5.16(a), then connect one lattice point with

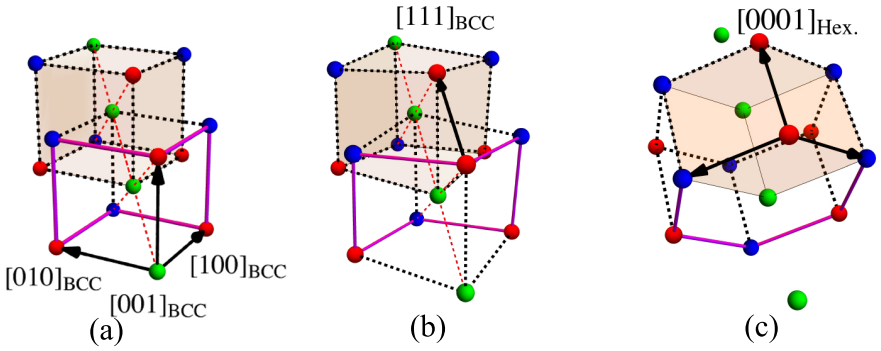


Figure 5.15 Sketch of the transformation from (a) a hexagonal lattice to (c) a BCC lattice. Atoms labelled red go along $[111]_{\mathrm{BCC}}$, blue go along $[\overline{1}\overline{1}\overline{1}]_{\mathrm{BCC}}$ and green stay at their original positions. The shaded area in (a) shows a primitive cell, and in (c) shows a unit cell. Magenta lines show how atoms at vertices of a hexagon displace during the lattice transformation.

its neighbours, we can obtain the skeletal graph of the H′-T surface. A similar argument can be made for the hexagonal lattice [Fig. 5.16(c,d)], but in this case, we can set the lattice points to be at the atoms. Referring to Table 5.1, it can be seen that both the I-WP surface (for the BCC Zr) and the H′-T surface (for the HCP or hexagonal Zr) have genus 4. Also, the skeletal graph of both the I-WP surface and the H′-T surface have eight arms connected to each end, and no change in the linking number means no change in genus. Hence, geometrically, a martensitic phase transformation is permitted. Similar arguments work for the BCC−HCP martensitic phase transformation of Na.

5.3.3 B2−B19′ Phase Transformations in NiTi

In this section, we will discuss martensitic phase transformations of NiTi. The austenite (B2) phase of NiTi has the space group $Pm\bar{3}m$. After a rapid quench, NiTi transforms to martensite B19′ ($P2_1/m$, experimental data from Ref. [183]) phase. Figure 5.17 shows the lattice deformation of B2−B19′ realised by a bilayer $\langle 100\rangle_{\mathrm{B2}}\{011\}_{\mathrm{B2}}$ shear together with a basal $\langle 100\rangle_{\mathrm{B2}}\{011\}_{\mathrm{B2}}$ shear/shuffle [184]. Before we go to the analysis using the skeletal graph, we would like to construct a more geometric description of crystal lattices using TPMS and lattice points. According to Sec. 4.2.5, we have the surface of given charge density ρ in B2 NiTi converging to the P surface. Considering the symmetry, we have the P surface formed by eight identical pieces. Such a piece is symmetric about the plane $(\overline{1}10)$ and has a 2-fold rotational symmetry about the axis $[1\overline{1}0]$ [Fig. 5.18(a)].

Connecting one Ni (labelled by A in Fig. 5.19) with the body-centre Ti (labelled by B Fig. 5.19), according to the symmetry, we have the line AB crossing at O (Fig. 5.19) sitting in the middle of one of these 8 pieces [$ABCDEF$ in Fig. 5.18(a)]. Calculating the Gaussian curvature here using Eq. E.11 and Eq. B.73, we have $K = 0$ (more details are in Appendix D.8). Hence, the

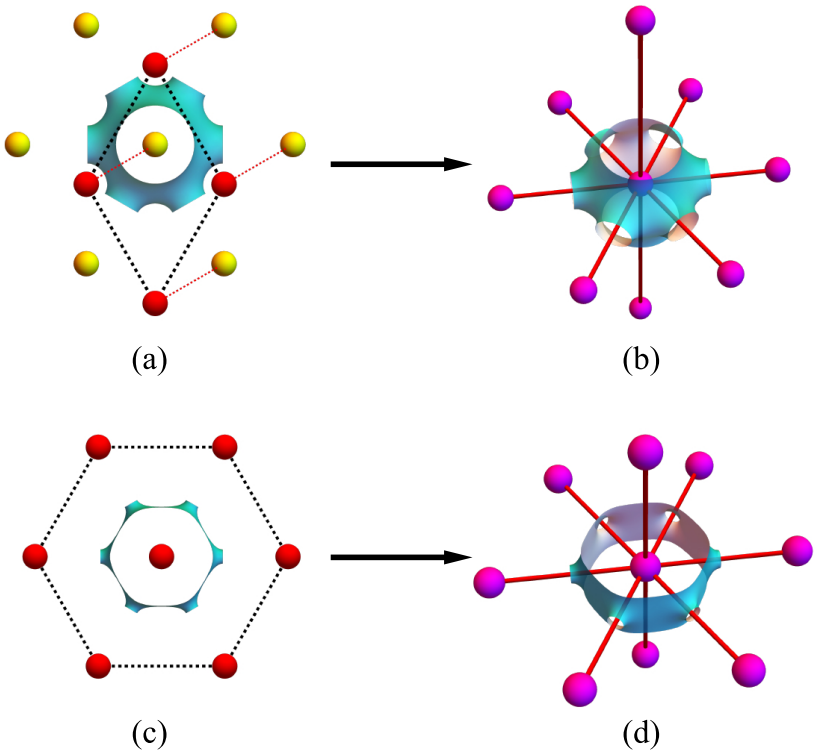


Figure 5.16 Sketch of hexagonal lattices. (a) Sketch of an HCP crystal lattice against the H′-T surface. Red and yellow spheres are for atoms on different layers. Atoms represented by yellow spheres are in the layer lying between the layers on which atoms represented by red spheres are located. Red dashed lines show the union of one atom coloured red and one atom coloured yellow that forms a lattice point. (b) Skeletal graph of the H′−T surface formed by connecting lattice points. Arms are labelled by red solid lines. (c) Sketch of a hexagonal crystal lattice against the H′−T surface. Red spheres represent atoms. (d) Skeletal graph of the H′−T surface formed by connecting lattice points. Arms are labelled by red solid lines. Magenta spheres label lattice points.

point O is a flat point. Since a lattice point, recalling its definition in Sec. 5.3, can be chosen at arbitrary positions as long as the Bravais lattice is preserved, we choose O to be a lattice point representing one Ni (labelled by A in Fig. 5.19) and one Ti (labelled by B in Fig. 5.19). Also,

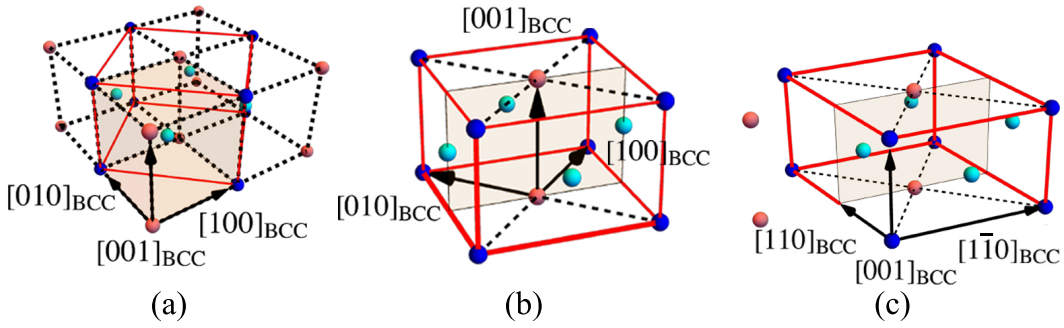


Figure 5.17 Sketch of the transformation from (a) a BCC lattice to (c) a monoclinic lattice. Ti atoms are labelled green and the rest are Ni atoms. Arrows label unit cell vectors. Shaded area in (a) shows the BCC unit cell. According to (b) and (c), we have $[110]_{\mathrm{BCC}} \| [100]_{\mathrm{B19'}}$.

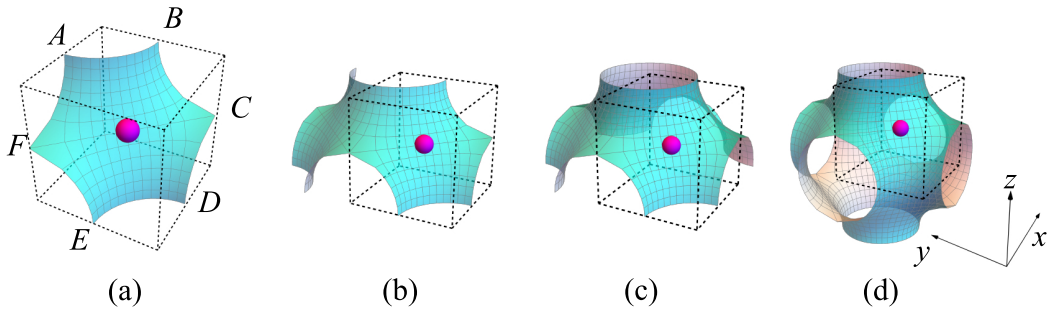


Figure 5.18 Formation of a P surface. (a) fundamental piece by taking symmetry operations on which one can obtain a P surface. The piece $ABCDEF$ is inscribed in a cube. (b) Surface obtained by taking a mirror reflection of $ABCDEF$ about the plane (010). (c) Surface obtained by taking a mirror reflection of (b) about the plane (100). (d) Surface obtained by taking a mirror reflection of (c) about the plane (001).Pink sphere represents the middle of the piece $ABCDEF$.

setting lattice points at flat points is for a mathematical convenience of describing its motion on a TPMS. Explanation to this part will be presented in Chapter 6. Similarly, we can find the rest of the lattice points [Fig. 5.20(a)] and hence draw the Bravais lattice [Fig. 5.20(b)] that is the same as the one formed only by atoms coloured red in Fig. 5.20(a). Parallel transporting the P surface and letting its centre coincide with one of the lattice points, we see the skeletal graph is now given by linking a lattice point to its neighbours [Fig. 5.20(c)]. The crystal structure of the B19′ NiTi, using data from Materials Project [185–199], is shown in Fig. 5.21(a). In B19′, one lattice point consists of two Ni atoms [yellow spheres in Fig. 5.21(a)] and two Ti atoms [orange spheres in Fig. 5.21(a)]. In Fig. 5.21(b), each lattice point is connected to 6 neighbours. However, since for B2 NiTi, one lattice point is formed by one Ni and one Ti atom, the variation in lattice point will contribute to one extra arm that "links" the two old "lattice points" [green spheres in Fig. 5.21(a)]. Therefore, the skeletal graph of B19′ should be described by the one depicted in Fig. 5.11(h), and can be deformed from the skeletal graph describing the B2 NiTi continuously.

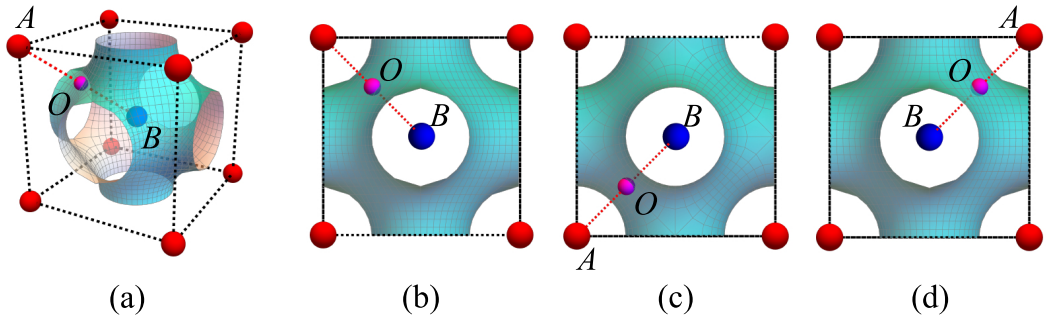


Figure 5.19 (a) Sketch of the B2 NiTi lattice. Red spheres and blue spheres represent the Ni and Ti atoms, respectively. (b,c,d) are the front, above and left orthogonal view of (a). Magenta spheres represent lattice points.

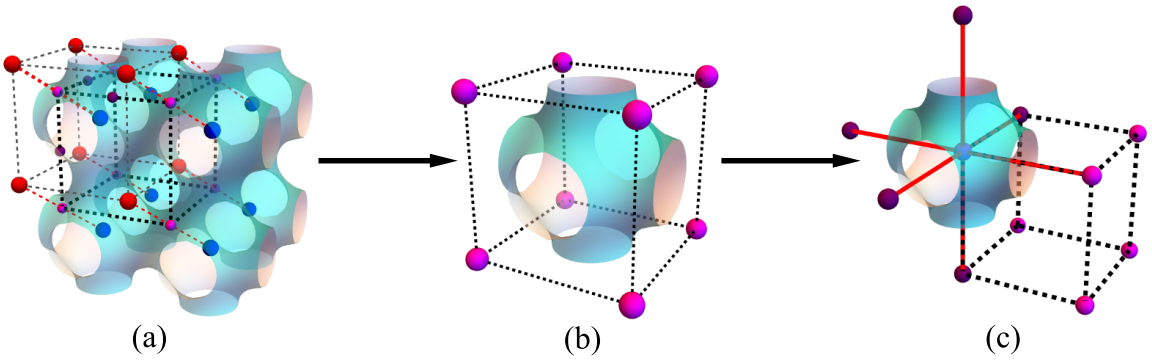


Figure 5.20 Sketch of the Bravais lattice of NiTi. (a) Sketch of the crystal lattice of B2 NiTi against the P surface. Red and Blue spheres represent Ni and Ti atoms, respectively. Pink spheres label lattice points sitting at flat points of the P surface, as given in Fig. 5.19. (b) The formed Bravais lattice by lattice points. (c) Parallel transporting the P surface to make its centre coincide with one of the lattice points. Red solid lines show the skeletal graph formed by lattice points coloured purple.

5.4 Conclusion

In this chapter, we gained more insight into the reason why martensitic phase transformations are topologically continuous. The key is in the diffusionless nature of martensitic phase transformations, *i.e.*, no atomic displacements larger than atomic spacing and no breaking in links between lattice points and their nearest neighbours. By continuously shrinking or stretching a TPMS, we can obtain its corresponding skeletal graph. Comparing the skeletal graph of a TPMS and the links between one lattice point and its nearest neighbours corresponding to a crystal, we see that they conform with each other. Hence, no discontinuous change in links between lattice points means no discontinuous change in skeletal graphs, which leads to a topologically continuous deformation between TPMS describing two end phases of a martensitic phase transformation.

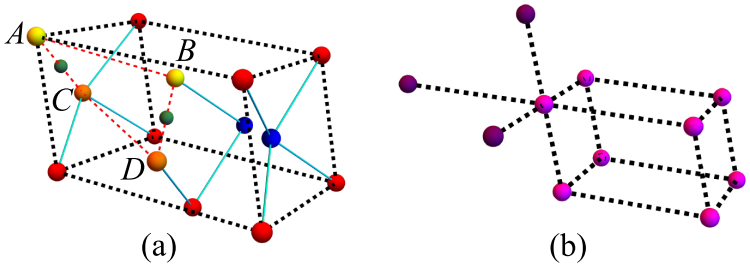


Figure 5.21 Sketch of (a) crystal lattice and (b) Bravais lattice of the B19′ NiTi. In (a), red and blue spheres are fore Ni and Ti atoms, respectively. Yellow and orange spheres are for Ni and Ti atoms forming one lattice point. Green spheres are the union of one Ni labelled by A (resp. B) and one Ti labelled by C (resp. D). In (b), purple and magenta spheres are for lattice points of the B19′ NiTi.



With swallows chirping and flying about,
All a sudden, morn is unfolding, in a vivid spring.
Flowers in gardens, blooming,
Like an iridescent scroll of silk brocade is unrolling.
Begonias, tipsy, are rouged by rains, swaying [4].

—-Song Qi

[4] Selected from *Jin Chan Dao, With swallows chirping and flying about*, translated by Yin Mengdi

Chapter 6

Convergence to Triply Periodic Minimal Surfaces in Crystals

ACCORDING to our results shown in Sec. 4.2, for crystals, we see that the shape of a surface S of a given charge density value is determined by the charge density distribution ρ. Such a surface is uniquely determined by the symmetry, the lattice parameter (which will be shown later in this chapter), as well as the connection between one lattice point and its nearest neighbours (which resembles the skeletal graph of the surface) of a crystal. Also, if we obtain a surface S of zero charge density of a crystal using computational methods, the genus of S converges to the genus of the corresponding TPMS of the crystal. Hence, although we may see TPMS having the same point group, such as the P surface and the I-WP surface [132], or even the same space group, such as the H′-T surface and the H″-R surface [132], these surfaces are distinguished from each other by their skeletal graphs or genus (or both), which means they can describe different crystals. Also, density functional theory (DFT) has pointed out that the ground state energy of a system is determined uniquely by ρ. Hence, we consider if it is possible to use a term having the dimension of energy but determined by the geometry of S to express the status of the crystal during a phase transformation. The idea is to represent the system status by two parts: 1) the shape of S and 2) the motion of lattice points on S. That is to say, during a phase transformation, due to the displacements of atoms, we have ρ, hence S, and positions of lattice points changed. We will show later in Sec. 6.3.1 that lattice points can be set at points with equal Gaussian curvature on S. Then, from the perspective of geometry, we can express the geometric energy as a sum of the bending energy of S and the kinetic and potential energy of lattice points moving on S.

Indeed, the bending energy of a surface can be expressed in a geometric way. This subject was first studied to understand the shape of a blood cell. That is to say, the energy of a lipid bilayer, which forms the outer shell of a red blood cell, can be written in terms of its

surface curvatures. According to Helfrich, the surface bending energy is a sum of two terms that are proportional to the mean curvature (M) and Gaussian curvature (K). Similarly, we describe our surface S based on Helfrich's model, with the surface energy of S being determined by the charge density distribution ρ that is determined by the crystal structure. Another property that characterises a crystal lattice is the arrangement of lattice points, which is determined by the symmetry of the crystal lattice, similar to S. Hence, we can project the lattice points of a crystal onto its corresponding TPMS, with all lattice points being geometrically equivalent. For instance, they should have the same Gaussian curvature. Moreover, the symmetry compatibility of S and the crystal lattice makes it possible for us to model the motion of lattice points during a phase transformation as being confined on S. Based on the statements above, we can build a term having the dimension of energy to learn the geometry of a crystal that is the sum of the bending energy of S and the motion of lattice points confined to S. With this term, we can study the geometric reversibility along different phase transformation paths.

We have discovered that, for martensitic phase transformations, paths with no variation in the geometric energy have zero energy barriers calculated using DFT and are allowed only by shape memory alloys. Materials discussed in Chapter 4 are employed as examples to clarify our model. We are interested in the shear-driven diffusionless progress determined by the DFT energy barrier and no attempt is made to simulate the temperature-driven nucleation progress. Thus, although the DFT calculations are conducted under 0K, we can still use them to assist our arguments.

This chapter will be arranged as follows. First, we will present the computational results to show the intuition behind building a geometric model to understand the shape memory effect. Then, we will present the mathematical background required to understand our model, followed by the modelling. Finally, we will use examples to relate the vanishing of the energy barrier in DFT calculations to the reversibility of the system's geometry.

6.1 Vanishing of Energy Barrier and Shape Memory Effect

6.1.1 Density-Functional Theory Calculations of Energy Barriers

The shape memory effect is a special case of martensitic phase transformations. To understand the mechanism that differentiates the shape memory alloys from other martensites, DFT has been employed to calculate the energy barrier along transformation paths realised by different shears. Roughly speaking, the path along which the energy barrier corresponding to a phase transformation can be obtained is searched as following. The first step is plotting the en-

ergy against varied parameters during the phase transformation (called an energy landscape). Next, the energy corresponding to the two end phases on the landscape can be located given the value of varied parameters are known for these two phases. Then, by searching for the saddle points, one can map out a path with the smallest energy barrier. One material that is widely studied is NiTi. DFT calculations [32,43,200,201] confirmed that NiTi has barrierless transformation paths. As shown in Table 6.1, for paths allowing both a bilayer and a basal

Table 6.1 Energy barriers calculated along different transformation paths. Space groups of NiTi phases listed in this table are given in Table 6.2. All shears are given in the coordinate of the B2 NiTi.

Material	Initial	Path	Final	Barrier(meV/atom)
NiTi [32]	B2	$\langle 100\rangle\{011\}^1$	B19′	∼37
	B2	$\langle 100\rangle\{011\}^2$	B19′	∼22
	B2	109°–B19′	B19′	0
NiTi [170]	B2	[111] elongation	R	0
NiTi [32,43,200,201]	B2	$\langle 100\rangle\{011\}^3$	B33	0
NiTi [172]	B2	$\langle \bar{1}10\rangle\{110\}^4$	B19	∼13
Ti-Ta [202]	BCC(β)	$[110]_\beta \| [100]^5_{\alpha''}$	$Cmcm(\alpha'')$	0
Ti-Ta [202]	BCC(β)	$[111]_\beta \| [0001]_\omega$	$P6/mmm(\omega)$	0
Ti [203]	HCP(α)	$[0001]_\alpha \| [01\bar{2}0]^6_\omega$	Hexagonal(ω)	∼9
Ti [204]	HCP(α)	$[11\bar{2}0]_\alpha \| [01\bar{1}1]^7_\omega$	Hexagonal(ω)	∼9.5
Ti [205]	HCP(α)	$[11\bar{2}0]_\alpha \| [0001]^8_\omega$	Hexagonal(ω)	58
Zr [206]	HCP(α)	$[11\bar{2}0]_\alpha \| [01\bar{1}1]^9_\omega$	Hexagonal(ω)	22
Zr [206]	HCP(α)	$[\bar{1}2\bar{1}0]_\alpha \| [0001]^{10}_\omega$	Hexagonal(ω)	58
Na [207]	BCC	Bain	FCC	∼0.79
Fe [208,209]	γ-Fe	Bain (FCT)	α-Fe	∼20
	γ–Fe	Bain	α–Fe	∼45
Fe [209]	γ–Fe	NW (FCT)	α–Fe	∼37
Fe [210]	γ–Fe	BB	α–Fe	∼25
Cu [211]	BCT(BCC)	Bain	FCC	1
Cu [212]	BCT(BCC)	Bain	FCC	1.3
Cu [213]	FCC	BB	BCC	∼700
Al [214]	BCT(BCC)	Bain	FCC	∼30
Al [214]	FCC	BB	BCC	∼100

[1] Basal (single-layer) shear.

[2] Bilayer shear.

[3] Alternate bilayer shear.

[4] Basal shear.

[5] Orientation relation, rest of which are: $[\bar{1}10]_\beta \| [010]_{\alpha''}$, $[001]_\beta \| [001]_{\alpha''}$. Various paths were generated in order to search for the lowest energy path (with the barrier provided above).

[6] Orientation relation, rest of which is: $\{11\bar{2}0\}_\alpha \| \{0001\}_\omega$. Various paths were generated in order to search for the lowest energy path (with the barrier provided above).

[7] Orientation relation, rest of which is: $(0001)_\alpha \| (0\bar{1}11)_\omega$. Lowest energy barrier has found.

[8] Orientation relation, rest of which is: $(0001)_\alpha \| (11\bar{2}0)_\omega$. [9] Orientation relation, rest of which is: $(0001)_\alpha \| (0\bar{1}11)_\omega$. [10] Orientation relation, rest of which is: $(0001)_\alpha \| (1\bar{2}10)_\omega$.

Table 6.2 Space groups and point groups of different triply periodic minimal surfaces corresponding to different NiTi phases. The column labelled 'Parameters" contains tuneable parameters whose variation produces a continuous surface deformation ensuring all associated surfaces are minimal surfaces. The corresponding Weierstrass representation can be found in Table 3.2. Group-subgroup relations between point groups can be found in Fig. 2.7.

Phase	SG	PG	Surface	PG(TPMS)	G_s	dh	Parameters
B2	$Pm\bar{3}m$	O_h	P	O_h	Eq. (3.61) [129,133]	1	a_s
B19′	$P2_1/m$	C_{2h}	oPb	C_{2h}	Eq. (3.63) [129,133]	1	a_s, b_s
B19	$Pcmm$	D_{2h}	oPa	D_{2h}	Eq. (3.61) [133]	1	a_s
B33	$Cmcm$	D_{2h}	oPa	D_{2h}	Eq. (3.61) [129,133]	1	a_s
R	$P3, P\bar{3}$	C_3, C_{3i}	oPb(H)	D_{3d}	Eq. (3.63) [129,133]	1	a_s, b_s
109°-B19′	$P2_1/m$	C_{2h}	oPb	C_{2h}	Eq. (3.63) [129,133]	1	a_s, b_s

$\langle 100\rangle_{\mathrm{B2}}\{011\}_{\mathrm{B2}}$ shear, the energy barrier is zero. While paths that allow only bilayer or single-layer (basal) $\langle 100\rangle_{\mathrm{B2}}\{011\}_{\mathrm{B2}}$ shear only have non-vanishing energy barriers. Ti and Zr, which undergo martensitic phase transformations, but are not shape memory alloys, have no barrierless deformation paths. Ordinary materials that cannot exhibit martensitic phase transformations, such as Cu and Al, also have no barrierless path. Moreover, topologically discontinuous paths, such as the Bain deformation path, have no zero energy barrier either. Iron here is calculated without considering its change in magnetism.

The DFT result is interesting, given that martensite (Ti and Zr), non-martensite (Cu and Al), and topologically discontinuous paths have no barrierless path, while only shape memory alloys (Ti-Ta and NiTi) can have barrierless paths. In the previous chapter, we discussed that what distinguishes martensitic phase transformations from ordinary transformations is the existence of geometric continuity in triply periodic minimal surfaces (TPMS) corresponding to the end phases. Here, we want to take a step further to find the geometric explanation for the vanishing of energy barriers along deformation paths shown only by shape memory alloys. The main idea is to treat the TPMS as a manifold and lattice points as massive particles moving on this manifold. Before we elaborate on this idea, we will provide more mathematical and physical background related to our modelling.

6.2 Helfrich Hamiltonian

The concept of *"Helfrich Hamiltonian"* was first proposed to study the elastic bending energy of a lipid bilayer in 1973 [55], and now is widely applied in studying the elastic energy of a membrane [7,215,216]. Helfrich took the analogy of the elastic energy of liquid crystals and then deduced that the elastic energy per area of a membrane can be expressed in terms of the mean curvature M and the Gaussian curvature K.

Bending will lead to variations in local curvatures. According to our definition of the first and second fundamental forms (Sec. 3.3), we see that a change in the normal vector field of this surface can reveal a change in local curvatures. Hence, Helfrich assumed that the bending energy of a membrane E_{bend} is a function of derivatives of its normal vector field. He chose a set of Cartesian coordinates (x, y, z) with the origin set to be a point on this surface. Then, the basis is $\{\boldsymbol{e}_i\}$, $i = 1, 2, 3$, with $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ lying in the tangent plane to this surface at the origin and $\boldsymbol{e}_3$ being the unit normal vector $\boldsymbol{n}$ at this point. $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are arbitrary orthonormal directions. We can express $\boldsymbol{n}$ as:

$$\boldsymbol{n} = (n_x, n_y, \sqrt{1 - n_x^2 - n_y^2}), \tag{6.1}$$

where n_x and n_y are x and y components of $\boldsymbol{n}$. Since $\boldsymbol{n}$ is perpendicular to a given surface, the curl of $\boldsymbol{n}$ should be zero, which means:

$$\partial_y n_x - \partial_x n_y = 0\,. \tag{6.2}$$

Because we are interested in the elastic bending energy, to define it, we can refer to the elastic energy of a spring system:

$$\varepsilon = \frac{1}{2} k \Delta x^2\,, \tag{6.3}$$

where Δx is the variation in length and k is the spring constant. Also, we see that the curvature can be expressed by the variation in $\boldsymbol{n}$ (Appendix D.3). Thus, we expect the bending energy at each point on a surface to be proportional to the quadratic form of the variation in $\boldsymbol{n}$:

$$\begin{pmatrix} \partial_x \boldsymbol{n} \\ \partial_y \boldsymbol{n} \end{pmatrix} = \begin{pmatrix} \partial_x n_x & \partial_x n_y \\ \partial_y n_x & \partial_y n_y \end{pmatrix} \begin{pmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \end{pmatrix}. \tag{6.4}$$

Here, n_z is eliminated because we set $n_z = \sqrt{1 - n_x^2 - n_y^2}$. Moreover, energy should be a scalar that is rotationally invariant. Hence, we want to have the quadratic forms written using matrix in Eq. 6.4 to be invariant under transformations:

$$\mathsf{P}^{-1} \nabla \boldsymbol{n} \mathsf{P}, \qquad \mathsf{P}^{-1} = \mathsf{P}^{\mathsf{T}}, \qquad \det(\mathsf{P}) = \pm 1\,, \tag{6.5}$$

where P is an orthonormal matrix describing a rotational operation. Thus, the possible terms that are invariant under such an operation are the trace and the determinant of $\nabla \boldsymbol{n}$:

$$\begin{cases} \partial_x n_x + \partial_y n_y, \\ (\partial_x n_x \partial_y n_y - \partial_x n_y \partial_y n_x). \end{cases} \tag{6.6}$$

Since we need quadratic forms, we can expect E_{bend} to take the form:

$$E_{\text{bend}} = \frac{1}{2}a_1(\partial_x n_x + \partial_y n_y - b_1)^2 + a_2(\partial_x n_x \partial_y n_y - \partial_x n_y \partial_y n_x), \tag{6.7}$$

where a_1 and a_2 are constants having the dimension of energy, and b_1 is a spontaneous curvature that will occur when the two layers of the lipid membrane are in different chemical environments [1]. Hence, b_1 will vanish when both sides of this membrane are chemically identical. Using Eq. D.3 given in Appendix D.3, we have:

$$\partial_x n_x + \partial_y n_y = M, \qquad \partial_x n_x \partial_y n_y - \partial_x n_y \partial_y n_x = K \tag{6.8}$$

Hence, we see that E_{bend} can be expressed as:

$$E_{\text{bend}} = \frac{1}{2}a_1(M - b_1)^2 + a_2 K \tag{6.9}$$

Then the Helfrich Hamiltonian of the system is expressed as:

$$H = \int_S E_{\text{bend}} \mathrm{d}S, \tag{6.10}$$

where the integral is taken over the membrane.

6.3 Geometric Description of System Energy

6.3.1 Revision of Lattice Points

To build our model, let us begin with the choice of lattice points. In the previous chapter, we set lattice points of NiTi as flat points on a TPMS as shown in Fig. 5.19(a) by magenta balls. Now, we would like to generalise this statement to other materials. Take BCC Al as an example, whose atoms sit at the centre of the I-WP surface [Fig. 6.1(a)]. Then, projecting one atom along the vertical direction onto the I-WP surface, we see the images of these atoms sitting right at flat points (calculations are given in Appendix D.8.2), as shown in Fig. 6.1(b). By taking the projection of all atoms, we will obtain a set of flat points on the I-WP surface. This set of flat points preserves the Bravais lattice structure; also, given that BCC Al consists of only one type of atom, we can choose this set of flat points to be lattice points of BCC Al. Similar arguments can be made for FCC, HCP and hexagonal lattices mentioned before (Fig. 6.1).

Then we see that, for a phase transformation, the end phases have:

[1] For instance, one lipid layer is facing water on one side while another lipid layer is facing air on the other side.

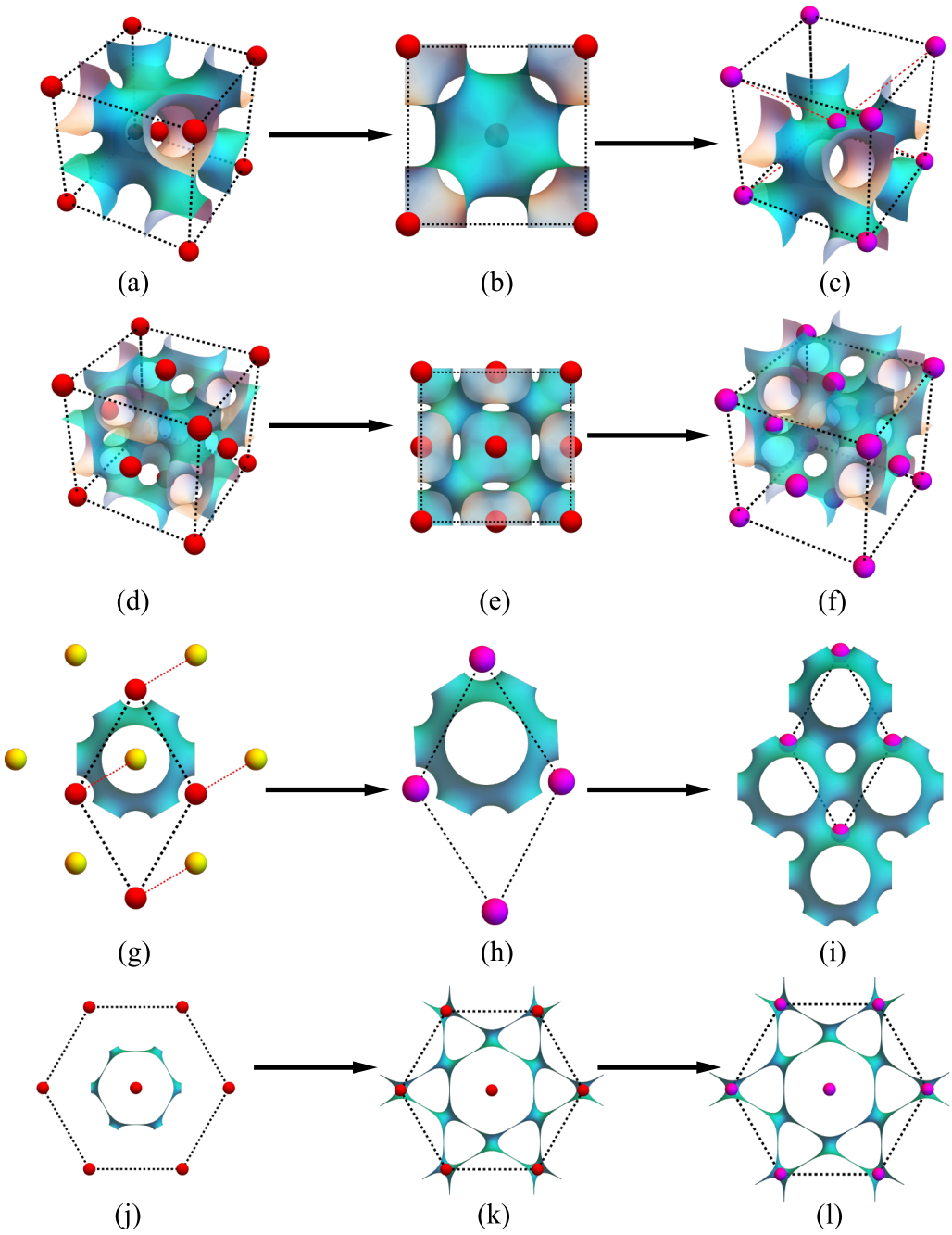


Figure 6.1 Sketch of crystal lattices and the choice of lattice points. Red and yellow spheres signify atoms and magenta spheres signify lattice points. (a) BCC lattice against the I-WP surface. (b) Above orthogonal view of (a). (c) BCC lattice formed by lattice points. All lattice points are at flat points of the I-WP surface. (d) FCC lattice against the F-RD surface. (e) Above orthogonal view of (d). Similar to the BCC lattice, all atoms of the FCC lattice can be projected to flat points on F-RD lattice with the corresponding Bravais lattice preserved. (g) Above orthogonal view of an HCP lattice. Yellow spheres label atoms lying on the layer in between. Red dashed lines show one lattice point is formed by a union of two atoms lying on different layers. (h) Lattice points locating at atoms labelled by red spheres in (g) against the H′-T surface. (i) Lattice points projected onto flat points on the H′-T surface. (j) Sketch of the hexagonal lattice against the H′-T surface. (k) Sketch of the hexagonal lattice against the H′-T surface that rotates by $\pi/6$. (l) Lattice points obtained by projecting atoms onto the rotated H′-T surface. The Bravais lattice is preserved.

1. Their surfaces S of a given charge density value ρ converging to TPMS, such that the point group of the crystal is a subgroup of the one of their corresponding TPMS;

2. Their lattice points being able to be projected to points having the same local curvature on S with their Bravais lattice structures being preserved, *i.e.*, lattice parameters and Bravais lattices are unchanged;

3. Their corresponding TPMS having the same genus, as long as the phase transformation is martensitic.

Moreover, since lattice points inherent the symmetry property of the crystal, which is the same as the TPMS, we can always find a TPMS on which all lattice points can reside. Information of the flat points are given in Appendix D.8.

6.3.2 Geometric Description of Ground State

In this section, we would like to explain why a surface S of given charge density ρ in crystals converges to a TPMS from the geometric perspective. The main idea is to use curvature to describe the geometric energy of the system. In DFT, the system energy is a functional of ρ. Since we see the geometry of S is also determined by ρ, we expect an expression of a "geometric energy" that can show the geometric status of the system and will minimise when the system is at its ground state. To do this, we will study the geometric energy from two aspects: 1) the bending energy of S and 2) the energy of lattice points located on this surface.

First, we would like to discuss the surface energy of $S \subset \mathbb{R}^3$ in crystals. Noticing that we are interested in a phase transformation rather than a thermal expansion and contraction, neither an enlargement or a downsizing of the surface will be considered. We will focus only on the evolution of a surface that can be reflected by the variation in its normal vector field. Therefore, we can employ the Helfrich Hamiltonian here to describe the surface energy. Since S is single-layered[2], we have the spontaneous curvature vanishing, which gives:

$$E_b = \frac{1}{2}\int_S \tau_1(\rho) M^2 \mathrm{d}S + \int_S \tau_2(\rho) K \mathrm{d}S = \frac{\tau_1(\rho)}{2}\int_S M^2 \mathrm{d}S + \tau_2(\rho) 2\pi\chi, \tag{6.11}$$

where $\tau_1 > 0$ and $\tau_2 < 0$ are functions of ρ. τ_1 and τ_2 depend on the types and positions of atoms forming this crystal and have the dimension of energy. Hence, when ρ is set to be a constant, so τ_1 and τ_2 can be pulled out of the integral. χ is the Euler characteristic of this surface, satisfying $2-2g_N = \chi$, where g_N is the genus of this surface. Here, $\tau_1 > 0$ and $\tau_2 < 0$ are chosen to make the energy non-negative. Because we see that for a TPMS having

[2]**Single-layered**: means unlike a lipid bilayer that is formed by two layers of membranes with each layer facing different chemical environments on either side.

at least genus 3, $\chi < 0$.

According to Eq. 6.11, we see why the surface of a given ρ converges to a TPMS. When the system is at its end phase, *i.e.*, no phase transformation takes place, then the term $\tau_2 2\pi\chi$ is determined. Hence, the minimisation of system bending energy will rely on the M^2 term. Since $\tau_1 \int_S M^2 \mathrm{d}S \geq 0$, the equal sign is taken if and only if S is a minimal surface. Hence, to minimise the surface energy, we require S to be a minimal surface.

The next step is looking at the energy of lattice points confined on a surface S. In general, we can obtain the system energy by calculating the Hamiltonian using the system Lagrangian (use Einstein summation convention) [217]:

$$L = \frac{1}{2}\dot{x}^i \dot{x}_i - V(\boldsymbol{x}), \qquad i = 1, 2, 3 \tag{6.12}$$

where x^i are Cartesian coordinates, $\boldsymbol{x} = (x_1, x_2, x_3)$ and $V(\boldsymbol{x})$ is the potential energy. However, the confinement of lattice points moving on a manifold requires an additional term. Hence, it is the *generalised Hamiltonian* [218] to be considered (details are in Appendix E.2). In our $\mathbb{R}^3$ space, let the surface S be described by an implicit function:

$$F(x_1, x_2, x_3) = F(\boldsymbol{x}) = q_0, \tag{6.13}$$

where x_1, x_2, x_3 are coordinates in Cartesian coordinate system and $q_0 \to 0$ is a constant. The system Lagrangian under this constraint is:

$$L = \frac{1}{2}\dot{x}^i \dot{x}_i - V(\boldsymbol{x}) + \lambda F(\boldsymbol{x}), \tag{6.14}$$

where λ is a function of coordinates $\{x^i\}$ and the time t. The primary constraint thus is (see Appendix E.2):

$$P_\lambda = \frac{\partial L}{\partial \dot{\lambda}} \approx 0. \tag{6.15}$$

Hence, the total Hamiltonian should be:

$$H' = \frac{1}{2}P^i P_i + V(\boldsymbol{x}) - \lambda F(\boldsymbol{x}) + \eta P_\lambda, \tag{6.16}$$

where η is the velocity function. Hence, according to the consistency condition (see Appendix E.2), we have:

$$\begin{cases} \phi_1 = P_\lambda \approx 0, \\ \phi_2 = \dot{P}_\lambda = \{P_\lambda, H'\}_P = F(\boldsymbol{x}) \approx 0, \\ \phi_3 = \ddot{P}_\lambda = \{F(x), H'\}_P = P^i \partial_i F(\boldsymbol{x}) \approx 0, \\ \phi_4 = \dddot{P}_\lambda = \{P^i \partial_i F(\boldsymbol{x}), H'\}_P = P^i P_j \partial_i \partial^j F(\boldsymbol{x}) + [\partial^k V(\boldsymbol{x}) - \lambda \partial^k F(\boldsymbol{x})] \partial_k F(\boldsymbol{x}) \approx 0, \end{cases} \tag{6.17}$$

where $\{,\}_P$ is the *Poisson bracket*. Given $\dddot{P}_\lambda$ contains a term $\eta \partial^b F(\boldsymbol{x}) \partial_b F(\boldsymbol{x})$, we only need four constraints as listed above. Simplifying $\partial_i F(\boldsymbol{x})$ by n_i, we now give the commutation relations using the Poission bracket $M_{ij} = \{\phi_i, \phi_j\}_P$ (Appendix E.2) as below:

$$\{M_{ij}\} = \begin{pmatrix} 0 & 0 & 0 & c_{14} \\ 0 & 0 & -c_{14} & c_{24} \\ 0 & c_{14} & 0 & c_{34} \\ -c_{14} & -c_{24} & -c_{34} & 0 \end{pmatrix}, \qquad \{M_{ij}\}^{-1} = \begin{pmatrix} 0 & \frac{c_{34}}{c_{14}^2} & -\frac{c_{24}}{c_{14}^2} & -\frac{1}{c_{14}} \\ -\frac{c_{34}}{c_{14}^2} & 0 & \frac{1}{c_{14}} & 0 \\ \frac{c_{24}}{c_{14}^2} & -\frac{1}{c_{14}} & 0 & 0 \\ \frac{1}{c_{14}} & 0 & 0 & 0 \end{pmatrix} \tag{6.18}$$

where:

$$\begin{aligned} c_{14} &= \{\phi_1, \phi_4\}_P = -n^i n_i, \\ c_{24} &= \{\phi_2, \phi_4\}_P = -P^i \partial_i c_{14}, \\ c_{34} &= \{\phi_3, \phi_4\}_P = 2P^k P_j \partial_i n_k \partial^j n_i - (\partial_i \partial_k n_j) n_i P^k P_j - n_j n_i \partial_i \partial^j V(\boldsymbol{x}) \\ &\quad - (\partial_i V(\boldsymbol{x}))(\partial_k n^i) n^k + 2\lambda n_j n_i \partial_i n_j - (\partial_i \partial_k n_j) n_i P^k P_j. \end{aligned} \tag{6.19}$$

Hence, we can see the geometric structure of our system is:

$$\begin{aligned} \{P_i, P_j\}_D &= \{P_i, P_j\}_P + \sum_{m,n} \{P_i, \phi_m\}_P M^{-1}_{mn} \{\phi_n, P_j\}_P, \\ &= n_i P^k \partial_j n_k - n_j P^k \partial_i n_k, \\ &= P^k (n_i \partial_j n_k - n_j \partial_i n_k), \end{aligned} \tag{6.20}$$

where $\{,\}_D$ is the *Dirac bracket* (see Appendix E.2). The first line of Eq. 6.20 gives the definition of the Dirac bracket. Now, taking a correspondence between the classic expression given above and a quantum mechanic expression. We set mass to 1 for simplicity because 1) all lattice points are of the same mass and 2) it is the geometry that will determine the final result here, and we will focus on the geometry in following calculations. Now, we see [217]:

$$P^i \to -i\hbar \partial^i, \qquad \{,\}_D \to [,], \tag{6.21}$$

and:

$$[P_i, P_j] = i\hbar\{P^k, (n_i\partial_j n_k - n_j\partial_i n_k)\}_S, \tag{6.22}$$

where:

$$\{A, B\}_S = \frac{1}{2}(AB - BA). \tag{6.23}$$

Hence, we have (derivation can be found in Appendix E.3.2):

$$[P_i, P_j] = \frac{1}{2}(-i\hbar\partial^k)i\hbar(\partial_i F\partial_j\partial_k F - \partial_j F\partial_i\partial_k F) = \frac{i\hbar}{2}(P^i n^k\partial_j n_k - P^j n^k\partial_i n_k). \tag{6.24}$$

Recalling that our constrain $F(\boldsymbol{x}) = q_0$, $q_0 \to 0$, q_0 is along a direction normal to the surface S, S is determined by $F(\boldsymbol{x}) = 0$, we use following relations here:

$$n_i = \partial_i F(\boldsymbol{x}) = \partial_i q^0, \qquad n^i n_i = 1. \tag{6.25}$$

Since we want to see how the geometry of S influences the lattice point energy, we now will do a coordinate transformation to write the system Hamiltonian in the *Gauss normal coordinates*. The *Gauss normal coordinates* are defined as follows. For all $p \in S$, the Gauss normal coordinates at p is a **coordinate system setting p to be the origin, taking the unit normal vector at p to S and two orthonormal vectors in the tangent plane at p to S as basis vectors.**

Now, in Gauss normal coordinates, we set the point on S that the studied lattice point initially locating at as the origin, and we have three orthonormal basis vectors set to be $\{\boldsymbol{g}_0, \boldsymbol{g}_1, \boldsymbol{g}_2\}$, with $\boldsymbol{g}_0 \parallel \nabla F(\boldsymbol{x})$ and $\boldsymbol{g}_1, \boldsymbol{g}_2 \in T_p(S)$.

Hence, the metric g (first fundamental form) is expressed as:

$$g = \begin{pmatrix} 1 & 0 & 0 \\ 0 & g_{11} & g_{12} \\ 0 & g_{21} & g_{22} \end{pmatrix}. \tag{6.26}$$

Let the coordinates in the Gauss normal coordinates be $\{q_0, q_1, q_2\}$, according to Eq. E.7, we have:

$$g_{\mu\nu} = \frac{\partial x^i}{\partial q^\mu}\frac{\partial x^i}{\partial q^\nu}, \qquad g^{\mu\nu} = \frac{\partial q^\mu}{\partial x^i}\frac{\partial q^\nu}{\partial x^i}, \qquad i = 1,2,3, \qquad \mu,\nu = 0,1,2. \tag{6.27}$$

Since under a coordinate transformation we have (see Appendix E.3.3):

$$\{p_\mu, p_\nu\}_D = p^\sigma(n_\mu\partial_\nu n_\sigma - n_\nu\partial_\mu n_\sigma). \tag{6.28}$$

Hence, following a similar procedure given in Eq. 6.22, Appendix E.3.3 and Appendix E.3.2, we see the correspondence:

$$[p_\mu, p_\nu] = \frac{\hbar^2}{2}(\partial_\mu \Gamma^\sigma{}_{\nu\sigma} - \partial_\nu \Gamma^\sigma{}_{\mu\sigma}). \tag{6.29}$$

Then the energy given by the Hamiltonian by taking $P_\lambda = 0, F(\boldsymbol{x}) = 0$ yields:

$$H = \frac{1}{2} g_{\mu\nu} p^\mu p^\nu + \widetilde{V}|_S(q) = \frac{1}{2} g_{ab} p^a p^b + \frac{1}{2} g_{00} p^0 p^0 + \widetilde{V}|_S(q), \qquad a, b = 1, 2, \tag{6.30}$$

where $\tilde{V}|_S(q)$ is the potential energy expressed in the Gauss normal coordinates confined on the surface S, and:

$$g_{00} p^0 p^0 = \frac{\hbar^2}{4} \Gamma^\mu_{\mu 0} \Gamma^\nu_{\nu 0}, \tag{6.31}$$

where:

$$\Gamma^\mu_{\mu 0} = \frac{1}{2} g^{\mu\sigma}(\partial_{q^0} g_{\sigma\mu} + \partial_{q^\mu} g_{\sigma 0} - \partial_{q^\sigma} g_{\mu 0}) = \frac{1}{2} g^{ab} \partial_{q^0} g_{ab} = -2M. \tag{6.32}$$

Here, M is the mean curvature and $\partial_{q^0} g_{ab} = -2h_{ab}$, $\{h_{ab}\}$ is the second fundamental form of the surface S in the Gauss normal coordinate (Appendix E.3.1). Thus, the energy is given by:

$$E_{\text{lat}} = \frac{1}{2} g_{ab} p^a p^b + \frac{\hbar^2}{2} M^2 + \widetilde{V}|_S(q) \tag{6.33}$$

Since in later calculations we are only interested in the variation in E_{lat}, and the potential energy on S is determined by ρ, which is zero through out the phase transformation, we can ignore the potential energy term. Now we obtain the total energy of lattice points:

$$E_{\text{lat,tot}} = \sum_{m=1}^{N_\Omega} \left(\frac{1}{2} g_{ab,m} p^a_m p^b_m + \frac{\hbar^2}{2} M^2_m \right), \tag{6.34}$$

where the summation runs over all lattice points, whose number is N_Ω, labelled by m in one unit cell. As required by the definition of lattice points, they should be equivalent to each other, hence the geometric energy per unit cell can be expressed as:

$$E_g = N_\Omega E_{\text{lat}} + E_b. \tag{6.35}$$

Hence, we see that, for a minimisation of E_g, we expect:

1. $M = 0$, which means the surface S should be a minimal surface;
2. $\frac{1}{2} g_{ab} p^a p^b = 0$, which means the lattice point should be relatively static with respect to the surface S.

6.3.3 Minimal Surface and Charge Density

Previously, we have discussed the system energy by taking a crystal unit cell as a combination of its lattice points and a surface of a given ρ. In this section, we will gain more insight into why when ρ goes towards zero, the corresponding surface S approaches a TPMS (Fig. 6.2).

As shown in Fig. 6.3(a), S can be treated as an interface between two regions A and B. The value of ρ on S, so does the total charge enclosed in A, are influenced by the current $\boldsymbol{j}$ flowing in the crystal. $\boldsymbol{j}$ consists of two parts: 1) the drift current $\boldsymbol{j}_{\text{drift}}$ and 2) the diffusion current $\boldsymbol{j}_{\text{diff}}$, such that:

$$\boldsymbol{j}_{\text{drift}} = q\mu \boldsymbol{E}\rho, \qquad \boldsymbol{j}_{\text{diff}} = -qD_c\nabla\rho, \tag{6.36}$$

where $\boldsymbol{E}$ is the electric field, q and μ are constants describing the unit charge and the mobility of electron, respectively. D_c is the diffusion constant. The variation in the total charge enclosed in A should equal the current "flowing " out of S [Fig. 6.3(b,c)], which means:

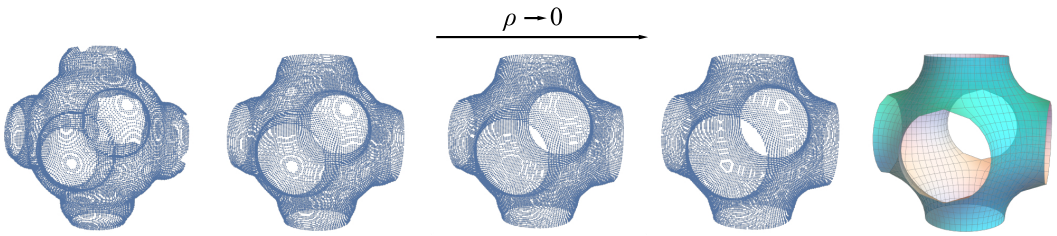
$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega} q\rho \mathrm{d}\Omega = -\int_{S} \boldsymbol{j}\cdot\boldsymbol{n}\mathrm{d}S, \tag{6.37}$$

The variation in $\int_{\Omega}\rho\mathrm{d}\Omega$ should go to zero, otherwise S will move. Thus, we have:

$$q\frac{\mathrm{d}}{\mathrm{d}t}\int_{A}\rho\mathrm{d}\Omega = 0, \tag{6.38}$$

which requires j_n, the normal component of $\boldsymbol{j}$, to be zero all over the surface. Locally, we have the continuity equation given in Eq. 6.37 expressed as:

$$\frac{\partial\rho}{\partial t} = \nabla\cdot\boldsymbol{j}. \tag{6.39}$$



(a) (b) (c) (d) (e)

Figure 6.2 Sketch of a surface of a given charge density ρ evolving along (a), (b), (c) and (d) when ρ goes towards zero. It can be seen that when ρ goes towards zero, the surface converges towards the P surface given in (e).

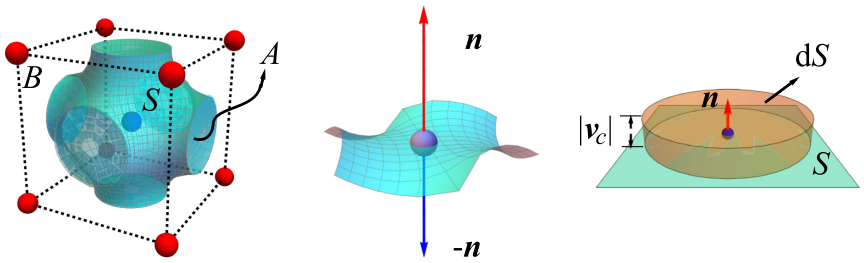


Figure 6.3 (a) Region within a NiTi unit cell divided by the P surface. Within the P surface and the area on the boundary of the unit cell coloured cyan is the region A, the rest of the unit cell is the region B. (b) Sketch of normal vectors pointing "outward" (red) and "inward" at a point labelled by a purple sphere on the surface. (c) Sketch showing the amount of charge "flowing" out of a surface S within an area dS. $|\boldsymbol{v}_c|$ is the flowing rate and $\boldsymbol{n}$ is the direction of dS (normal vector).

Hence, the local invariance of ρ, *viz.*, the invariance of ρ at every point on S, requires:

$$\nabla \cdot \boldsymbol{j} = 0. \tag{6.40}$$

Now let us look at an arbitrary point $p \in S$. As shown in Fig. 6.4(a) that sketches a small piece on S, let us chose a local coordinate with two orthonormal axis $\boldsymbol{x}$ and $\boldsymbol{y}$ lying in the tangent plane $T_p(S)$ at p to S, and the third axis $\boldsymbol{n}$, which is the unit normal axis at p to S. $\boldsymbol{x}$ and $\boldsymbol{y}$ are not required to be along principal directions.

Then because on S, ρ is given, we have at any point $p \in S$:

$$\partial_x \rho = \partial_y \rho = 0. \tag{6.41}$$

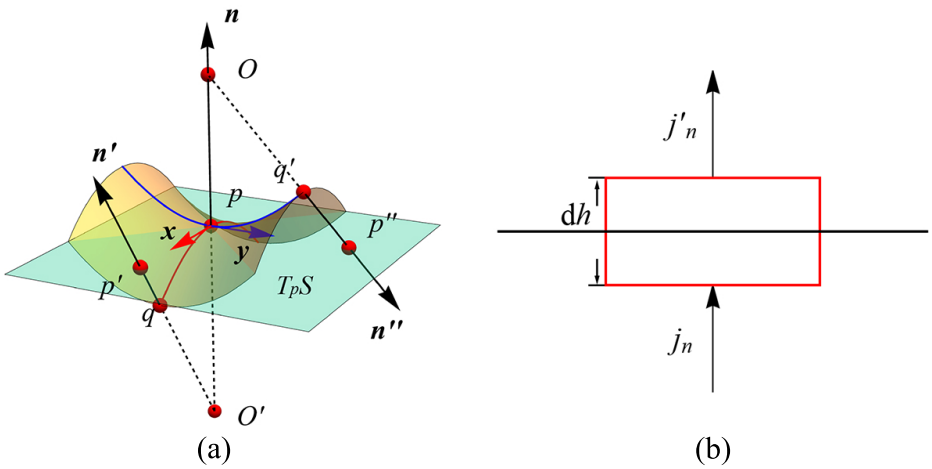


Figure 6.4 Sketch of (a) a small piece on S and (b) a thin pill box chosen enclose a small piece of S.

Now let us consider a pill box [Fig. 6.4(b)] with $\mathrm{d}h \to 0$ enclosing a small area of S. Because $\mathrm{d}h$ is very small, *viz.*, the cross section is very small, the variation of the total charge enclosed in this pill box contributed by the in-plane current is negligible. Hence, to make the charge contained in this pill box invariant, we require j_n flowing into this box to balance j_n' flowing out of this box [Fig. 6.4(b)]. Thus, $j_n = j_n'$, and because $\mathrm{d}h \to 0$, we have the variation of j_n along $\boldsymbol{n}$ goes to zero, which means:

$$\partial_n j_n = 0. \tag{6.42}$$

When $\rho = 0$, according to Eq. 6.36 and the requirement that $j_n = 0$, the drift current, hence the normal component of the diffusion current, goes to zero. Therefore, we will see $\partial_n \rho = 0$ when $\rho = 0$. However, we want to study the geometric property of S by looking at the variation in ρ on S. Thus, we need to go to the vicinity of the $\rho = 0$ surface, where the drift current is small but non-zero, and $\partial_n \rho \neq 0$.

Next, let us look at the $\boldsymbol{xpn}$ plane, as shown in Fig. 6.5(a). The point p' can be obtained by displacing $p \in S$ along $\boldsymbol{x}$ by δx or by displacing $q \in S$ along $\boldsymbol{n}'$ by $\delta n'$. δx and $\delta n'$ are very small. $\boldsymbol{n}'$ is the unit normal vector at q to S. Geometrically, we have:

$$\delta x^2 + R_1^2 = (\delta n' + R_1)^2, \tag{6.43}$$

where R_1 is the radius at p of curve C_1 (the red curve in Fig. 6.4(a)]). Because $\delta x \to 0$, under a Taylor expansion around $\delta x = 0$, we have:

$$\delta n' \approx \frac{\delta x^2}{2R_1}. \tag{6.44}$$

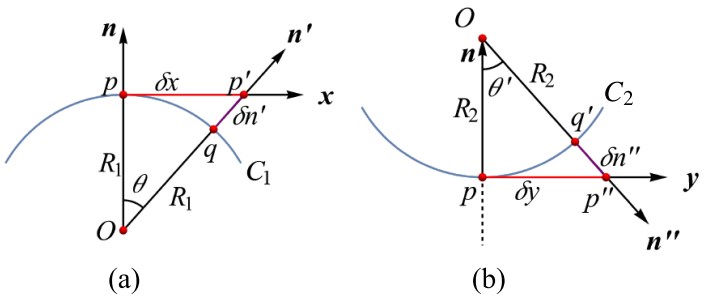


Figure 6.5 (a) Sketch of viewing Fig. 6.4(a) from the $\boldsymbol{xpn}$ plane. p and q are very close to each other. The radius at p, so does q, of the curve C_1 is R_1. (b) Sketch of Sketch of viewing Fig. 6.4(a) from the $\boldsymbol{ypn}$ plane. p and q' are very close to each other. The radius at p, so does q', of the curve C_2 is R_2.

Also, under Taylor expansions around $\delta x = 0$ and $\delta n' = 0$, respectively, we have:

$$\begin{aligned}\rho(p') &= \rho(p) + \partial_x\rho|_p \delta x + \frac{1}{2}\partial_{xx}\rho|_p \delta x^2 + O(\delta x^3), \\ &= \rho(q) + \partial_{n'}\rho|_q \delta n' + O(\delta x^4),\end{aligned} \tag{6.45}$$

Ignoring the higher order terms because they are too small, with $\rho(p) = \rho(q)$ and $\partial_x\rho|_p = 0$, we have:

$$\partial_{xx}\rho|_p = \frac{\partial_{n'}\rho|_q}{R_1}. \tag{6.46}$$

Similarly, in the yon plane [Fig. 6.5(b)], we have:

$$\partial_{yy}\rho|_p = \frac{\partial_{n''}\rho|_{q'}}{R_2}, \tag{6.47}$$

where R_2 is the radius at p of curve C_2 (the blue curve in Fig. 6.4(a)]), as shown in Fig. 6.5(b). Because the surface piece is very small and q (q') approaches to p, we have:

$$\partial_n\rho|_p \approx \partial_{n'}\rho_q \approx \partial_{n''}\rho|_{q'}, \tag{6.48}$$

which means:

$$\partial_{xx}\rho|_p + \partial_{yy}\rho|_p = \left(\frac{1}{R_1} + \frac{1}{R_2}\right)\partial_n\rho|_p. \tag{6.49}$$

Moreover, in the vicinity of $\rho = 0$, the variation of charge is contributed principally by the diffusion current. Take the x component of the drift current as an example. According to Eq. 6.39, we have:

$$\partial_x j_{\text{drift},x} = q\mu\partial_x(E_x\rho) = q\mu\rho\partial_x E_x + q\mu E_x\partial_x\rho. \tag{6.50}$$

Because $\rho \to 0$, the term $q\mu\rho\partial_x E_x$ goes towards zero. And because $\partial_x\rho = 0$ at every point on S, we see that $\partial_x j_{\text{drift},x} \approx 0$. Similar argument holds for the y component of the drift current. Hence, given $\partial_n j_n = 0$, Eq. 6.39 is re-written as:

$$\partial_x j_{\text{diff},x} + \partial_y j_{\text{diff},y} = 0, \tag{6.51}$$

at every point on S. According to Eq. 6.36, we see that:

$$\partial_{xx}\rho|_p + \partial_{yy}\rho|_p = \left(\frac{1}{R_1} + \frac{1}{R_2}\right)\partial_n\rho|_p = 0. \tag{6.52}$$

Because $\partial_n \rho|_p \neq 0$, using Euler formula (Eq. 3.32) and the definition of mean curvature M, we see that:

$$\frac{1}{R_1} + \frac{1}{R_2} = 2M = 0. \tag{6.53}$$

Thus, we have that when $\rho \to 0$, the stability of its corresponding surface of a given charge density value requires it to converge to a TPMS. If this surface is not a minimal surface, then we will witness a variation in ρ on this surface. And now we can answer why DFT results give a better convergence to TPMS than ZEP, which is because in real systems surfaces of zero charge density value are TPMS and DFT settings are closer to real systems.

6.4 Geometric Explanation of the Shape Memory Effect

6.4.1 Model Construction, Geometric Entropy

Based on our energy modelling, we now would like to discuss the condition for a shape memory effect to occur. First, we review the results we have obtained using a flow chart, as shown in Fig. 6.6. According to DFT calculations, we have at end states, the surface of zero charge density S conforming to a TPMS that is compatible with the symmetry of the crystal. When the Euler characteristic χ (equivalently, the genus) of the TPMS describing the two end phases are the same, then the phase transformation is martensitic; otherwise, it is not. To study the phase transformation geometrically, we propose a "geometric energy" based on the fact that S is determined by the charge density distribution ρ. This energy consists of two parts: 1) the bending energy of the surface S and 2) the energy of lattice points moving on S. The bending energy depends on the mean curvature M and Gaussian curvature K. The energy of lattice points depends on the gliding of lattice points on S, the geometry of S and the potential energy of lattice points in the crystal.

Hence, we see the first condition for a shape memory effect to take place is χ, the Euler characteristic, being invariant. In this section, we are going to discuss other conditions that can ensure the occurrence of a shape memory effect using the geometric energy obtained above. The main idea is based on the "reversibility" of a shape memory alloy. First, we look at the energy variation during a martensitic phase transformation. Noticing that the phase transformation is realised by a rapid quenching, we can approximate the temperature during the phase transformation by the lower temperature at which the phase transformation completes. According to the results obtained above, we see that:

$$E_g = N_\Omega E_{\text{lat}} + E_b. \tag{6.54}$$

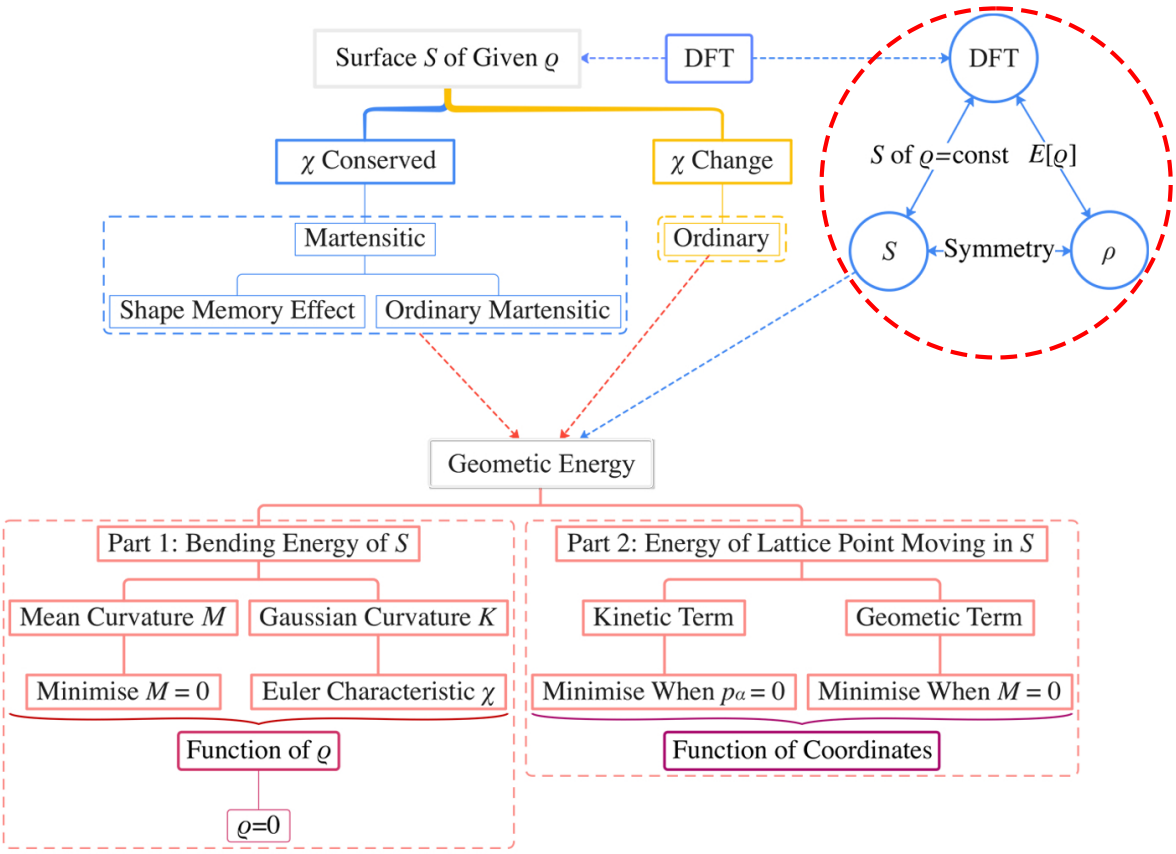


Figure 6.6 Conclusion of results obtained in previous sections for the minimal surface modelling. By calculating surfaces of a given charge density in various crystals without change in magnetism using DFT, we find that crystals undergoing a martensitic phase transformation have the end phases topologically equivalent, *i.e.*, having the same Euler characteristic. Based on the geometric equivalence required by a martensitic phase transformation, we propose a geometric energy that is a function of ρ. This energy is formed by two parts: 1) the bending energy of S and 2) the energy of lattice points moving in S.

where:

$$E_b = \frac{1}{2}\int_S \tau_1(\rho)M^2 \mathrm{d}S + \int_S \tau_2(\rho)K\mathrm{d}S, \tag{6.55}$$

and:

$$E_{\text{lat}} = \frac{1}{2}g_{ab}p^a p^b + \frac{\hbar^2}{2}M^2. \tag{6.56}$$

Taking an analogy with the thermodynamic entropy, to study the reversibility of a phase transformation, we introduce the geometric entropy here, which is given as:

$$T\Delta S_g = \Delta E_g, \tag{6.57}$$

where T is the phase transformation temperature and can be taken as a constant since the martensitic phase transformation is rapid. Moreover, the temperature is to balance the dimension of both sides of the equation, and the focus here is the variation in E_g. Thus, we do

not need to consider the variation of T here. The vanishing of ΔS_g means $\Delta E_g = 0$. Referring back to Eq. 6.55 and Eq. 6.56, let us parameterise this phase transformation by a parameter t. When $t = 0$, we are at the initial phase and when $t = t_f$, we are at the final phase. We divide the interval $[0, t_f]$ into a set of small intervals:

$$[0, t_f] = \cup_{i=1}^{n} [t_{i-1}, t_i], \qquad 0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t_f, \tag{6.58}$$

with:

$$t_i - t_{i-1} = \mathrm{d}t \to 0, \qquad \Delta E_g = \sum_i [|\mathrm{d}E_g(t_i)| + |\mathrm{d}E_{\mathrm{lat}}(t_i)|], \tag{6.59}$$

where:

$$\mathrm{d}E_b(t_i) = E_b(t_i) - E_b(t_{i-1}), \qquad \mathrm{d}E_{\mathrm{lat}}(t_i) = E_{\mathrm{lat}}(t_i) - E_{\mathrm{lat}}(t_{i-1}). \tag{6.60}$$

The reason why the absolute values are taken in Eq. 6.59 is it is not just at end points, but the variation in E_g at each time in $[t, t_f]$ to be considered. Looking at the interval $[t_{i-1}, t_i] \subset [t, t_f]$, let us consider the general case, *i.e.*, non-martensitic phase transformation first. In this case, the phase transformation is not topologically continuous. Also, because the charge density is set to be zero on S, the parameter $\tau_2(\rho)$ can be pulled out of the integral. Hence, using the Gauss-Bonnet theorem, we can see that there exists at least one interval $[t_{i-1}, t_i]$, such that:

$$\int_{S'} \tau_2(\rho', t_i) K' \mathrm{d}S' - \int_S \tau_2(\rho, t_{i-1}) K \mathrm{d}S = 2\pi \tau_2(\rho)(\chi' - \chi) \neq 0. \tag{6.61}$$

Therefore, a non-martensitic phase transformation has $\Delta E_g \neq 0$. Then, let us consider a general martensitic phase transformation. In this case, the variation in genus vanishes. Hence, we can focus on the first term that depends on M^2. Now we have:

$$\mathrm{d}E_b(t_i) = \int_{S'} \frac{\tau_1(\rho', t_i)}{2} M'^2 \mathrm{d}S' - \int_S \frac{\tau_1(\rho, t_{i-1})}{2} M^2 \mathrm{d}S. \tag{6.62}$$

A general continuous deformation is not required to stay in a space of TPMS. Hence, there can be an interval $[t_{i-1}, t_i]$ such that the surface S at t_{i-1} is a minimal surface while the one S' at t_i is not. Hence, the first term in Eq. 6.62 will vanish (given S is a minimal surface) while the second term will not. Therefore, we have this interval gives $|\mathrm{d}E_b(t_i)| \neq 0$, which means $\Delta E_g \neq 0$.

Now, if the martensitic phase transformation path is restricted to a space of TPMS, we have $\sum_i |\mathrm{d}E_b(t_i)| = 0$. Then, let us look at the energy of lattice points E_{lat}. We see that for a continuous deformation in a space of TPMS:

$$E_{\mathrm{lat}}(t_i) = N_\Omega(t_i) \left[\frac{1}{2} g_{ab,t_i} p^a_{t_i} p^b_{t_i} + \frac{\hbar^2}{2} M'^2 |_{M'=0} \right]. \tag{6.63}$$

Since $S^{'}$ and S are minimal surfaces, we see the term $\hbar^2 M^2/2 = \hbar^2 M^{'2}/2 = 0$. $N_\Omega(t_\mathrm{i})$ (resp. $N_\Omega(t_{\mathrm{i}-1})$) is the number of lattice points per unit cell, and we should have $N_\Omega(t_\mathrm{i}) = N_\Omega(t_{i-1})$ if our deformation is continuous (Sec. 5.3). If we expect a martensitic phase transformation that is geometrically reversible, which means we should have $\Delta S_g = 0$, we need $\mathrm{d}E_{\mathrm{lat}}(t_i) = 0, \forall t_i \in [t, t_f]$, such that:

$$g_{ab,t_\mathrm{i}} p^a_{t_\mathrm{i}} p^b_{t_\mathrm{i}} = g_{ab,t_{\mathrm{i}-1}} p^a_{t_{\mathrm{i}-1}} p^b_{t_{\mathrm{i}-1}}. \tag{6.64}$$

Let $t_\mathrm{i} = t_0$, since at the initial state $g_{ab} p^a p^b = 0$, we now see that $\mathrm{d}E_{\mathrm{lat}}(t_\mathrm{i}) = 0$ requires $g_{ab,t_\mathrm{i}} p^a_{t_\mathrm{i}} p^b_{t_\mathrm{i}} = 0$, for all $t_\mathrm{i} \in [0, t_f]$. It means the images of lattice points on a TPMS should exhibit no gliding on the TPMS. To be specific, let us consider a deformation between two TPMS S and S', such that:

$$S = \boldsymbol{x}(u,v),\ (u,v) \in \mathbb{R}^2, \qquad S' = \boldsymbol{x}'(u,v),\ (u,v) \in \mathbb{R}^2. \tag{6.65}$$

A continuous deformation between S and S' means there exists an invertible mapping φ, such that $\varphi \circ \boldsymbol{x} = \boldsymbol{x}'$. Then under φ, $p \in S, p = \boldsymbol{x}(u_0, v_0)$ will be mapped to:

$$\varphi(p) = \varphi \circ \boldsymbol{x}|_{(u_0,v_0)} = \boldsymbol{x}'(u_0, v_0) = p' \in S'. \tag{6.66}$$

In conclusion, the geometric reversibility of a martensitic phase transformation expects:

- TPMS corresponding to two end phases can deform into each other continuously within a space of TPMS,
- Lattice points corresponding to phases occurring during the phase transformation should be compatible with this family of TPMS, with preimages of these lattice points in $\mathbb{C}$ being invariant.

6.4.2 Shape Memory Effect in NiTi

In this section, we use NiTi as an example to illustrate the statements given above. According to Sec. 6.3.1, we can build the Bravais lattice of NiTi in different phases using flat points on their corresponding TPMS. For example, we can build the Bravais lattice of B2 NiTi [Fig. 6.7(a,c)] using flat points on the P surface [Fig. 6.7(d)]. Hence, the ratios AB/AC and AC/AD should equal to the ratio of lattice parameters b/a and a/c [Fig. 6.7(b)], respectively. Also, because E, G and F are middle points[3] of AB, AD and AC, respectively [Fig. 6.7(c,d)], we have:

[3] Middle points are chosen here because they are on the P surface as shown in Fig. 6.7(d).

$$\frac{a}{c}=\frac{AC}{AD}=\frac{AF}{AG},\qquad \frac{b}{a}=\frac{AB}{AC}=\frac{AE}{AF}. \tag{6.67}$$

We can also see that the angles σ, φ and ω in Fig. 6.7(c) satisfy:

$$\alpha=\sigma,\qquad \beta=\omega,\qquad \gamma=\varphi, \tag{6.68}$$

where $\varphi=\arccos[(OC^2+OB^2-BC^2)/2OC\cdot OB]$. A similar argument can be made for materials in different phases discussed above (Sec. 6.3.1). Hence, by:

1. Calculating whether the lattice formed by flat points on TPMS can conform to the ratio and angle given by the lattice parameters of the crystal lattice;
2. Examining whether the point group of the crystal lattice is a subgroup of the point group of the TPMS;
3. Checking whether the skeletal graph of the TPMS conforms to the connection between lattice points and their nearest neighbours,

we can see whether a crystal can be described by a given TPMS or not. For the B19′ NiTi, we find that the flat points on the oPb surface can describe the corresponding crystal lattice [Fig. 6.8(a)]. Calculations related to the corresponding Gaussian curvature are given in Appendix D.8.5. Lattice parameters as labelled in Fig. 6.8(a) satisfy:

$$\frac{a}{c}=1.13,\qquad \frac{a}{b}=1.60,\qquad \gamma=97.82^\circ. \tag{6.69}$$

Comparison between the theoretical results and experimental/computational results are given in Table 6.3. We find that all end phases of NiTi have their Bravais lattices described by TPMS, with the point group of each phase being a subgroup of the point group of its corresponding TPMS (Table 6.2). The skeletal graph [Fig. 6.8(b)] is similar to the one given in Fig. 5.11(h,i,j),

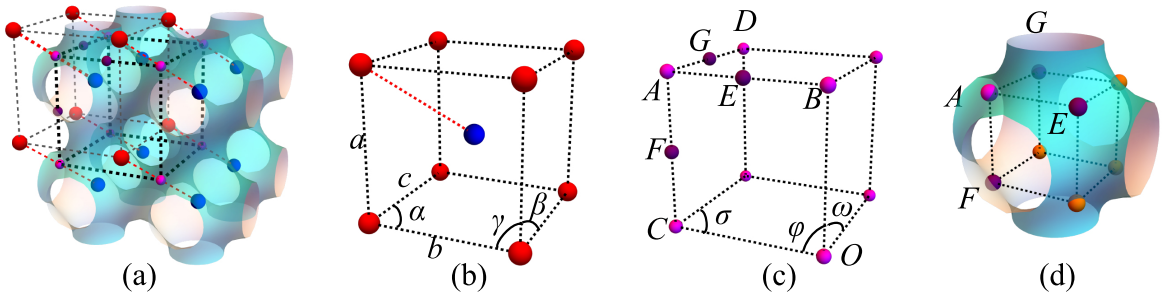


Figure 6.7 (a) Plot of the NiTi lattice (red and blue spheres represent Ni atoms and Ti atoms, respectively), lattice points (magenta spheres) and other flat points (purple spheres) against the P surface. (b) Lattice of a NiTi unit cell, with lattice parameters $(a, b, c; \alpha, \beta, \gamma)$ being labelled. (c) Sketch of the Bravais lattice of B2 NiTi given in (a). (d) Plot of one cell in (a) with lattice points.

Table 6.3 Parameters for the TPMS corresponding to phases along martensitic transformation paths of NiTi. For each phase, the references for computed and/or measured values of the ratios a/c and b/c of lattice parameters and values γ of the non-right angles are indicated. The corresponding values of these quantities obtained from our minimal surface theory are indicated by (T). The parameters a_s, b_s, and ψ determine each minimal surface from expressions given in Table 3.2. Weierstrass representations of TPMS describing these phases are given in Table 6.2.

Phase	a/c	a/b	a/c (T)	a/b (T)	γ	γ (T)	Parameters	TPMS
B2 [32,183,219]	1	1	1	1			$a_s = 0.5$, $\psi = 0$	P
B19′ [32,220]	1.13	1.60	1.13	1.60	98°	97.82°	$a_s \approx 0.096$, $b_s \approx 0.198$	oPb
B19′ [32,221]	1.13	1.60	1.13	1.60	97.78°	97.70°	$a_s \approx 0.096$, $b_s \approx 0.198$	oPb
B19′ [32]	1.15	1.60	1.15	1.60	98°	97.69°	$a_s \approx 0.098$, $b_s \approx 0.197$	oPb
B19′ [219]	1.16	1.60	1.16	1.60	98.97°	97.92°	$a_s \approx 0.099$, $b_s \approx 0.195$	oPb
B19 [32,222]	1.56	1.07	1.56	1.07			$a_s \approx 0.151$	oPa
B19 [32]	1.62	1.11	1.62	1.11			$a_s \approx 0.158$	oPa
B19 [219]	1.75	1.08	1.75	1.08			$a_s \approx 0.175$	oPa
B33 [201]	1.24	1.67	1.24	1.67	107.4°	107.4°	$a_s \approx 0.153$	oPa
B33 [201]	3.19	2.36	3.19	2.36			$a_s \approx 0.153$	oPa
R1 [32,223]	1.39	1	1.39	1	120°	120°	$a_s \approx 0.113, b_s \approx 0.220$	oPb
R1 [32]	1.36	1	1.36	1	120°	120°	$a_s \approx 0.111, b_s \approx 0.223$	oPb
R2 [43]	1.37	1	1.37	1	120°	120°	$a_s \approx 0.112, b_s \approx 0.222$	oPb
R2 [32]	1.39	1	1.39	1	120°	120°	$a_s \approx 0.113, b_s \approx 0.221$	oPb
R3 [32]	1.35	1	1.35	1	120°	120°	$a_s \approx 0.110, b_s \approx 0.224$	oPb
109°–B19′ [43]	1.13	1.61	1.13	1.61	108.88°	109.9°	$a_s \approx 0.070$, $b_s \approx 0.192$	oPb

conforming to the lattice of B19′ NiTi shown in Fig. 6.8(c) (also Fig. 5.21 in Sec. 5.3.3). This oPb surface, as pointed out by Chen and Weber [130], can continuously deform into the P surface within a space of TPMS. Similarly, we find other phases of NiTi can be located on TPMS in the same way. NiTi phases and their corresponding TPMS are given in Table 6.3. Sketches of Bravais lattices of NiTi in different phases against their corresponding TPMS will be shown later in this discussion. All lattice points on TPMS describing aforementioned NiTi phases are all flat points, and can be mapped to their corresponding TPMS by the same point in $\mathbb{C}$ (Appendix D.8.1, Appendix D.8.5 and Appendix D.8.6). We will then gain more insight into the phase transformation paths mentioned in Table 6.1 and explain what the vanishing

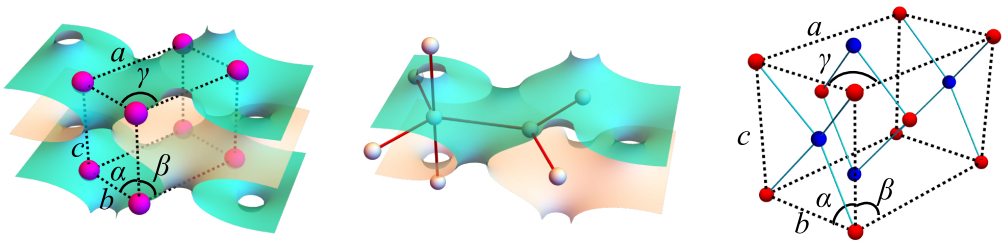


Figure 6.8 (a) Sketch of the Bravais lattice of the B19′ NiTi against the oPb TPMS. (b) Skeletal graph of this oPb TPMS. (c) Sketch of the crystal lattice of the B19′ NiTi.

of energy barrier means.

Path via a Single-layer or a Bilayer Shear

If we look at the deformation of the crystal lattice of B2 NiTi under a single-layer (resp. bilayer) $\langle 100\rangle_{B2}\{011\}_{B2}$, shear, as shown in Fig. 6.9(b) (resp. 6.9(c)), we find no minimal surface can describe such a lattice. For the single-layer shear, we see the deformed lattice needs the co-existence of a B19′ NiTi unit cell marked by the black line and a unit cell marked by the red (or blue) line. Treating the lattice marked by black lines given in Fig. 6.9(b) as an intermediate state of the B2−B19′ NiTi phase transformation, we see that according to Fig. 6.9(b), we can obtain the ratio condition when γ is set to be 98°:

$$\frac{l_1}{l_2} = 1.40, \qquad \gamma = 98^\circ, \tag{6.70}$$

we find a TPMS, as shown in Fig. 6.10(a), that is closest to the requirement given in Eq. 6.70, in which $l_1'/l_2' = 1.40, \gamma = 98^\circ$. In Fig. 6.9(b), γ_2 and γ_3 are required to be 141.89° and 98°, respectively. However, in Fig. 6.10(a), γ_2 and γ_3 equal to 125.09° and 112.86°, respectively. Moreover, $180^\circ - \gamma_2' = 151.31^\circ$, which means the TPMS given in Fig. 6.10(a) can not conform to the lattice given in Fig. 6.9(b). Thus, the deformation path via a single-layer $\langle 100\rangle_{B2}\{011\}_{B2}$ shear is not compatible with a space of TPMS.

Similarly, the bilayer $\langle 100\rangle_{B2}\{011\}_{B2}$ shear shown in Fig. 6.9(c) is not compatible with a space of TPMS. As shown in Fig. 6.9(c), $\gamma = \gamma_3 = 98^\circ$ and $\gamma_2 = 109.47^\circ$. The ratio $l_1/l_2 = 0.70$. The TPMS that we find can satisfy the aforementioned conditions is given in Fig. 6.10(b), with $l_1'/l_2' = 0.70, \gamma = \gamma_3 = 98^\circ$. However, we see $\gamma_2 = 112.86^\circ$. Moreover, for another choice of lattice points on a TPMS, as shown in Fig. 6.10(c), the conditions

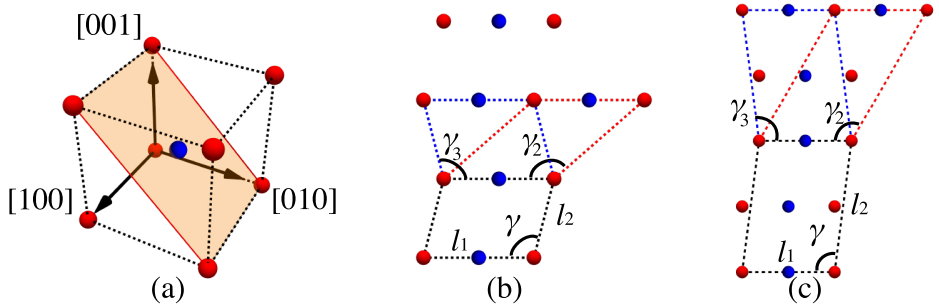
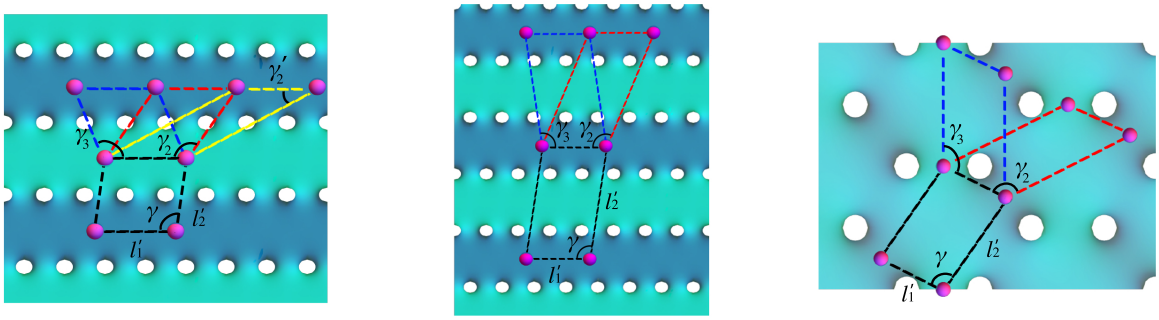


Figure 6.9 (a) B2 NiTi. Red and blue spheres representing Ni and Ti atoms, respectively. The orange plane labels the (011) plane. (b) and (c) are viewed from the $[0\bar{1}1]$ direction of (a). (b) Deformed B2 NiTi lattice under a $\langle 100\rangle_{B2}\{011\}_{B2}$ single-layer (basal) shear. The angle $\gamma = 98^\circ$, $\gamma \neq \gamma_2$. (c) Deformed B2 NiTi lattice under a $\langle 100\rangle_{B2}\{011\}_{B2}$ bilayer-shear. $\gamma = 98^\circ$, $\gamma \neq \gamma_2$.



(a) (b) (c)

Figure 6.10 Sketch of TPMS against deformed lattice of the B2 NiTi under (a) a basal $\langle 100\rangle_{\mathrm{B2}}\{011\}_{\mathrm{B2}}$ shear and (b,c) a bilayer $\langle 100\rangle_{\mathrm{B2}}\{011\}_{\mathrm{B2}}$ shear.

$l_1'/l_2' = 0.70$ and $\gamma = 98°$ cannot be satisfied simultaneously. Thus, this TPMS cannot suit the lattice given in Fig. 6.9(c). Hence, the deformation path caused by a bilayer $\langle 100\rangle_{\mathrm{B2}}\{011\}_{\mathrm{B2}}$ shear is not compatible with a space of TPMS.

Therefore, the path linking B2 NiTi and B19′ NiTi via a basal (resp. bilayer) $\langle 100\rangle_{\mathrm{B2}}\{011\}_{\mathrm{B2}}$ shear is not compatible with a space of TPMS. Hence, the corresponding value of the variation in the geometric entropy $\Delta S_{\mathrm{g}} > 0$.

Path B2–109°-B19′–B19′

This path is obtained by Hatcher and his colleagues [32] based on DFT calculations. They discovered that by allowing the co-existence of the bilayer and basal $\langle 100\rangle_{\mathrm{B2}}\{011\}_{\mathrm{B2}}$ shear, the B2 NiTi can transform to the B19′ phase along a path with vanishing energy barrier (Table 6.1). The 109°-B19 phase is an intermediate phase that occurs in this path [32, 43] when the displacement of the bilayer shear is about $0.46a$, where a is the lattice parameter. As shown in Fig. 6.11, (a,b,c) are viewed from the $[0\bar{1}1]_{\mathrm{B2}}$ direction. The lattice of the 109°-B19′ NiTi [Fig. 6.11(b)] can be described by a TPMS belonging to the oPb family [Fig. 6.11(d)], so can the lattice of the B19′ NiTi [Fig. 6.11(e)]. Lattice points on the oPb family surface forming the B19′ (resp. 109°-B19′) phase are mapped from $z = 0 \in \mathbb{C}$ by the Weierstrass representation, the same as the P surface (Appendix D.8.5). Hence, the compatibility between lattices occurring during the phase transformation and a space of TPMS means, for this transformation path, that $\Delta S_g = 0$.

Path B2–R

Consider the deformation of B2 NiTi by an elongation along $[111]_{\mathrm{B2}}$ to the R phase [170], as shown in Fig. 6.12. We can see the $[111]_{\mathrm{B2}}$ elongation [Fig. 6.13(a)] can be viewed as a combination as an extension of the lattice along $[110]_{\mathrm{B2}}$ and an extension along $[001]_{\mathrm{B2}}$.

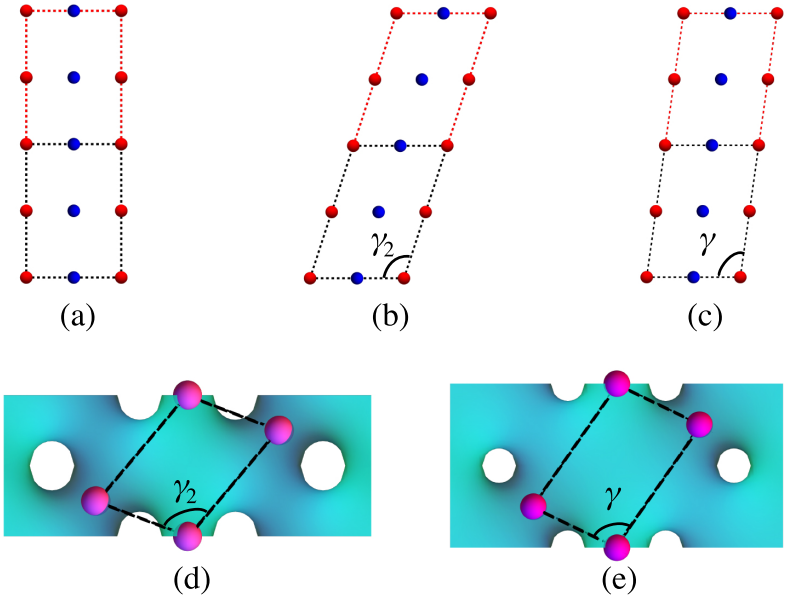


Figure 6.11 (a) Original B2 NiTi lattice viewed from $[0\bar{1}1]_{B2}$. (b) 109-B19′ lattice deformed from the B2 lattice under a combination of the bilayer and basal $\langle 100\rangle_{B2}\{011\}_{B2}$ shear. $\gamma_2 = 109°$ (c) B19′ lattice deformed from the B2 lattice under a combination of the bilayer and basal $\langle 100\rangle_{B2}\{011\}_{B2}$ shear. $\gamma = 98°$. (d) and (e) are TPMS belonging to the oPb family describing the 109°-B19′ and B19′ phase, respectively. The angles are $\gamma \approx 98°$, $\gamma_2 \approx 109°$.

When it comes to the corresponding deformation from the P surface to the R phase, we can see that it is a combination of a shear along the $[110]_{B2}$ direction [Fig. 6.13(a,b)] and a shear along the $[001]_{B2}$ direction [Fig. 6.13(c,d)]. The shear along $[001]_{B2}$ direction is for a compression along the z direction of the P surface. Thus, we see the shear path along $[111]_{B_2}$ conforms to deformation path from the P surface to the H surface within the oPb family. According to the statement we made in Sec. 5.3, this transformation path is continuous. Therefore, we have $\Delta S_g = 0$.

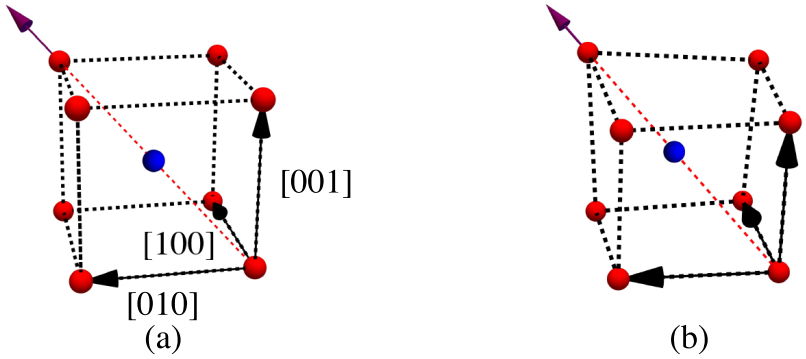


Figure 6.12 Sketch of (a) B2 NiTi deforming into (b) R NiTi via an $[111]_{B2}$ elongation. Position information of atoms in (b) are obtained from Materials Project [185–199]. Red and blue spheres are the Ni and Ti atoms, respectively. Purple arrows point in $[111]_{B2}$.

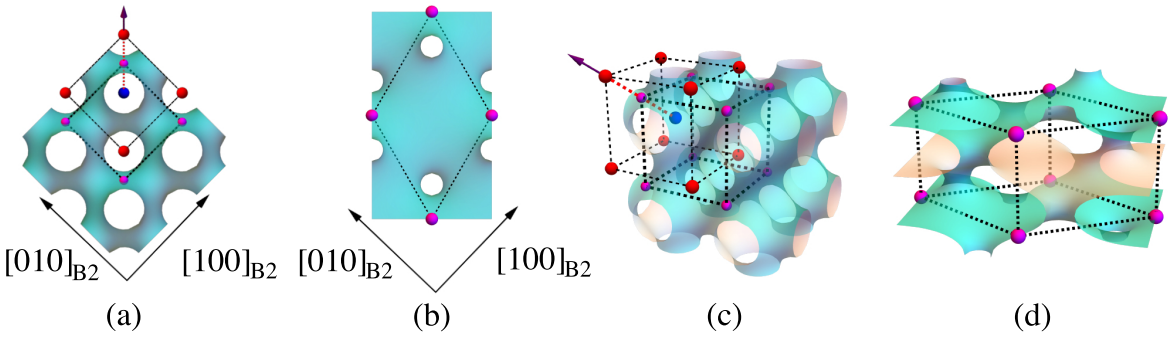


Figure 6.13 (a,b) Sketch of the NiTi crystal lattice and Bravais lattice against the P surface. (a) is viewed along $[00\bar{1}]_{B2}$. Red and blue spheres signify the Ni and Ti atoms. Magenta spheres signify the lattice point. (b,c) Sketch of the Bravais lattice of R NiTi against the corresponding H surface. (b) is viewed along $[00\bar{1}]_{B2}$ direction.

Path B2–B33

According to Morris *et al.* [200], we have the deformation as an alternate bilayer shear $\langle 100\rangle_{B2}\{011\}_{B2}$. That is to say, as sketched in Fig. 6.14(a) by black dashed lines, we see the new cell chosen in B2 NiTi has the angle $\gamma = 109.47°$. Let the bilayer marked by a blue dotted rectangle and the bilayer marked by a red dotted rectangle shear in different directions as indicated by black arrows; we will obtain a B33 NiTi lattice with the angle γ going to 107.41° [Fig. 6.14(b)]. Accordingly, we see that the P surface corresponding to the B2 phase deforms within a family of oPa surfaces. Red dotted lines in Fig. 6.15(b) give the new cell chosen in the B2 phase that will deform to the unit cell of the B33 phase [red dotted lines in Fig.6.15(d)]. Since the direction of the $\langle 100\rangle_{B2}\{011\}_{B2}$ shear is $[100]_{B2}$, we see that a compression of the P surface along $[100]_{B2}$ will conform to this shear. Hence we see a change from $\gamma = 109.47°$ in Fig. 6.15(b) to $\gamma = 107.41°$ in Fig. 6.15(d). Also, comparing Fig. 6.15(a,c), we see an expansion of the lattice along $[\bar{1}10]_{B2}$. Indeed, it conforms to the

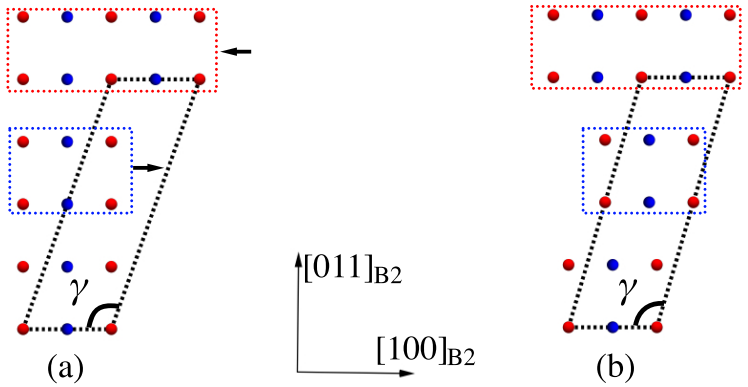


Figure 6.14 Sketch of (a) B2 NiTi viewed from the $[0\bar{1}1]_{B2}$ direction deforming under an alternate bilayer shear $\langle 100\rangle_{B2}\{011\}_{B2}$ to (b) B33 NiTi. Arrows in (a) denotes the shear direction. Bilayers marked by a red (resp. blue) dotted rectangle move in the same direction. The angle γ in (a) is 109.47° while in (b) is 107.41°.

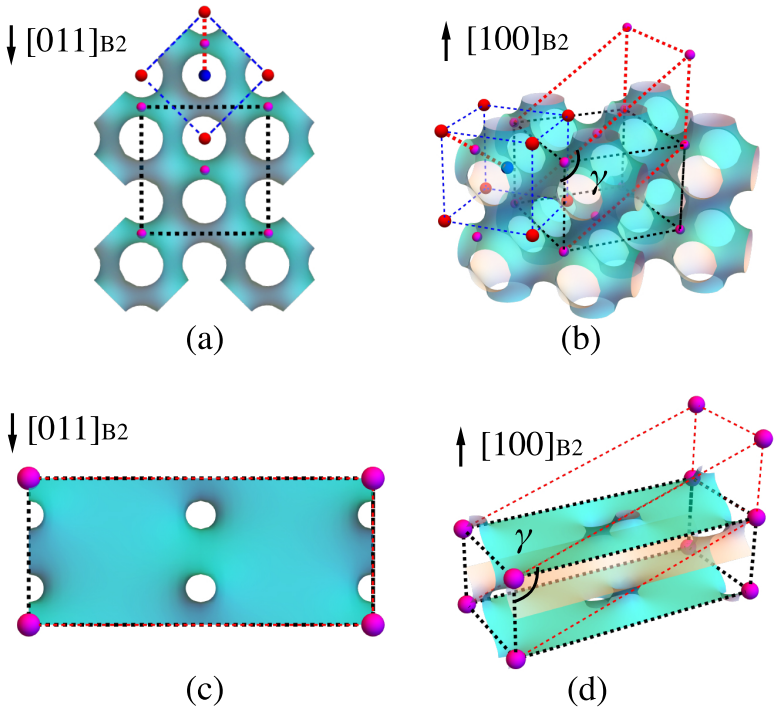


Figure 6.15 (a) Sketch of the crystal lattice and Bravais lattice of B2 NiTi against the P surface. Red and blue spheres are Ni and Ti atoms, respectively. Magenta spheres are lattice points. (a) is (b) viewed from the $[100]_{B2}$ direction. Blue dotted lines depict the unit cell of B2 NiTi. Red (resp. black) dotted lines in (b) depict the new cell that will deform into the unit cell of B33 NiTi sketched by red (resp. black) dotted lines in (d). In (b), $\gamma = 109.47°$. (c) is (d) viewed from the $[100]_{B2}$ direction. In (d), $\gamma = 107.41°$.

computational result that an increase in lattice parameters will occur when the shear takes place [32, 200]. In conclusion, we find the deformation path realised by an alternate bilayer shear $\langle 100\rangle_{B2}\{011\}_{B2}$ from B2 NiTi to B33 NiTi can reside in a space of TPMS. Hence, $\Delta S_g = 0$.

Path B2–B19

The deformation path we will discuss here is realised by a basal $\langle \bar{1}10\rangle_{B2}\{110\}_{B2}$ shear. As shown in Fig. 6.16, we see although the B33 NiTi can be described by an oPa surface, this transformation path cannot be realised within a space of TPMS. Hence for this path, we see $\Delta S_g > 0$. Referring to Table 6.3, we see the energy barrier is larger than zero.

6.4.3 ΔS_g and Energy Barrier

Comparing the energy barrier given in Table 6.1, we see paths with $\Delta S_g = 0$ should have their corresponding energy barrier go to zero. In addition to the argument we made in Sec. 6.4.2, other deformation paths, *i.e.*, Bain deformation, Bogers–Burgers deformation, Nishiyama–Wassermann deformation and Kurdjumow−Sachs deformation, are not topolo-

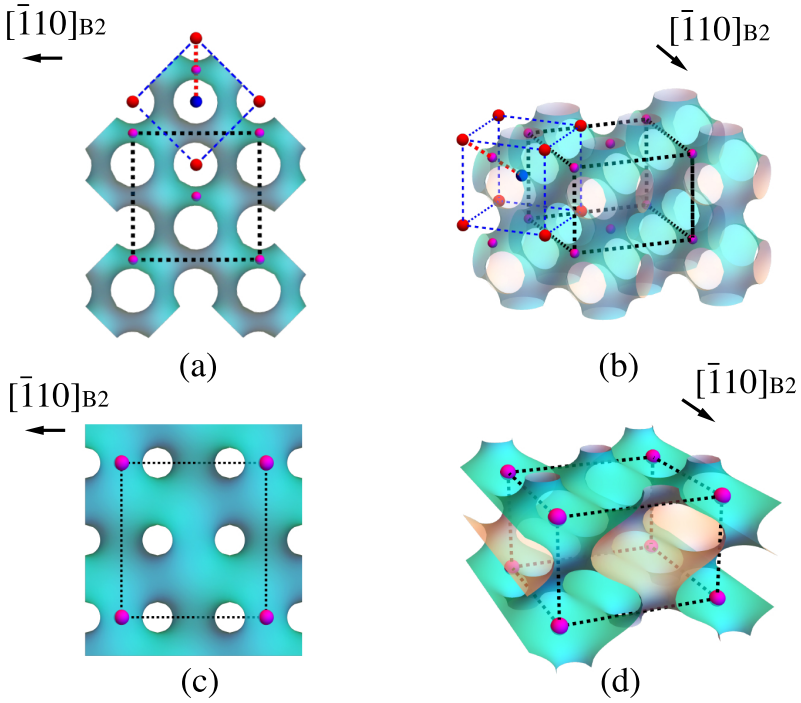


Figure 6.16 (a) Sketch of the crystal lattice and Bravais lattice of B2 NiTi against the P surface. Red and blue spheres are Ni and Ti atoms, respectively. Magenta spheres are lattice points. (a) is (b) viewed from the $[001]_{B2}$ direction. (c) is (d) viewed from the $[100]_{B2}$ direction. Black lines depict the unit cell of the B19 phase.

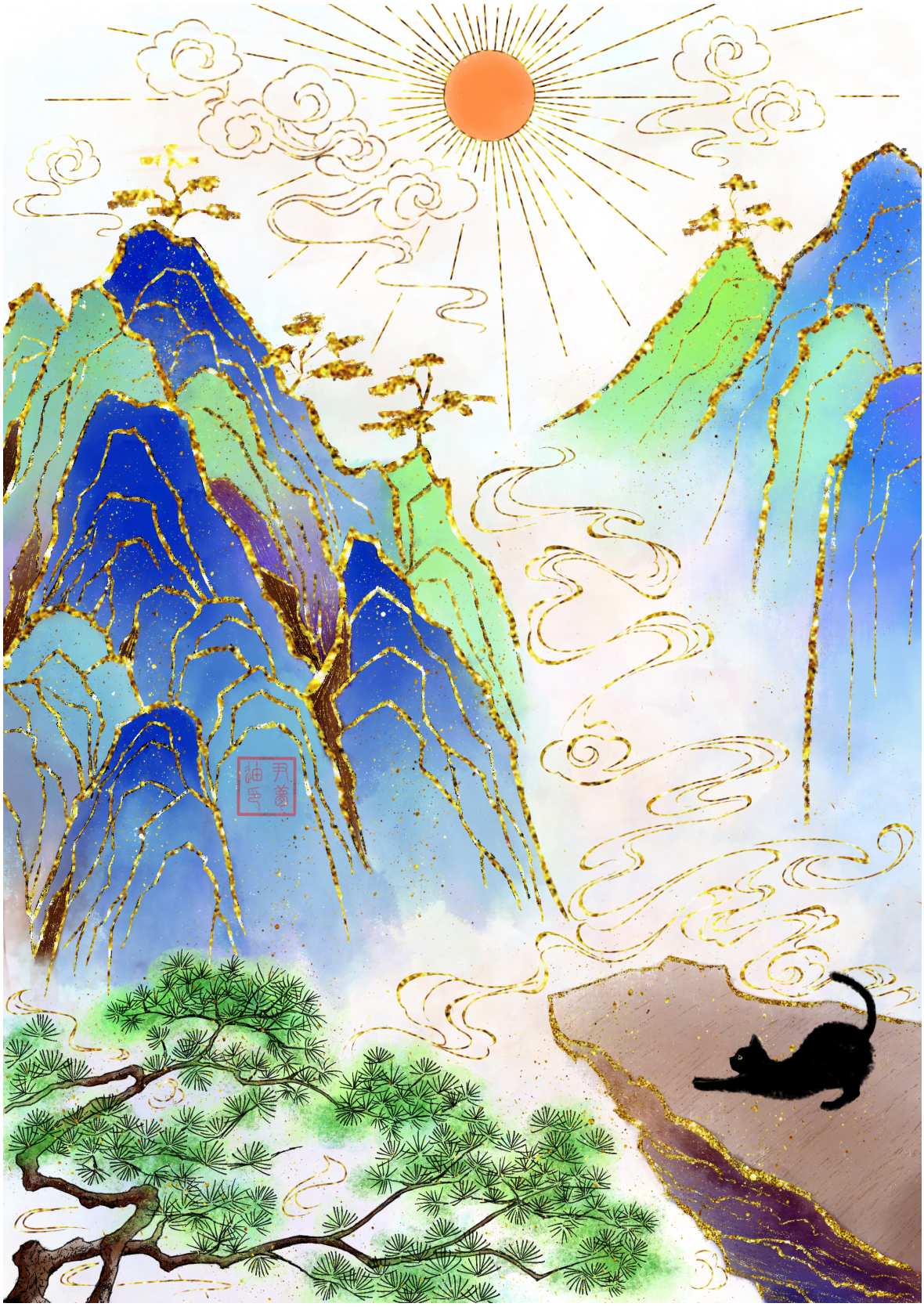
gically continuous (Sec. 6.4.2), given they always have $\Delta S_g > 0$, and hence, a non-zero energy barrier. Thus, we can conclude that a shape memory alloy has martensitic phase transformation paths with zero energy barriers. The vanishing of the energy barrier means $\Delta S_g = 0$, *i.e.*, this transformation path can live in the space of TPMS without lattice points gliding and is geometrically reversible. Moreover, although various transformation paths are listed above, it is not clear yet what the actual transformation path is. And this is principally because no experimental evidence has been provided.

6.5 Conclusion

In this chapter, we have discussed how to distinguish ordinary martensites from shape memory alloys with the help of geometry. According to DFT calculations, only shape memory alloys can have deformation paths with a vanishing energy barrier. Hence, in addition to the topological continuity, there are some other geometric factors that make shape memory alloys different from ordinary martensites. Hence, if we construct an "entropy" term, which is analogous to the thermodynamic entropy, based on these geometric properties, we can expect no variation in this geometric entropy for shape memory alloys that are deemed as undergoing "reversible" martensitic phase transformations.

According to our previous results, a crystal can find a correspondence to a TPMS if there are compatibilities between the symmetry, lattice parameters and connections between lattice points of this crystal and the symmetry, parameters (Sec. 6.4.2) and skeletal graph of this TPMS. Moreover, by looking at the surface of a given charge density value, we find that such a surface will converge to the corresponding TPMS when the value goes towards zero. Indeed, this convergence can be understood by considering the variation of the charge density ρ caused by the drift current (proportional to ρ) and the diffusion current (proportional to $\nabla\rho$). It can be proven that only when the surface is a minimal surface, we have no variation in ρ on this surface and ρ stays at zero on this surface (Sec. 6.3.3).

Hence, we proposed a term called “geometric energy" that consists of two parts: 1) bending energy E_b of the surface S of zero charge density and 2) kinetic energy $E_{\text{lat,tot}}$ of lattice points gliding on this surface. The variation in the geometric entropy is determined by the variation in the “geometric energy”. The term E_b is constructed using the Helfrich Hamiltonian and depends on the mean curvature M and the Gaussian curvature K over S. The term $E_{\text{lat,tot}}$, because it is confined on S, depends on the motion of the point confined on this surface and the local mean curvature where the point locates on S. For an ordinary phase transformation, its topological discontinuity can be told by a non-vanishing variation in E_b contributed by the integral of K over S, which, according to the Gauss-Bonnet theorem, gives a variation in genus. The ordinary martensitic phase transformation, however, has a non-vanishing variation in E_b contributed by M going from zero to a non-zero value. Variation in $E_{\text{lat,tot}}$ will be non-zero when there is a gliding of lattice points on S. Thus, if we want this phase transformation to be geometrically reversible, *i.e.* with no variation in the geometric entropy, then we expect this transformation path to stay in a space of TPMS with no gliding of lattice points on S. Examples of NiTi were provided in Sec. 6.4.2. All paths with a zero energy barrier calculated using DFT are geometrically reversible.



Clouds rise therefrom and lave my breast, my eyes are strained to see birds fleet. Try to ascend the mountain's crest, It dwarfs all peaks under our feet.[4] —-Du Fu

[4]Selected from *Gazing on Mountain Tai*, translated by Xu Yuanchong [224].

Chapter 7

Conclusion and Outlook

IN this chapter, we draw conclusions of our work and then discuss the application and research potential of this topic. The main idea of our work is using geometry to understand martensitic transformations and the shape memory effect. Martensitic phase transformations are diffusionless phase transformations that are mostly driven by a rapid change in the temperature. The parent (high-temperature) and the daughter (low-temperature) phases are named as the austenite and the martensite, respectively. Since martensitic phase transformations are diffusionless, the austenite and martensite phases are related to each other by a shear. A shear can be understood as a relative motion of two neighbour atomic layers in a crystal in opposite directions. If a deformed martensite can recover its original shape by heating, it is said to be a shape memory material and is thermoelastic. Theories such as the Landau's theory of phase transformations, the reconstructive theory and the continuum theory based on energy minimisation were proposed to study martensitic phase transformations and were reviewed in Chapter 2. These are all phenomenological theories and perform well in understanding martensitic phase transformations from the perspective of symmetry. Theories such as phonon softening that can be applied to specific materials were also proposed. Our theory is to look at martensitic phase transformations and the shape memory effect in materials involving no change in magnetism from geometry that not only can explain the phenomena, but can also be applied to specific materials. Our theory is built as follows.

First we studied the end phases of phase transformations with the help of density functional theory (DFT). We calculated the charge density distribution ρ in Cu, Al, Na, Zr and NiTi in different phases using the Vienna *Ab initio* Simulation Package (VASP) [138–141], and found that the surface S of a given charge density should converge to a triply periodic minimal surface (TPMS), such that the point group of the crystal is a subgroup, either proper or improper, of the point group of its corresponding TPMS. The convergence is calculated by evaluating the root-mean-square (RMS) of the deviation of S from its corresponding TPMS.

The deviation is obtained by calculating the minimal distance between each point on S and the TPMS. Moreover, we have discovered and then proved later that martensitic phases are required to have two end phases described by TPMS of the same genus, while ordinary phase transformations are not. To be specific, Cu and Al cannot undergo martensitic phase transformations, and we have them in FCC and BCC lattices described by the F-RD and the I-WP surfaces having genus 6 and 4, respectively. The BCC$-$HCP phase transformation in Na, and HCP$-$BCC/hexagonal phase transformation in Zr are martensitic and have the end phases described as the I-WP (BCC) and the H$'$-T (HCP/hexagonal) surfaces, both of which have genus 4.

Then, we used the result obtained in Chapter 4 to discuss why the end phases of a martensitic phase transformation are topologically equivalent. The main idea is based on the lattice points defined as a union of atoms such that in a crystal, all lattice points have the same physical and chemical environment. Because martensitic phase transformations are diffusionless and the relative arrangement of atoms is preserved, we can expect the connection between each lattice point and its nearest neighbours to be invariant. Since the connection between each lattice point and its neighbours conforms to the skeletal graph corresponding to the TPMS describing the crystal lattice, the invariance of such a connection means the skeletal graphs of the end phases are topologically equivalent to each other. Hence, the TPMS should undergo a continuous deformation. Like TPMS, a continuous deformation of a skeletal graph should involve no breaking but allow stretching and twisting. Since the "endpoint" of a skeletal graph is a junction of all arms and is part of a TPMS, we can stretch one endpoint to form new arms. The arm formed by stretching a junction is not a real connection between lattice points. Hence, it does not mean atoms are "broken" in such a process. However, it indicates lattice points in the new phase is formed by a combination of lattice points in the old phase. Given "lattice points" are only mathematical points in the same structural environment representing a union of atoms, such a "combination" of lattice points can take place during a phase transformation. For instance, in B19$'$ NiTi, a lattice point consists of two Ni atoms and two Ti atoms, which includes two lattice points in B2 NiTi, each of which consists of one Ni and one Ti atom. In other words, one lattice point in B19$'$ NiTi is formed by two lattice points in B2 NiTi. We used the transformation path of Cu, Al, Na, Zr and NiTi based on their crystal lattices corresponding to the end phases of a phase transformation to illustrate the argument we made in this chapter. Moreover, we can also find a unique correspondence between a crystal and a TPMS. That is to say, the TPMS correspond to the zero charge density surface S of a crystal satisfies: 1) the isogonal group of the crystal being a subgroup, either proper or improper, of the point group of the TPMS, 2) the connections between on lattice point and its nearest neighbours in this crystal conforming to the skeletal graph of this TPMS, 3) the

genus of S converging to the genus of the TPMS and 4) the Bravais lattice of this crystal can find a residence on this surface (as discussed in Sec. 6.4.2).

Based on the topological continuity required by martensitic phase transformations, in Chapter 6, we took a step further and looked at the shape memory effect. We compared the energy barrier calculated using DFT along various paths of phase transformations and discovered that only shape memory materials can have their energy barriers vanishing. Therefore, there should be additional geometric details that distinguish shape memory alloys from ordinary martensites. Since shape memory alloys are said to be "thermoelastic" and thus can undergo a "reversible" martensitic phase transformation, we would like to construct a quantity that is determined by the geometric properties of the system and can reveal the reversibility of a martensitic phase transformation. Thus, by looking at this quantity, we can tell whether a martensitic phase transformation is reversible or not. In thermodynamic physics, we use entropy S_p to distinguish reversible progresses ($\Delta S_p \leq 0$) from irreversible ones ($\Delta S_p > 0$). Hence, the term we proposed in this work to distinguish shape memory alloy from ordinary martensites is the "geometric entropy" S_g. Similar to the entropy defined in thermodynamic physics, the variation in S_g is expressed as $\Delta S_g = T\Delta E_g$. $\Delta E_g = \sum_i |\mathrm{d}E_g(t_i)|$ is a term having the dimension of energy depending on the geometric properties of a crystal system. t_i are parameters appearing in the interval $[0, t_f]$ used to parameterise the phase transformation process. We have $t_i = 0$ at the initial state and $t_i = t_f$ at the final state. The absolute value is taken here because what we care about here is the variation in E_g. E_g consists of two parts: 1) the bending of the surface S of given charge density ρ and 2) the gliding of lattice points on S. E_g is constructed in such a way for the following reasons.

First, according to Chapter 4, we see that S is determined by ρ that is influenced principally by the positions of atoms in a crystal. Therefore, the shape of S can reveal the structure and phase of the crystal. Secondly, based on the argument we made in Chapter 5 and Chapter 6, lattice points can be chosen to be points having the same Gaussian curvature on S as long as the Bravais lattice is preserved. When there is a phase transformation, given the positions of atoms will change, S will deform. Also, the motion of lattice points during a phase transformation can be treated as a combination of the motion together with S and a gliding on S. Hence, we write E_g as a sum of the bending energy of S and the energy of lattice points confined on S. By looking at the expression of E_g, we found that the vanishing of ΔE_g requires the end phases being linked by a continuous deformation within a space of TPMS without lattice points gliding on S. $\Delta E_g = 0$ gives $\Delta S_g = 0$, which means this phase transformation path is "reversible" or, equivalently, can exhibit the shape memory effect. Then we analysed the deformation paths of NiTi from its B2 phase to B19′, B33, B19 and 109°-B19′ phases via different shears.

We found that transformation paths having $\Delta S_g = 0$ also have their energy barrier vanishing. Additionally, for materials that cannot undergo a martensitic phase transformation, since they are topologically discontinuous, their energy barriers are also non-vanishing. In conclusion, topological continuity distinguishes martensites from ordinary materials, and geometric reversibility distinguishes shape memory alloys from ordinary materials. Phase transformation paths that are geometrically reversible also have their energy barrier vanishing. Additionally, by treating ρ as a continuous fluid and employing the continuity equation, we have found that when S is a minimal surface, ρ should go to zero. A flow chart showing the structure of this work is given in Fig. 7.1.

Our work provides a new perspective on studying martensitic phase transformations and the shape memory effect.

1. This method does not care about phase group-subgroup relations; hence, it can be applied to phase transformations such as the BCC–HCP phase transformation.

2. This method is time-saving and can provide a fast way to search for possible martensites and shape memory alloys.

3. This method does not need a large number of calculations and can not only explain the mechanism behind martensitic phase transformations but also distinguish possible martensites from ordinary materials. For instance, when the end phases of a material are provided, then by looking at their corresponding TPMS, we can tell whether this material can be a martensite or not. If this material is a martensite, then we can tell whether it is a shape memory alloy by looking at the corresponding orientation relationship that can be observed experimentally of this phase transformation and analysing whether it is compatible with a transformation path within a space of TPMS (as has been done in Sec. 6.4.2).

4. Although high precision DFT or methods beyond DFT (such as quantum Monte Carlo) can provide more precise results, the geometric method proposed in this thesis does not require that high precision. The precision will be enough as long as the calculation can be used to plot out the surface of a given charge density value S, and we can tell to which TPMS it converges. Such a convergence can be told by: 1) getting a mathematical description of this surface according to a fitting to S and calculating the genus g_N, 2) comparing the symmetry, links between one lattice point and its nearest neighbours of this crystal with the symmetry and the skeletal graph of TPMS having genus g_N and 3) checking whether the Bravais lattice of this crystal can be formed on TPMS satisfying these two conditions by locating lattice points at flat points (or points having the same geometry) on this surface.

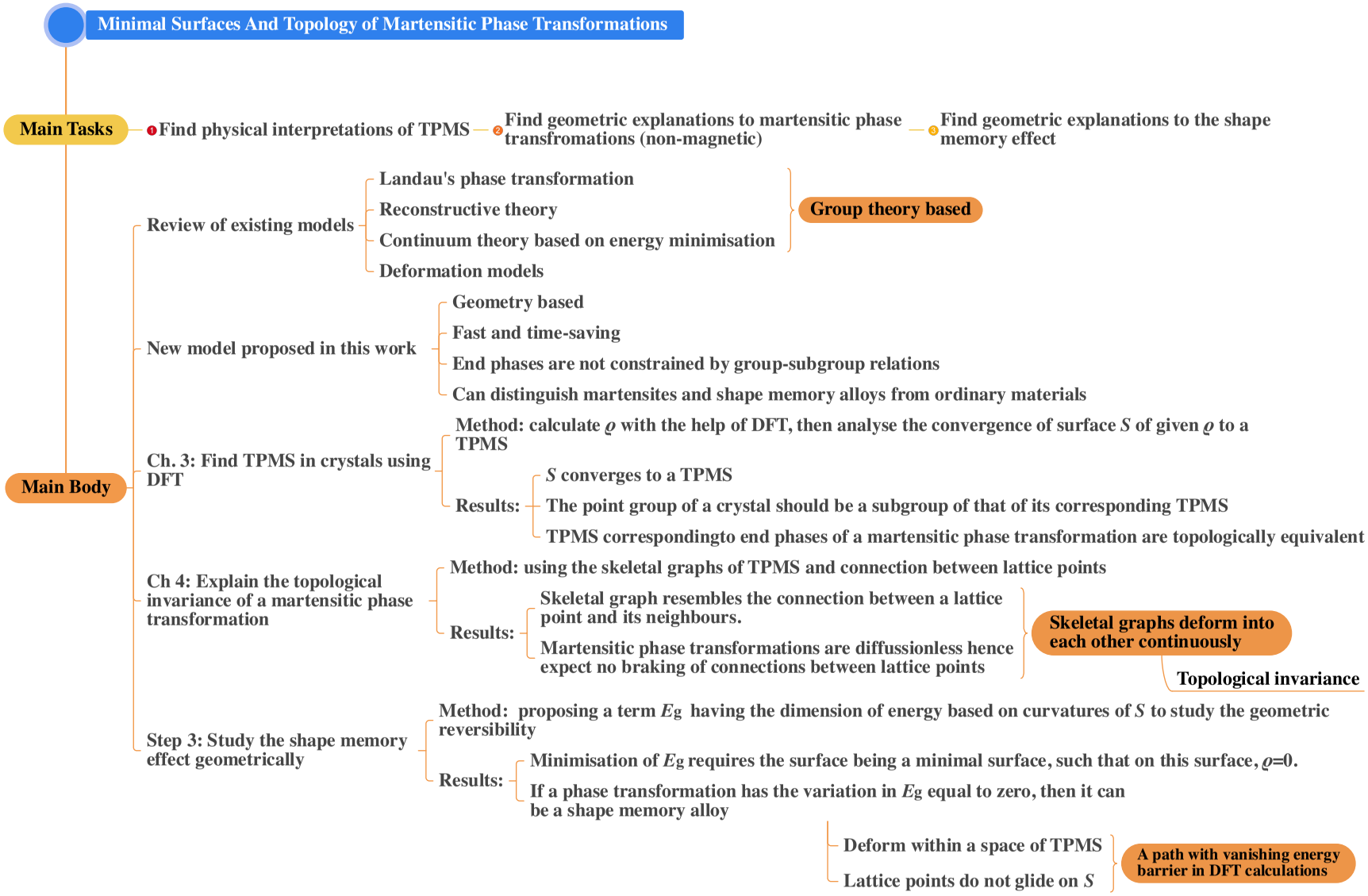

Minimal Surfaces And Topology of Martensitic Phase Transformations
Main Tasks
1 Find physical interpretations of TPMS
2 Find geometric explanations to martensitic phase transfromations (non-magnetic)
3 Find geometric explanations to the shape memory effect
Main Body
Review of existing models
Landau's phase transformation
Reconstructive theory
Continuum theory based on energy minimisation
Deformation models
Group theory based
New model proposed in this work
Geometry based
Fast and time-saving
End phases are not constrained by group-subgroup relations
Can distinguish martensites and shape memory alloys from ordinary materials
Ch. 3: Find TPMS in crystals using DFT
Method: calculate ϱ with the help of DFT, then analyse the convergence of surface S of given ϱ to a TPMS
Results:
S converges to a TPMS
The point group of a crystal should be a subgroup of that of its corresponding TPMS
TPMS correspondingto end phases of a martensitic phase transformation are topologically equivalent
Ch 4: Explain the topological invariance of a martensitic phase transformation
Method: using the skeletal graphs of TPMS and connection between lattice points
Results:
Skeletal graph resembles the connection between a lattice point and its neighbours.
Martensitic phase transformations are diffussionless hence expect no braking of connections between lattice points
Skeletal graphs deform into each other continuously
Topological invariance
Step 3: Study the shape memory effect geometrically
Method: proposing a term Eg having the dimension of energy based on curvatures of S to study the geometric reversibility
Results:
Minimisation of Eg requires the surface being a minimal surface, such that on this surface, ϱ=0.
If a phase transformation has the variation in Eg equal to zero, then it can be a shape memory alloy
Deform within a space of TPMS
Lattice points do not glide on S
A path with vanishing energy barrier in DFT calculations

Figure 7.1

Last but not least, our theory has enormous potential for future research.

1. We can gain more insight into how magnetism will influence this model, how the surface of a given charge density will be related to the Fermi surface, and how a TPMS will look like after a Fourier transformation from real space to the reciprocal space.

2. Since this model is geometrically based and not confined by a lattice periodicity, it can also be applied in the study of liquid crystals, liquid-solid phase transformations and the martensitic phase transformation in molecules.

A. Group Theory and Crystallography

A.1 Group Theory and Representation Theory

A.1.1 Abstract Groups

An abstract group is a set of elements $G = \{g_1, g_2, \cdots, g_n\}, n \in \mathbb{Z}^+$, such that we can define group multiplication $\circ$ that associates a pair of group elements to another. This set has the following properties:

1. Closure: $\forall g_i, g_j \in G, g_i \circ g_j = g_k, g_k \in G$;
2. Associative law: $\forall g_i, g_j, g_k \in G, (g_i \circ g_j) \circ g_k = g_i \circ (g_j \circ g_k)$;
3. Identity: $\exists E \in G, \forall g_i \in G, E \circ g_i = g_i \circ E = g_i$;
4. Inverse: $\forall g_i \in G, \exists g_i^{-1} \in G, g_i \circ g_i^{-1} = g_i^{-1} \circ g_i = E$

The number of elements in G is defined as the *order* of G, denoted as $|G|$. If a group G' finds all its elements belonging to G and has its order smaller or equal to $|G|$, G' is said to be a *subgroup* of G, *i.e.*, $G' \subseteq G$. If $|G'| = |G|$, G' is an *improper* subgroup of G, if $|G'| < |G|$, then G' is a *proper* subgroup of G [68].

According to the closure of a group, we can derive the *rearrangement theorem* [81], which states that if $g_1, g_2, \cdots, g_n$ are elements of a group G, then for all $g \in G$, $gg_1, gg_2, \cdots, gg_n$ is a rearrangement of $g_1, g_2, \cdots, g_n$. Here we use $gg_i, i = 1, 2, \cdots, n$ to represent $g \circ g_i$ and will use this convention in following contents.

If $G' \subset G$, then the order of G' is a factor of the order of G, and $|G|/|G'|$ is an **integer** named as the *index* of G' in G. This is called *Lagrange's theorem* [68]. Taking all elements $h = \{h_1, h_2, \cdots, h_k\}, k = |G|/|G'|$ in G but not in G', then by multiplying them on the left/right

hand side of G', we obtain two sets of *cosets* of G' ($|G'| = q$):

$$h_j G' = \{h_j, h_j g'_1, \cdots, h_j g'_q\}, \qquad h_j \in h, \qquad \text{left coset} \tag{A.1}$$

$$G' h_j = \{h_j, g'_1 h_j, \cdots, g'_q h_j\}, \qquad h_j \in h, \qquad \text{right coset} \tag{A.2}$$

Here, k is an integer and is the index of G' in G. If $\forall h_i \in h, h_i G' = G' h_i$, G' is said to be an *invariant* subgroup of G ($G' \trianglelefteq G$). The identity and G itself are two *trivial* invariant subgroups of G. **Except the identity and** G, other invariant subgroups G' of G are said to be *non-trivial* invariant subgroups of G ($G' \triangleleft G$). When $G' \trianglelefteq G$, G/G' is defined as a *quotient* group. Similarly, we can write a group as the product of another two groups. G is a *direct product* of G_1, G_2, *viz.*, $G = G_1 \otimes G_2$, if:

1. $G_1 \cap G_2 = \{E\}$;
2. $G_1 \circ G_2 = G_2 \circ G_1$;
3. $\forall g \in G$, it can be written uniquely as the product of $g_1 \in G_1$ and $g_2 \in G_2$;
4. $G_1 \triangleleft G, G_2 \triangleleft G$;

Moreover, G is a *semidirect product* of G_1, G_2, *viz.*, $G = G_1 \rtimes G_2$, if:

1. $G_1 \cap G_2 = \{E\}$;
2. $G_1 \triangleleft G$;
3. $\forall g$, it can be written uniquely as the product of $g_1 \in G_1$ and $g_2 \in G_2$, but in general $g_1 \circ g_2 \neq g_2 \circ g_1$.

Consider two groups G_1 and G_2. If we can construct a map $\varphi : G_1 \to G_2$, such that there is a correspondence:

$$\forall g_i, g_j \in G_1, g_i g_j = g_k, g_k \in G_1 \Rightarrow \varphi(g_i)\varphi(g_j) = \varphi(g_k), \tag{A.3}$$

where $\varphi(g_j), \varphi(g_k) \in G_2$, then we say G_1 and G_2 are *homomorphic* to each other. A homomorphism is a **many-to-one** correspondence [Fig. A.1(a)]. If we take a step further to require φ to be a **one-to-one** correspondence, then we say G_1 and G_2 are *isomorphic* to each other [Fig. A.1(b)] [68, 81].

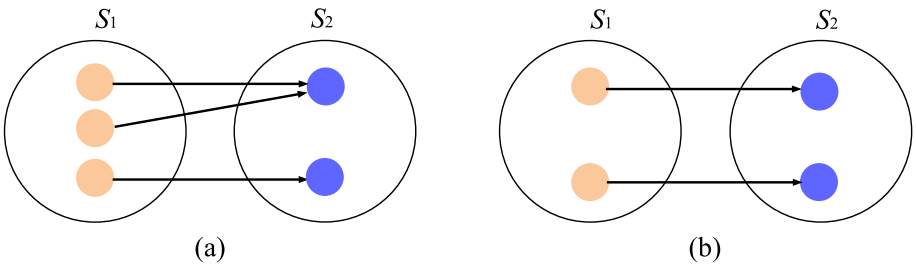


Figure A.1 Sketch of (a) the homomorphism, S_1 and S_2 are homomorphic to each other and (b) the isomorphism, S_1 and S_2 are isomorphic to each other.

A.1.2 Representation Theory

Beginning with a linear space $V \neq \{0\}$ defined on a field F (such as $\mathbb{R}$ and $\mathbb{C}$), then all invertible linear transformations on V can form a group $\mathrm{GL}(V)$. Here, "GL" stands for all "general linear transformations" and V has the dimension n. A *homomorphism* $\varphi : G \to \mathrm{GL}(V)$ is a **linear representation** of G on F and V is said to be the *representation space* of φ. If V is of finite dimension, then **the dimension of** V is said to be the *degree* of the representation φ. The *kernel* of φ is defined as $\mathrm{Ker}\varphi := \{g \in G | \varphi(g) = E_\varphi\}$, and is an invariant subgroup of G. E_φ is the identity of $\mathrm{GL}(V)$. Moreover, we have:

1. If $\mathrm{Ker}\varphi = \{E_\varphi\}$, then φ is a one-to-one correspondence, and φ is said to be *faithful*;
2. If $\mathrm{Ker}\varphi = G$, then φ is said to be *trivial*;
3. If the degree of φ equal to one and $\mathrm{Ker}\varphi = G$, φ is said to be the *identity representation* of G.

Since there is an isomorphism between $\mathrm{GL}(V)$ and $\mathrm{GL}_n(F)$, where $\mathrm{GL}_n(F)$ is a group formed by all $n \times n$ invertible matrices defined on F, then we can also define a *homomorphism* $\psi : G \to \mathrm{GL}_n(F)$ as a **matrix representation** of G on F. If two matrix representations ψ_1, ψ_2 defined on F are of the same dimension and can transform into each other by a **similarity transformation**, then ψ_1 and ψ_2 are said to be *equivalent*.

Consider a representation $\varphi : G \to \mathrm{GL}(V)$. If we can find a subspace $U \subset V, U \neq \{0\}$ and for all $g \in G$, $u \in U$, $\varphi(g)u \in U$, then φ is *reducible*. Otherwise φ is said to be *irreducible*. Moreover, the representation space V can be written as the direct sum of its invariant subspace U and the complementary space of U in V (W). Thus, for all vector $v \in V$, it can be uniquely expressed as the sum of a vector in U and a vector in W.

All reducible representations can be expressed as a **direct sum** of irreducible representations (IRREPs). The *character* χ of a matrix representation is defined as the **trace of the**

matrix of this representation, given it will be invariant under similarity transformations. If we want to decompose a reducible representation D of G to IRREPs D^i of G, we can refer to the *orthogonality relation* [81]:

$$n_i = \frac{1}{|G|}\sum_g \chi_D(g)\chi^*_{D_i}(g), \tag{A.4}$$

where D_i appears n_i times in D.

A.1.3 Crystallography and Crystallographic Groups

In crystals, a lattice point is defined as the union of atoms such that all lattice points are equivalent to each other, from the perspective of both symmetry and chemical environments. Due to the lattice periodicity, only 1,2,3,4,6-fold rotational symmetries, but no 5-fold rotational symmetry, are allowed. Together with mirror reflection symmetry and inversion symmetry, 32 point groups can be obtained. Point group operations are given in Table A1.

According to Fedorov [225], by adding translational symmetries one can extend 32 point groups to 230 space groups. Translations can contribute to the appearance of two symmetry operations:

1. Skew axis, which is a combination of a rotation by $2\pi/n$ about an n-fold rotational symmetry axis then a translation along this axis by $l\boldsymbol{\tau}/n, l \leq n, l \in \mathbb{Z}^+$, and $\boldsymbol{\tau}$ is a periodic vector along this axis;

2. Glide reflection, which is a combination of mirror reflection and translation by $\boldsymbol{\tau}/n$, $n = 2,4$ along a periodic vector $\boldsymbol{\tau}$ parallel to this mirror plane.

Hence, a space group operation $\{R|\boldsymbol{t}\}$ consists of two parts: 1) a point group operation R and 2) a translation $\boldsymbol{t}$. A space group has $\boldsymbol{t} = 0$ for all of its operations is said to be *symmorphic*, otherwise it is said to be *non-symmorphic* [81]. **The group formed by all point group operations presented in all operations of a space group** is defined as the *isogonal point group* of this space group [68]. Generally, there are two sets of symbols employed to describe space groups and point groups [226]:

Table A1 List of point group operations

Symbol	Operation
$C^m_{n\alpha}$	$n-$fold rotation around axis α by $2\pi m/n, m = 1,2,\cdots,n-1$
M_β	Mirror reflection about a plane whose normal is β
I	Inversion about the centre point
$IC^m_{n\alpha}$	Combination of a rotation $C^m_{n\alpha}$ and an inversion

1. **Schoenflies symbol**, written in the form $X_{n\alpha}^{j}$, where $X = C, S, D$ that stand for 'cyclic' (only rotations along the same axis), 'spiegel' (rotations along the same axis and an inversion) and 'dihedral' (rotations along different axes). $n = 2, 3, 4, 6, i$, representing a 2, 3, 4, 6-fold rotational symmetry or an inversion ("i"). $\alpha = h, v, d$ stand for 'horizontal', 'vertical' and 'diagonal'. j is a positive integer used to distinguish the space groups whose isogonal groups are the same;

2. **Hermann-Mauguin symbol**, with "P", "I", "F", "R", "C" ("A" or "B") denoting simple cubic, body-centred, face-centred, rhombohedral, and base-centred lattice, followed by symmetry operations along principal directions of the crystal (Table. A2).

Space group operations can be treated as a semidirect product of a point group and a translation group.

A.2 Lifshitz Criterion

Let us treat the density of free energy as a function of pressure, temperature, order parameter and spatial differentials of order parameters:

$$\varphi = \varphi\Big(P, T, \eta_i(\boldsymbol{r}), \nabla_{\boldsymbol{r}}\eta_i(\boldsymbol{r})\Big) = \int_{\Omega} \phi\Big(P, T, \eta_i(\boldsymbol{r}), \nabla_{\boldsymbol{r}}\eta_i(\boldsymbol{r})\Big)\mathrm{d}\boldsymbol{r}, \tag{A.5}$$

Table A2 Table of principal directions in different crystal systems [122, 226].

Crystal System	Principal directions
Cubic	$\langle 100\rangle$ $\langle 111\rangle$ $\langle 1\bar{1}0\rangle$
Hexagonal(Rhombohedral)	$\langle 001\rangle$ $\langle 010\rangle$ $\langle 1\bar{1}0\rangle$
Orthorhombic	$\langle 001\rangle$ $\langle 010\rangle$ $\langle 1\bar{1}0\rangle$
Tetragonal	$\langle 100\rangle$ $\langle 010\rangle$ $\langle 001\rangle$
Monoclinic	$\langle 010\rangle$
Triclinic	none

where i runs over the dimension of the order parameter space and ϕ is the free energy density, Ω is the system volume. Expanding ϕ around the critical point, we have that:

$$\phi = \phi_0 + (\partial_{\eta_j})\phi\eta_j + \partial_{\partial_{x_p}\eta_j}\phi(\partial_{x_p}\eta_j) + \frac{1}{2}\left[\frac{\partial^2\phi}{\partial\eta_i\partial\eta_j}\eta_i\eta_j + \frac{\partial^2\phi}{\partial\eta_i\partial(\partial_{x_p}\eta_j)}\eta_i(\partial_{x_p}\eta_j)\right. \tag{A.6}$$

$$\left. + \frac{\partial^2\phi}{\partial(\partial_{x_p}\eta_j)\partial\eta_i}(\partial_{x_p}\eta_j)\eta_i + \frac{\partial^2\phi}{\partial(\partial_{x_p}\eta_i)\partial(\partial_{x_q}\eta_j)}(\partial_{x_p}\eta_i)(\partial_{x_q}\eta_j)\right] + \cdots \tag{A.7}$$

At the critical point, we have the second and third terms vanishing, given the first-order derivative of ϕ_0 should equal zero. We can ignore quadratic terms because they are relatively small. Now setting $\partial^2\phi/\partial\eta_i\partial(\partial_{x_p}\eta_j) = A_{i,j,p}$, we have:

$$\phi = \phi_0 + A_{i,j,p}\eta_i(\partial_{x_p}\eta_j) + \cdots \tag{A.8}$$

$$= \phi_0 + \frac{1}{2}A_{i,j,p}\left[\eta_i(\partial_{x_p}\eta_j) + \eta_j(\partial_{x_p}\eta_i)\right] + \frac{1}{2}A_{i,j,p}\left[\eta_i(\partial_{x_p}\eta_j) - \eta_j(\partial_{x_p}\eta_i)\right] + \cdots \tag{A.9}$$

Because:

$$\eta_i(\partial_{x_p}\eta_j) + \eta_j(\partial_{x_p}\eta_i) = \partial_{x_p}(\eta_i\eta_j), \tag{A.10}$$

the integral $\int_\Omega \partial_{x_p}(\eta_i\eta_j)\mathrm{d}\boldsymbol{r}$ will go to $\eta_i\eta_j$ thus can be absorbed in φ_0. However, the term proportional to $\eta_i(\partial_{x_p}\eta_j) - \eta_j(\partial_{x_p}\eta_i)$ cannot go to $\eta_i\eta_j$ upon integration, hence will contribute inhomogeneity to the phase transformation. The **term** $\eta_i(\partial_{x_p}\eta_j) - \eta_j(\partial_{x_p}\eta_i)$ is the *Lifshitz invariant* [56, 227], and we call [80, 227]:

$$A_{i,j,p}\int_\Omega \mathrm{d}\boldsymbol{r}\left(\eta_i\frac{\partial\eta_j}{\partial x_p} - \eta_j\frac{\partial\eta_i}{\partial x_p}\right) = 0, \tag{A.11}$$

the *Lifshitz criterion*.

A.3 Powers of Irreducible Representations

Supposing that D is a representation of G that acts on a linear space L, with basis vectors $\{\psi_l\}, l = 1, 2, \cdots, \dim L$. Then the nth power of D, D^n (take direct product of D by n times), should act on L^n (take direct product of L n times). The basis of L^n is written as:

$$e^n_{i,j,\cdots,q} = (\psi^1_i, \psi^2_j, \cdots, \psi^n_q), \qquad i, j, \cdots, q = 1, 2, \cdots, \dim L, \tag{A.12}$$

dim$L$ is the dimension of L. Let S_n be a permutation group of degree n that is a group of all permutations of n objects. Then for any vector $v^n = \lambda^n_{i,j,\cdots,q} e^n_{i,j,\cdots,q} \in L^n$, $p \in S_n$:

$$pv^n = \lambda^n_{p(i,j,\cdots,q)} e^n_{i,j,\cdots,q}, \tag{A.13}$$

where $p(i,j,\cdots,q)$ gives a permutation of these q coordinates. Recall our definition of the representation in Sec. A.1.2. We see that by performing all operations $p \in S_n$ according to Eq. A.13, we can obtain a representation of S_n on L^n. Labell this representation by P. Let τ be an irreducible representation (IRREP) of S_n contained in P. Let the direct sum of all irreducible subspaces of L^n transforming according to τ be L^0. Because the permutation p commutes with operations in G, we can see that L^0 is invariant under all operations $D^n(g)$, $\forall g \in G$, such that:

$$D^n(g)[\tau(p)L^0] = D^n(g)L^0 = \tau(p)[D^n(g)L^0] \Rightarrow D^n(g)L^0 = L_0, \tag{A.14}$$

Because $D^n(g)L^0$ is still invariant under $\tau(p)$, we see that $D^n(g)L^0 = L^0$. Also, since $L^0 \subset L^n$, D^n can be decomposed into the direct sum of several representations, one of which is defined on L^0. Labell this representation by D^0. Let τ be a unit representation that **maps all permutations in S_n to the identity**, then D^0 is said to be an *symmetric nth power of D*, denoted by [83]:

$$D^0 = [D]^n. \tag{A.15}$$

Otherwise, If τ maps each **even permutation** in S_n to the **identity** while **odd permutations** to the **minus identity**, then D^n is said to be an *anti-symmetric nth power of D*, denoted as:

$$D^0 = \{D\}^n. \tag{A.16}$$

Here, an even (resp. odd) permutation means a permutation that can be decomposed into even (resp. odd) number of transpositions between two elements. For example, a permutation on the set $\{1,2,8\}$ that maps $1 \to 8, 2 \to 1, 8 \to 2$ gives $\{8,1,2\}$, which can be written as the product of $(21)(18)$ and is an even permutation.

Consider the order parameter space $\widetilde{L}$ with dimension n. Then polynomials of degree p formed by these order parameters will form a linear space $\widetilde{L}^p$. If $\widetilde{L}$ is invariant under a group G, since G acts linearly on $\widetilde{L}$, according to our previous statements, we have: 1) G acts linearly on $\widetilde{L}^p$ and 2) $\widetilde{L}^p$ carries a representation D^p of G, which is generally reducible.

Label the symmetry group of the parent phase by G. Let us decompose D^p into IRREPs of G. A decomposition containing the trivial representation of G means that there are polynomials of degree p being invariant under G [37]. Hence, for a decomposition of $[D]^3$, how

many times that the trivial representation D_1 of G occurs means how many linear-independent 3-degree polynomials we will have.

A.4 Group Theoretic Expression of the Landau Criterion and the Lifshitz Criterion

According to the statement given in Sec. A.3, we will see the group theoretical expression of the Landau criterion is [80]:

$$D_1 \notin [D]^3, \tag{A.17}$$

D_1 is the identity representation (which is trivial). Similarly, for the Lifshitz criterion, if we want $\int_\Omega \mathrm{d}\boldsymbol{r}(\eta_i \partial_{x_p}\eta_j - \eta_j\partial_{x_p}\eta_i) = 0$, then we must have $\eta_i\partial_{x_p}\eta_j - \eta_j\partial_{x_p}\eta_i \to \{D\}^2$ and $\boldsymbol{r} \to V$ (V is the vector representation of G). Therefore the vanishing of Lifshitz invariant means that we can find no D_1 in $V \otimes \{D\}^2$ [80]:

$$D_1 \notin V \otimes \{D\}^2. \tag{A.18}$$

Derivation of characters of $[D]^n$ can be found in Ref. [83]. Here we present the result:

$$[\chi(g)]^n = \sum_{\text{partition}} \frac{\chi^{r_1}(g^{q_1})\chi^{r_2}(g^{q_2})\cdots\chi^{r_\sigma}(g^{q_\sigma})}{r_1!q_1^{r_1}r_2!q_2^{r_2}\cdots r_\sigma!q^{r_\sigma}}, \tag{A.19}$$

where the sum runs over all possible different partitions of n in the form of:

$$n = r_1q_1 + r_2q_2 + \cdots + r_\sigma q_\sigma, \qquad q_i, r_i \in \mathbb{Z}^+, \tag{A.20}$$

where r_i are numbers of permutations with length equal to q_i. For example, when $n = 3$, then the partitions can be:

$$n = 1\times 3 = 3\times 1 = 1\times 1 + 1\times 2, \tag{A.21}$$

which means the character $[\chi]^2$:

$$[\chi(g)]^3 = \frac{\chi(g^3)}{1!3^1} + \frac{\chi^3(g^1)}{3!1^3} + \frac{\chi^1(g^1)\chi^1(g^2)}{1!1^1 1!2^1} \tag{A.22}$$

$$= \frac{1}{3}\chi(g^3) + \frac{1}{6}\chi^3(g) + \frac{1}{2}\chi(g)\chi(g^2) \tag{A.23}$$

As for $\{D\}^2$, however, a permutation between two elements is odd, hence will contribute a minus sign to the term $\chi(g^2)$, which means:

$$\{\chi(g)\}^2 = \frac{1}{2}[\chi^2(g) - \chi(g^2)] \tag{A.24}$$

Using the orthogonality condition (Eq. A.4), we now see:

1. Landau criterion:

$$\sum_g \left[\frac{1}{3}\chi(g^3) + \frac{1}{6}\chi^3(g) + \frac{1}{2}\chi(g)\chi(g^2)\right] = 0. \tag{A.25}$$

2. Lifshitz criterion:

$$\sum_g \left\{\frac{1}{2}\chi_V[\chi^2(g) - \chi(g^2)]\right\} = 0. \tag{A.26}$$

A.4.1 Irreducible Representation X_5^-

In this section, we provide the character table used in Sec. 2.2.1.

Table A3 Matrices of the X_5^- representation [226]

Operation	Rep. Matrix	$V(g)$	Operation	Rep. Matrix	$V(g)$
$\{E\|000\}$	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}$	$\{I\|000\}$	$\begin{pmatrix}-1&0\\0&-1\end{pmatrix}$	$\begin{pmatrix}-1&0&0\\0&-1&0\\0&0&-1\end{pmatrix}$
$\{C_{2z}\|000\}$	$\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\begin{pmatrix}-1&0&0\\0&-1&0\\0&0&1\end{pmatrix}$	$\{M_z\|000\}$	$\begin{pmatrix}0&-1\\-1&0\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&1&0\\0&0&-1\end{pmatrix}$
$\{C_{2y}\|000\}$	$\begin{pmatrix}-1&0\\0&-1\end{pmatrix}$	$\begin{pmatrix}-1&0&0\\0&1&0\\0&0&-1\end{pmatrix}$	$\{M_y\|000\}$	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&-1&0\\0&0&1\end{pmatrix}$
$\{C_{2x}\|000\}$	$\begin{pmatrix}0&-1\\-1&0\end{pmatrix}$	$\begin{pmatrix}1&0&0\\0&-1&0\\0&0&-1\end{pmatrix}$	$\{M_x\|000\}$	$\begin{pmatrix}0&1\\1&0\end{pmatrix}$	$\begin{pmatrix}-1&0&0\\0&1&0\\0&0&1\end{pmatrix}$
$\{C_{4y}\|000\}$	$\begin{pmatrix}0&1\\-1&0\end{pmatrix}$	$\begin{pmatrix}0&0&1\\0&1&0\\-1&0&0\end{pmatrix}$	$\{IC_{4y}\|000\}$	$\begin{pmatrix}0&-1\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0&-1\\0&-1&0\\1&0&0\end{pmatrix}$
$\{C_{2,101}\|000\}$	$\begin{pmatrix}1&0\\0&-1\end{pmatrix}$	$\begin{pmatrix}0&0&1\\0&-1&0\\1&0&0\end{pmatrix}$	$\{M_{101}\|000\}$	$\begin{pmatrix}-1&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0&-1\\0&1&0\\-1&0&0\end{pmatrix}$
$\{C_{4y}^3\|000\}$	$\begin{pmatrix}0&-1\\1&0\end{pmatrix}$	$\begin{pmatrix}0&0&-1\\0&1&0\\1&0&0\end{pmatrix}$	$\{IC_{4y}^3\|000\}$	$\begin{pmatrix}0&1\\-1&0\end{pmatrix}$	$\begin{pmatrix}0&0&1\\0&-1&0\\-1&0&0\end{pmatrix}$
$C_{2,\bar{1}01}\|000\}$	$\begin{pmatrix}-1&0\\0&1\end{pmatrix}$	$\begin{pmatrix}0&0&-1\\0&-1&0\\-1&0&0\end{pmatrix}$	$\{M_{\bar{1}01}\|000\}$	$\begin{pmatrix}1&0\\0&-1\end{pmatrix}$	$\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix}$

B. Differential Geometry Background

B.1 Quadratic Form and Congruence

For a linear space V of dimension n defined on a field F (such as $\mathbb{C}$ and $\mathbb{R}$), a *bilinear form* B on V should satisfy [47]:

1.
$$B(k_1u_1+k_2u_2,v)=k_1B(u_1,v)+k_2B(u_2,v),\qquad \forall k_1,k_2\in F,\,u_1,u_2,v\in V, \tag{B.1}$$

2.
$$B(u,k_1v_1+k_2v_2)=k_1B(u,v_1)+k_2B(u,v_2),\qquad \forall k_1,k_2\in F,\,u,v_1,v_2\in V, \tag{B.2}$$

such that B is linear in both u and v. Then let us chose a set of basis vectors $\{\alpha_i\}$, $i=1,2,\cdots,n$, for all $v\in V$, this vector v can be expressed as:

$$v=(\alpha_1,\alpha_2,\cdots,\alpha_n)(v_1,v_2,\cdots,\cdots,v_n)^{\mathrm{T}}. \tag{B.3}$$

Thus, we see the bilinear form B in basis $\{\alpha_i\}$ can be expressed in a matrix form $\bar{B}$, such that:

$$B(u,v)=B\left(\sum_{i=1}^{n}u_i\alpha_i,\sum_{j=1}^{n}v_j\alpha_j\right)=\sum_{i,j=1}^{n}u_iv_jB(\alpha_i,\alpha_j)=x^{\mathrm{T}}\bar{B}_{ij}y, \tag{B.4}$$

where:

$$x^{\mathrm{T}}=(u_1,u_2,\cdots,u_n),\qquad y=(v_1,v_2,\cdots,v_n)^{\mathrm{T}}. \tag{B.5}$$

Hence, we see that:

$$\bar{B}_{ij}=B(\alpha_i,\alpha_j). \tag{B.6}$$

Consider another set of basis vectors $\{\beta_i\}$, $i=1,2,\cdots,n$ in V. Let:

$$(\alpha_1,\alpha_2,\cdots,\alpha_n)C=(\beta_1,\beta_2,\cdots,\beta_n)\Rightarrow\beta_i=\sum_j\alpha_jC_{ji}, \tag{B.7}$$

where C is an $n \times n$ transition matrix between these two sets of basis. Labelling the matrix form of B in $\{\beta_i\}$ as $\widetilde{B}$, we see that:

$$\widetilde{B}_{ij} = B(\beta_i, \beta_j) = B\left(\sum_k \alpha_k C_{ki}, \sum_l \alpha_l C_{lj}\right) = \sum_{k,l=1}^{n} (C^{\mathrm{T}})_{ik} B(\alpha_k, \alpha_l) C_{lj} = (C^{\mathrm{T}} \overline{B} C)_{ij}. \tag{B.8}$$

Thus, we have:

$$C^{\mathrm{T}} \bar{B} C = \widetilde{B}. \tag{B.9}$$

If C is **nonsingular**, then we say $\widetilde{B}$ and $\bar{B}$ are *congruent to each other*. Specially, if B is **symmetric**, such that [47]:

$$B(u, v) = B(v, u), \tag{B.10}$$

then a *quadratic form* Q on V is defined as:

$$Q(v) = B(v, v), \qquad \forall v \in V. \tag{B.11}$$

B.2 Tensor Algebra and Differential Geometry

B.2.1 Definition of Metric Space

If we can define the **metric** $d : S \times S \to \mathbb{R}$ in a topological space S to measure the distance between between two elements $A, B \in S$, d should satisfy [124]:

1. Symmetric: $d(A,B) = d(B,A)$;
2. Positive-defined: $d(A,B) \geq 0$, $d(B,A) = 0$ only if $A = B$;
3. Triangle inequality: $d(A,B) \leq d(B,C) + d(C,A), C \in S$.

The 2D Euclidean space $\mathbb{R}^2$ is an example of the metric spaces, in Cartesian coordinates, let $A = (x_A, y_A), B = (x_B, y_B)$, then the distance between two points is given by:

$$d(A,B) = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2} \geq 0. \tag{B.12}$$

If we put A infinitely close to B, then the distance $d(A,B)$ is an infinitesimal element, with:

$$\mathrm{d}s = \lim_{A \to B} d(A,B), \mathrm{d}x = \lim_{A \to B} x_A - x_B, \mathrm{d}y = \lim_{A \to B} y_A - y_B. \tag{B.13}$$

Thus we can rewrite Eq. B.12 as:

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + 0\,\mathrm{d}x\,\mathrm{d}y + 0\,\mathrm{d}y\,\mathrm{d}x, \tag{B.14}$$

which, in matrix form:

$$ds^2 = \begin{pmatrix} dx & dy \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix}. \tag{B.15}$$

Now, come back to the basis vector expression of ds^2. Since we have the basis vectors $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}, \boldsymbol{e}_1 = (1,0)^{\mathrm{T}}, \boldsymbol{e}_2 = (0,1)^{\mathrm{T}}$, we can see that:

$$\begin{aligned} ds^2 &= \langle dx\boldsymbol{e}_1 + dy\boldsymbol{e}_2, dx\boldsymbol{e}_1 + dy\boldsymbol{e}_2 \rangle, \\ &= \begin{pmatrix} dx & dy \end{pmatrix} \begin{pmatrix} \langle \boldsymbol{e}_1, \boldsymbol{e}_1 \rangle & \langle \boldsymbol{e}_1, \boldsymbol{e}_2 \rangle \\ \langle \boldsymbol{e}_2, \boldsymbol{e}_1 \rangle & \langle \boldsymbol{e}_2, \boldsymbol{e}_2 \rangle \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix}. \end{aligned} \tag{B.16}$$

Comparing Eq. B.15 and Eq. B.16, we see that the metric here is in accordance with the first fundamental form given in Sec. 3.3.1. Moreover, we can now understand in a more direct way how the metric serves to measure the distance in a metric space.

Two topological spaces $(S_1, \{U_i^1\}), (S_2, \{U_j^2\})$ are said to be *homeomorphic* to each other if there exists a map $\sigma : S_1 \to S_2$ satisfying: 1) σ is **one-to-one** [Fig. B.1(a)] and **onto** (any $s_2 \in S_2$ has a preimage in S_1) [Fig. B.1(b)]; 2) σ and σ^{-1} are **continuous**. σ here is said to be a *homeomorphism*. If σ and σ^{-1} are differentiable, then σ is said to be a *diffeomorphism*. A 3D *manifold* is a **metric space** S such that for all $s \in S$, there exists a neighbourhood $U \subset S$ of s, U is **homeomorphic** to $\mathbb{R}^3$ [126]. The 3D Euclidean space $\mathbb{R}^3$ is a manifold. Every point in this space can be mapped to $(x,y,z), x,y,z \in \mathbb{R}$ in Cartesian coordinates. Moreover, the 2D Euclidean space $\mathbb{R}^2$ is a 2D manifold.

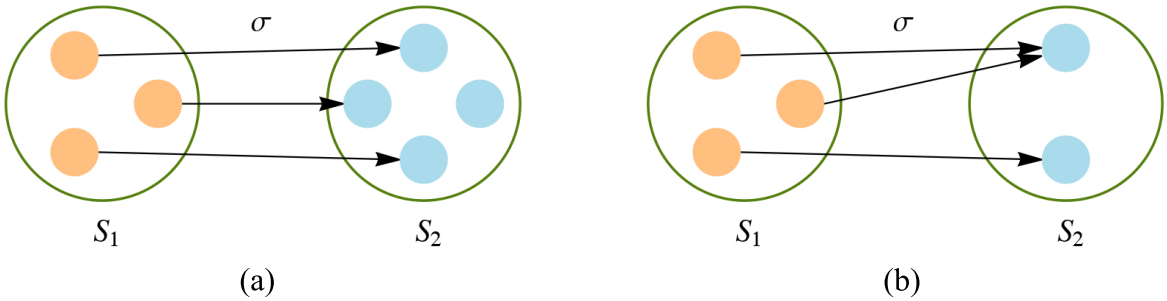


Figure B.1 Sketch of (a) a one-to-one map, every element in S_1 can find its image under σ in S_2; (b) an onto map, every element in S_2 can find its inverse image(s) under σ^{-1}.

B.2.2 Metric Differentiation

When it comes to the evolution of a surface $S = \boldsymbol{x}(u,v)$, $(u,v) \in \mathbb{R}^2$ obtained by a map $\boldsymbol{x} : \mathbb{R}^2 \to \mathbb{R}^3$, $\forall p \in S$, we can chose the basis $\{\boldsymbol{x}_u, \boldsymbol{x}_v, \boldsymbol{N}_p\} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_0\}$. $\boldsymbol{N}_p$ is an unit normal vector to S. In this basis, the corresponding coordinates are (x^1, x^2, t). Therefore we can parameterise an evolution of the surface along the normal vector at any $p \in S$ by t, and when $t = 0$, the point p is on the original surface (which is a 2D manifold). The boundary of S is empty given we are talking about a manifold (Appendix B.2.1).

Now we can discuss the evolution of g_{ij} (g is the metric defined in S, $i, j = 1, 2$, see Sec. 3.3.1 and Eq. E.7) with respect to t at $t = 0$. To describe an evolution along the normal vector parameterised by t, we can define a field $\boldsymbol{F} = \partial_t \boldsymbol{x}_t$, $\boldsymbol{x}_t = \boldsymbol{x}_t(u,v,t)$.

$$\begin{aligned}
\partial_t g_{ij} &= \partial_t \langle \partial_{x^i} \boldsymbol{x}_t, \partial_{x^j} \boldsymbol{x}_t \rangle_{t=0} \\
&= \langle \partial_{x^i} \boldsymbol{F}, \partial_{x^j} \boldsymbol{x}_t \rangle_{t=0} + \langle \partial_{x^i} \boldsymbol{x}_t, \partial_{x^j} \boldsymbol{F} \rangle_{t=0} \\
&= \partial_{x^i} \langle \boldsymbol{F}, \partial_{x^j} \boldsymbol{x}_t \rangle_{t=0} + \partial_{x^j} \langle \boldsymbol{F}, \partial_{x^i} \boldsymbol{x}_t \rangle_{t=0} - 2\langle \boldsymbol{F}, \partial_{x^i} \partial_{x^j} \boldsymbol{x}_t \rangle_{t=0}
\end{aligned} \tag{B.17}$$

Let $\boldsymbol{N}$ be the unit vector normal to S at $t = 0$, using Eq. E.5, we see [228]:

$$\partial_t g_{ij} = \partial_{x^i} \langle \boldsymbol{F}, \partial_{x^j} \boldsymbol{x}_t \rangle_{t=0} + \partial_{x^j} \langle \boldsymbol{F}, \partial_{x^i} \boldsymbol{x}_t \rangle_{t=0} - 2\Gamma^q{}_{ij} \langle \boldsymbol{F}, \partial_{x^q} \boldsymbol{x}_t \rangle_{t=0} - 2h_{ij} \langle \boldsymbol{F}, \boldsymbol{N} \rangle_{t=0}, \tag{B.18}$$

where $\{h_{ij}\}$ represent elements in the matrix of second fundamental form Π_2^p in basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ and [228]:

$$\frac{1}{2} g^{ij} h_{ij} = M, \tag{B.19}$$

M is the mean curvature. Using Eq. B.26, labelling $\omega_\alpha = \langle \boldsymbol{F} | \partial_{x^\alpha} \boldsymbol{x}_t \rangle$, $\alpha = i, j$, we have Eq. B.18 expressed as:

$$\begin{aligned}
\partial_t g_{ij} &= (\partial_i \omega_j - \Gamma^q{}_{ij} \omega_q) + (\partial_j \omega_i - \Gamma^q{}_{ij} \omega_q) - 2h_{ij} \langle \boldsymbol{F}, \boldsymbol{N} \rangle_{t=0}, \\
&= \nabla_i \omega_j + \nabla_j \omega_i - 2h_{ij} \langle \boldsymbol{F}, \boldsymbol{N} \rangle_{t=0}.
\end{aligned} \tag{B.20}$$

Hence, with Eq. B.24 [228]:

$$\begin{aligned}
\partial_t \sqrt{g(t)} \Big|_{t=0} &= \frac{g(g^{ij} \partial_t g_{ij})}{2\sqrt{g}} \Big|_{t=0}, \\
&= \frac{\sqrt{g}}{2} (g^{ij} \nabla_i \omega_j + g^{ij} \nabla_j \omega_i - 4M \langle \boldsymbol{F}, \boldsymbol{N} \rangle_{t=0}), \\
&= \sqrt{g} (g^{ij} \nabla_i \omega_j - 2M \langle \boldsymbol{F}, \boldsymbol{N} \rangle_{t=0}), \\
&= \sqrt{g} (\operatorname{div} \boldsymbol{F}_{T_p(S)} - 2M \langle \boldsymbol{F}, \boldsymbol{N} \rangle_{t=0}),
\end{aligned} \tag{B.21}$$

where $g = \det g_{ij}$, $\boldsymbol{F}_{T_p(S)}$ is the component of $\boldsymbol{F}$ in $T_p(S)$ (see Eq. B.28). Then Eq. 3.51 can be calculated as:

$$\begin{aligned}
\partial_t A(t)\Big|_{t=0} &= \int_S \partial_t \mathrm{d}S(t)\Big|_{t=0} \\
&= \int_S (\partial_t \sqrt{g})\mathrm{d}u\,\mathrm{d}v\Big|_{t=0} \\
&= \int_S (\operatorname{div} \boldsymbol{F}_{T_p(S)} - 2M\langle \boldsymbol{F}, \boldsymbol{N}\rangle)\mathrm{d}u\,\mathrm{d}v \\
&= -\int_S 2M\langle \boldsymbol{F}, \boldsymbol{N}\rangle \mathrm{d}u\,\mathrm{d}v
\end{aligned} \tag{B.22}$$

Here, we use following equations:

1. According to Eq. 3.45, labelling Π_2^p by $\{h_{ij}\}$, $i,j = 1,2$, we have [228]:

$$M = \frac{1}{2} g^{ij} h_{ij}, \tag{B.23}$$

2. The differential of determinant of a matrix gives [228]:

$$\partial_t(\det D) = (\det D)\ \mathrm{Tr}(D^{-1}\partial_t D) \tag{B.24}$$

3. $g^{ij}g_{ij} = \delta^i{}_i = 1$.

4. ω_i are the dual vector (also called 1-form) components, such that the dual vector basis $\{\omega_i^*\}$ corresponding to $\boldsymbol{e}_i$ is in the cotangent space $T_p^*(S)$, satisfying [229]:

$$\langle \omega_i^*, \boldsymbol{e}_j\rangle = \delta_{ij}, \tag{B.25}$$

   and for any $\boldsymbol{\omega} \in T_p^*(S), \boldsymbol{v} \in T_p(S)$, $\langle \boldsymbol{\omega}, \boldsymbol{v}\rangle = g_{ij}\omega^i v^j$.

5. The covariant derivative ∇_i is defined as [229]:

$$\nabla_i \omega_j = \partial_i \omega_j - \Gamma^q{}_{ij}\omega_q, \tag{B.26}$$

   which satisfies:

$$\nabla_i g_{ij} = \nabla_i g^{ij} = 0, \qquad \nabla_i \omega_j = \nabla_j \omega_i. \tag{B.27}$$

6. $\operatorname{div} \boldsymbol{F}_{T_p(S)}$ is the **divergence** of $\boldsymbol{F}_{T_p(S)}$, and in Eq. B.21 $\boldsymbol{\omega}^i$ describes basis vectors in $T_p(S)$. Here, g^{ij} raises the index, and transforms a dual vector component to a vector component. More details can be found in Ref. [229]. Here we have:

$$g^{ij}\nabla_i \omega_j = \nabla_i(g^{ij}\omega_j) = \nabla_i \omega^i = \operatorname{div} \boldsymbol{F}_{T_p(S)}. \tag{B.28}$$

7. The divergence theorem: $\int_S \text{div}\ \boldsymbol{F}_{T_p(S)} \mathrm{d}S = \int_{\partial S} \boldsymbol{F}_{T_p(S)} \cdot \boldsymbol{N} \mathrm{d}l$, when $\partial_S = \emptyset$:

$$\int_S \text{div}\ \boldsymbol{F}_{T_p(S)} \mathrm{d}S = 0 \tag{B.29}$$

B.2.3 Derivation of Weierstrass Representation

As shown in Fig. B.2, after a stereographic projection, $p \in S^2$ is projected onto p' in the equatorial plane, which is the intersection of $p_N p$ and the equatorial plane. Hence, in Cartesian coordinates, let $p = (X, Y, Z)$ and $p' = (x_c, y_c, 0)$. Therefore $p_0 = (X, Y, 0)$, and we see:

$$Z = \cos\theta. \tag{B.30}$$

Therefore, we have that:

$$x_c = \frac{X}{1-Z}, \qquad y_c = \frac{Y}{1-Z}, \tag{B.31}$$

and the element corresponding to $p' \in \mathbb{C}$:

$$\xi = x_c + \mathrm{i}y_c = \frac{X + \mathrm{i}Y}{1-Z}. \tag{B.32}$$

Now come back to a surface $S = \boldsymbol{x}(u,v), (u,v) \in U \subset \mathbb{R}^2$. Consider $w = u + \mathrm{i}v, w \in \mathbb{C}$ and then let (u,v) be **conformal**. Any surface $S = \boldsymbol{x}(u,v) \in \mathbb{R}^3$, if all partial differentials of $\boldsymbol{x}$ with respect to (u,v) up to order 2 exist and are continuous, then for any $p \in S$, there exists a neighbourhood U of p and local coordinates (u_l, v_l), such that the induced metric can be written as: $\mathrm{d}s^2 = \lambda^2(u_l, v_l)(\mathrm{d}u_l^2 + \mathrm{d}v_l^2)$, $\lambda^2(u_l, v_l)$ is a positive function. (u_l, v_l) is said to be

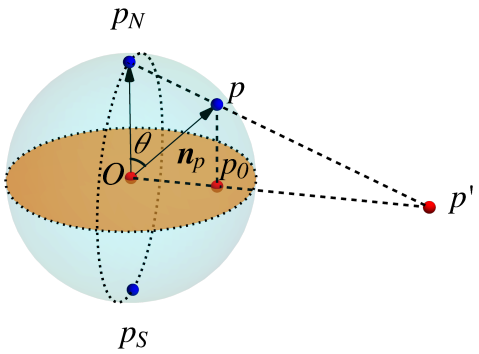


Figure B.2 Sketch of the stereographic projection. p_N, p_S are the North and South poles of the unit sphere, respectively. p is a point on this sphere, p' is the intersection of the line $p_S P$ and the equatorial plane, pp_0 is perpendicular to the equatorial plane and θ is the angle between pO and $p_N O$.

conformal [47,51]. We have [51]:

$$\frac{\partial}{\partial w} = \frac{1}{2}\left(\frac{\partial}{\partial u} - \mathrm{i}\frac{\partial}{\partial v}\right), \tag{B.33}$$

and:

$$\partial_w \boldsymbol{x} = \boldsymbol{f}(x,y,z) = (f_1, f_2, f_3), \tag{B.34}$$

Since the normal vector at $p \in S = \boldsymbol{x}(u,v), (u,v) \in U \subset \mathbb{R}^2$ is given by:

$$\boldsymbol{N} = \frac{\partial_u \boldsymbol{x} \times \partial_v \boldsymbol{x}}{|\partial_u \boldsymbol{x} \times \partial_v \boldsymbol{x}|}, \tag{B.35}$$

and let $\partial_u \boldsymbol{x} \times \partial_v \boldsymbol{x} = (\partial_u x, \partial_u y, \partial_u v) \times (\partial_v x, \partial_v y, \partial_v z) = (\Delta^1, \Delta^2, \Delta^3)$, where:

$$\begin{aligned} \Delta^1 &= 4\mathrm{Im}(f_2 \overline{f}_3), \\ \Delta^2 &= 4\mathrm{Im}(f_3 \overline{f}_1), \\ \Delta^3 &= 4\mathrm{Im}(f_1 \overline{f}_2). \end{aligned} \tag{B.36}$$

Taking Δ^1 as an example, we see:

$$\begin{aligned} \Delta^1 &= \partial_u y \partial_v z - \partial_v y \partial_u z, \\ &= \mathrm{Im}[(\partial_u y - \mathrm{i}\partial_v y)(\partial_u z + \mathrm{i}\partial_v z)], \\ &= 4\mathrm{Im}\partial_w y \overline{\partial_w z}, \\ &= 4\mathrm{Im}(f_2 \overline{f}_3). \end{aligned} \tag{B.37}$$

Hence, we can see that:

$$\begin{aligned} \partial_u \boldsymbol{x} \times \partial_v \boldsymbol{x} &= 4\mathrm{Im}(f_2 \overline{f}_3, f_3 \overline{f}_1, f_1 \overline{f}_2), \\ &= 2\left(f_2 \overline{f}_3 - (f_2 \overline{f}_3)^*, f_3 \overline{f}_1 - (f_3 \overline{f}_1)^*, f_1 \overline{f}_2 - (f_1 \overline{f}_2)^*\right), \\ &= 2\boldsymbol{f} \times \overline{\boldsymbol{f}}. \end{aligned} \tag{B.38}$$

Also, because (u,v) can chosen to be **conformal**:

$$|\partial_u \boldsymbol{x} \times \partial_v \boldsymbol{x}| = |\partial_u \boldsymbol{x}||\partial_v \boldsymbol{x}| = 2|\boldsymbol{f}|^2. \tag{B.39}$$

The normal vector $\boldsymbol{N}$ is now expressed as:

$$\boldsymbol{N} = \frac{\boldsymbol{f} \times \overline{\boldsymbol{f}}}{|\boldsymbol{f}|^2}. \tag{B.40}$$

Therefore, after a Gauss map, we have points on S represented by points on a unit sphere, with:

$$\boldsymbol{N} = (X, Y, Z), \tag{B.41}$$

we can see with Eq. B.32:

$$\xi = \frac{2\mathrm{Im}(f_2\overline{f}_3) + 2\mathrm{i}\ \mathrm{Im}(f_3\overline{f}_1)}{|\boldsymbol{f}|^2 - 2\mathrm{Im}(f_1\overline{f}_2)}, \tag{B.42}$$

where:

$$2\mathrm{Im}(f_2\overline{f}_3) + 2\mathrm{i}\ \mathrm{Im}(f_3\overline{f}_1) = f_3(\overline{f}_1 + \mathrm{i}\overline{f}_2) - \overline{f}_3(f_1 + \mathrm{i}f_2), \tag{B.43}$$

Because (u, v) is chosen to be conformal, we have:

$$\begin{aligned} \boldsymbol{f}\cdot\boldsymbol{f} &= \frac{1}{4}(\partial_u\boldsymbol{x} - \mathrm{i}\partial_v\boldsymbol{x})\cdot(\partial_u\boldsymbol{x} - \mathrm{i}\partial_v\boldsymbol{x}), \\ &= \frac{1}{4}(\langle\partial_u\boldsymbol{x}, \partial_u\boldsymbol{x}\rangle - \langle\partial_v\boldsymbol{x}, \partial_v\boldsymbol{x}\rangle), \\ &= 0. \end{aligned} \tag{B.44}$$

This means we have (f_1, f_2, f_3) satisfying the following equation:

$$f_1\cdot f_1 + f_2\cdot f_2 + f_3\cdot f_3 = 0 \Rightarrow f_1 + \mathrm{i}f_2 = -\frac{f_3^2}{f_1 - \mathrm{i}f_2}. \tag{B.45}$$

Thus, plugging Eq. B.45 into Eq. B.43, we can see that:

$$\begin{aligned} f_3(\overline{f}_1 + \mathrm{i}\overline{f}_2) - \overline{f}_3(f_1 + \mathrm{i}f_2) &= f_3(\overline{f}_1 + \mathrm{i}\overline{f}_2) + \frac{\overline{f}_3 f_3^2}{f_1 - \mathrm{i}f_2}, \\ &= \frac{f_3[(\overline{f}_1 + \mathrm{i}\overline{f}_2)(f_1 - \mathrm{i}f_2) + |f_3|^2]}{f_1 - \mathrm{i}f_2}, \\ &= \frac{f_3}{f_1 - \mathrm{i}f_2}[|\boldsymbol{f}|^2 - 2\mathrm{Im}(f_1\overline{f}_2)]. \end{aligned} \tag{B.46}$$

Combining with Eq. B.46 and Eq. B.43, we see that $\xi \in \mathbb{C}$ can be understood as the image of the Gauss map under a stereographic projection, and ξ can be written as:

$$\xi = \frac{f_3}{f_1 - \mathrm{i}f_2}. \tag{B.47}$$

Moreover, since we want the North/South pole to be mapped to ∞, the image of the sphere under a stereographic projection is extended to $\mathbb{C}^* = \mathbb{C}\cup\{\infty\}$ and is identified with the

Riemann sphere. Let us define:

$$g_s = \frac{f_3}{f_1 - \mathrm{i} f_2}, \qquad h = f_1 - \mathrm{i} f_2, \tag{B.48}$$

The pair of functions obtained in this way (h, g_s) is defined as the **Weierstrass representation** [51]. Equivalently, we can also define the pair as $(G_s, \mathrm{d}h)$ [129]:

$$G_s = -\frac{f_1 + \mathrm{i} f_2}{f_3}, \qquad \mathrm{d}h = f_3, \tag{B.49}$$

$\mathrm{d}h$ is the Z component of $\partial_w \boldsymbol{x}$, and is called the *height differential* [129]. Now we have the projection of S onto $\mathbb{C}^*$, and:

$$f_1 = \frac{\mathrm{d}h}{2}\left(\frac{1}{G_s} - G_s\right), \qquad f_2 = \frac{\mathrm{i}\,\mathrm{d}h}{2}\left(\frac{1}{G_s} + G_s\right), \tag{B.50}$$

Hence, by taking an integral, we can obtain the expression of S in Cartesian coordinates:

$$\begin{aligned} x &= \int \frac{\mathrm{d}h}{2}\left(\frac{1}{G_s} - G_s\right)\mathrm{d}w, \\ y &= \int \frac{\mathrm{i}\,\mathrm{d}h}{2}\left(\frac{1}{G_s} + G_s\right)\mathrm{d}w, \\ z &= \int \mathrm{d}h\,\mathrm{d}w. \end{aligned} \tag{B.51}$$

B.2.4 Gaussian Curvature in Weierstrass Representation

Consider a surface S parameterised by $\boldsymbol{x}(u,v), (u,v) \in U \subset \mathbb{R}^2$. In conformal coordinates, we have $|\partial_u \boldsymbol{x}|^2 = |\partial_v \boldsymbol{x}|^2 = \lambda^2$, $\langle \partial_u \boldsymbol{x}, \partial_v \boldsymbol{x} \rangle = 0$. Hence, the area $|\partial_u \boldsymbol{x} \times \partial_v \boldsymbol{x}| = \lambda^2 = 2|\boldsymbol{f}|^2$, which is combined with Eq. 3.59, yields:

$$\begin{aligned} |\boldsymbol{f}|^2 &= \frac{|\mathrm{d}h|^2}{4}\left\langle \left(\frac{1}{G_s} - G_s, \frac{\mathrm{i}}{G_s} + \mathrm{i}G_s, 2\right), \left(\overline{\frac{1}{G_s} - G_s}, \overline{\frac{\mathrm{i}}{G_s} + \mathrm{i}G_s}, 2\right)\right\rangle \\ &= \frac{|\mathrm{d}h|^2}{4}\left(\left|\frac{1}{G_s}\right|^2 + |G_s|^2 + \left|\frac{1}{G_s}\right|^2 + |G_s|^2 + 4\right) \\ &= \frac{|\mathrm{d}h|^2}{2}\left(\left|\frac{1}{G_s}\right|^2 + |G_s|^2 + 2\right) \\ &= \frac{|\mathrm{d}h|^2}{2}\left(\left|\frac{1}{G_s}\right| + |G_s|\right)^2. \end{aligned} \tag{B.52}$$

In the basis $\{\partial_u \boldsymbol{x}, \partial_v \boldsymbol{x}\}$ we have:

$$\begin{aligned} \partial_u \boldsymbol{N} &= k_{11}\partial_u \boldsymbol{x} + k_{12}\boldsymbol{x}_v, \\ \partial_v \boldsymbol{N} &= k_{21}\partial_u \boldsymbol{x} + k_{22}\boldsymbol{x}_v, \end{aligned} \tag{B.53}$$

Also, according to the definition of the second fundamental form we can see that:

$$\begin{aligned} \langle \partial_u \boldsymbol{N}, \partial_u \boldsymbol{x} \rangle &= -\langle \boldsymbol{N}_p, \partial_{uu}\boldsymbol{x} \rangle = -h_{11} = k_{11}g_{11} + k_{12}g_{21}, \\ \langle \partial_u \boldsymbol{N}, \partial_v \boldsymbol{x} \rangle &= -\langle \boldsymbol{N}_p, \partial_{uv}\boldsymbol{x} \rangle = -h_{12} = k_{11}g_{12} + k_{12}g_{22}, \\ \langle \partial_v \boldsymbol{N}, \partial_u \boldsymbol{x} \rangle &= -\langle \boldsymbol{N}_p, \partial_{vu}\boldsymbol{x} \rangle = -h_{21} = k_{21}g_{11} + k_{22}g_{21}, \\ \langle \partial_v \boldsymbol{N}, \partial_v \boldsymbol{x} \rangle &= -\langle \boldsymbol{N}_p, \partial_{vv}\boldsymbol{x} \rangle = -h_{22} = k_{21}g_{12} + k_{22}g_{22}. \end{aligned} \tag{B.54}$$

Therefore, with $g_{12} = g_{21}, h_{12} = h_{21}$:

$$\begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix} \begin{pmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{pmatrix} = - \begin{pmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{pmatrix}. \tag{B.55}$$

And we see:

$$\begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix} = \frac{1}{g} \begin{pmatrix} g_{22} & -g_{12} \\ -g_{12} & g_{11} \end{pmatrix} \begin{pmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{pmatrix}. \tag{B.56}$$

Thus, noticing that on a minimal surface $K \leq 0$, we have the normal vector satisfying:

$$\begin{aligned} |\partial_u \boldsymbol{N} \times \partial_v \boldsymbol{N}| &= |k_{11}k_{22}\ \partial_u \boldsymbol{x} \times \partial_v \boldsymbol{x} + k_{12}k_{21}\ \partial_v \boldsymbol{x} \times \partial_u \boldsymbol{x}|, \\ &= \left| \frac{\det \overline{\Pi}_2^p(S)}{\det \Pi_1^p(S)} \partial_u \boldsymbol{x} \times \partial_v \boldsymbol{x} \right|, \\ &= -K\lambda^2. \end{aligned} \tag{B.57}$$

In the Weierstrass representation, substituting Eq. B.52 into Eq. B.57, we have:

$$|\partial_u \boldsymbol{N} \times \partial_v \boldsymbol{N}| = -K\ |\mathrm{d}h|^2 \left(\left| \frac{1}{G_s} \right| + |G_s| \right)^2. \tag{B.58}$$

Next we construct another expression for $|\partial_u \boldsymbol{N} \times \partial_v \boldsymbol{N}|$ in terms of $|G_s|$, which is the stereographic projection $\pi_{\boldsymbol{N}}$ of $\boldsymbol{N}$. As shown in Fig. B.3(a), after a stereographic projection, a displacement along a meridian from p to q will be projected to $p'q'$ and as shown in Fig. B.3(b), a displacement $\overset{\frown}{AB}$ along a parallel will be projected to $\overset{\frown}{p'q'}$. Geometrically, the angle $\angle pp_N O = \theta/2$, which means:

$$p'O = \cot\frac{\theta}{2} \tag{B.59}$$

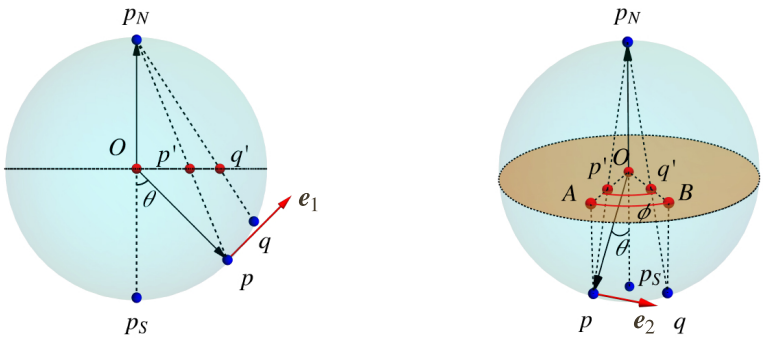


Figure B.3 Sketch of how points moving along the parallel and the meridian on a sphere.

Let $\{e_1, e_2\}$ be unit vectors along the meridian m and the parallel l at p [Fig. B.4(b)]. Then the stereographic projection of e_1, e_2 onto $\mathbb{C}^*$ yields $\pi_N(e_1), \pi_N(e_2)$. In spherical coordinates, $l = r\theta = \theta$ and $m = r\phi \sin\theta = \phi \sin\theta$. If we move p an infinitesimal length $\mathrm{d}t$ along e_1, then the image of m will move along $\pi_N(e_1)$ at a speed:

$$|\pi_N(e_1)| = \left|\frac{\mathrm{d}}{\mathrm{d}t} p'O\right| = \left|\frac{\mathrm{d}}{\mathrm{d}t} \cot\frac{\theta(t)}{2}\right| = \frac{1}{2\sin^2\frac{\theta}{2}}. \tag{B.60}$$

Similarly, when we move p a unit length along e_2, which means $\overset{\frown}{AB}$=1, we have the image of e_2 [as shown in Fig. B.3(b), represented by $\overset{\frown}{p'q'}$ and $OA = OB = \sin\theta$] changing along $\pi_N(e_2)$ at a speed:

$$|\pi_N(e_2)| = \left|\frac{p'O}{\sin\theta}\right| = \left|\frac{\cot\frac{\theta}{2}}{\sin\theta}\right| = \frac{1}{2\sin^2\frac{\theta}{2}} \tag{B.61}$$

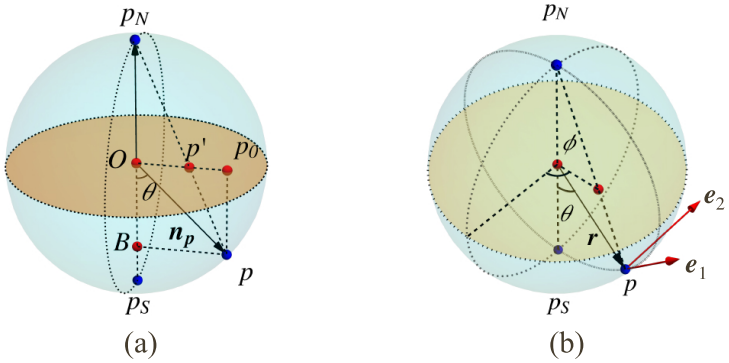


Figure B.4 Stereographic projection of p on to the equatorial plane. e_1 and e_2 are unit vectors along the meridian and the parallel.

Therefore, if $|\boldsymbol{e}_1| = |\boldsymbol{e}_2|$ before a stereographic projection, then after the projection, we still have: $|\pi_N(\boldsymbol{e}_1)| = |\pi_N(\boldsymbol{e}_2)|$. In the conformal coordinate, Eq. B.53 can be re-written as:

$$\begin{aligned}\partial_u \boldsymbol{N} &= -\frac{h_{11}}{\lambda}\partial_u \boldsymbol{x} - \frac{h_{12}}{\lambda}\partial_v \boldsymbol{x},\\ \partial_v \boldsymbol{N} &= -\frac{h_{12}}{\lambda}\partial_u \boldsymbol{x} - \frac{h_{22}}{\lambda}\partial_v \boldsymbol{x}.\end{aligned} \tag{B.62}$$

Recalling that $\partial_u \boldsymbol{N}, \partial_v \boldsymbol{N} \in T_p(S)$, we have:

$$\partial_u \langle \boldsymbol{N}, \partial_u \boldsymbol{N}\rangle = |\partial_u \boldsymbol{N}|^2 + \langle \partial_{uu}\boldsymbol{N}, \boldsymbol{N}\rangle = 0, \tag{B.63}$$

similar for $\partial_v \boldsymbol{N}$. We now have:

$$\begin{aligned}|\partial_u \boldsymbol{N}|^2 &= -\langle \partial_{uu}\boldsymbol{N}, \boldsymbol{N}\rangle = \langle -\frac{h_{11}}{\lambda}\partial_{uu}\boldsymbol{x} - \frac{h_{12}}{\lambda}\partial_{vu}\boldsymbol{x}, \boldsymbol{N}\rangle = \frac{h_{11}^2 + h_{12}^2}{\lambda},\\ |\partial_v \boldsymbol{N}|^2 &= -\langle \partial_{vv}\boldsymbol{N}, \boldsymbol{N}\rangle = \langle -\frac{h_{12}}{\lambda}\partial_{uv}\boldsymbol{x} - \frac{h_{22}}{\lambda}\partial_{vv}\boldsymbol{x}, \boldsymbol{N}\rangle = \frac{h_{12}^2 + h_{22}^2}{\lambda}.\end{aligned} \tag{B.64}$$

Noticing that in this conformal coordinate, the mean curvature of this minimal surface now is expressed as:

$$M \propto k_{11} + k_{22} = \frac{1}{\lambda}(h_{11} + h_{22}) = 0 \Rightarrow h_{11} = -h_{22}, \tag{B.65}$$

we see that:

$$|\partial_u \boldsymbol{N}|^2 = |\partial_v \boldsymbol{N}|^2, \tag{B.66}$$

where we have $\langle \partial_u \boldsymbol{N}, \partial_v \boldsymbol{N}\rangle = 0$. Hence, the image $\{\partial_u \boldsymbol{N}, \partial_v \boldsymbol{N}\}$ on $\mathbb{C}$ satisfies:

$$|\pi_N(\partial_u \boldsymbol{N})| = |\pi_N(\partial_v \boldsymbol{N})|. \tag{B.67}$$

Moreover, according to our statement in Appendix B.2.3, we have:

$$\pi_N(\boldsymbol{N}) = \xi = G_s, \tag{B.68}$$

Referring to Fig. B.5, we see that the point $p' = \pi_N(p) = G_s$, where $p = \boldsymbol{N}$ is a point on the Gauss map sphere. Labelling $\boldsymbol{e}_1 || \partial_u \boldsymbol{N}$, $\boldsymbol{e}_2 || \partial_v \boldsymbol{N}$, $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ are orthonormal vectors. Then moving p to q along $\boldsymbol{e}_1$, we can map out $\partial_u \boldsymbol{N}$, whose the image under π_N draws an arc $\overset{\frown}{p'q'}$ on the complex plane and $|\pi_N(\partial_u \boldsymbol{N})| = |\overset{\frown}{p'q'}|$ can be represented by $|\mathrm{d}G_s/\mathrm{d}w|$, $w \in \mathbb{C}$ (the orange plane in Fig. B.5). Similarly, moving p to C, we find the image of $\partial_v \boldsymbol{N}$ draws a line $|Ap'| = |\mathrm{d}G_s/\mathrm{d}w|$. Thus, the area spanned by $\{\pi_N(\partial_u \boldsymbol{N}), \pi_N(\partial_v \boldsymbol{N})\}$, $ABq'p'$ marked by the

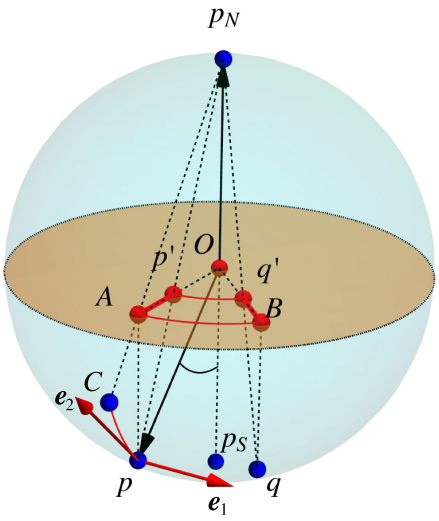


Figure B.5 Sketch of points on a Gauss map sphere and their images under a stereographic projection.

red line in Fig. B.5, can be written as:

$$|\pi_N(\partial_u \boldsymbol{N}) \times \pi_N(\partial_v \boldsymbol{N})| = \left|\frac{\mathrm{d}G_s}{\mathrm{d}w}\right|^2. \tag{B.69}$$

Now, we substitute Eq. B.60 and Eq. B.61 into Eq. B.32 and find:

$$|\xi| = |p'O| = \cot\frac{\theta}{2}, \tag{B.70}$$

and:

$$\frac{1}{2\sin^2\frac{\theta}{2}} = \frac{1}{2}\frac{\sin^2\frac{\theta}{2}+\cos^2\frac{\theta}{2}}{\sin^2\frac{\theta}{2}} = \frac{1+|\xi|^2}{2} = \frac{1+|G_s|^2}{2}. \tag{B.71}$$

We now see that:

$$|\partial_u \boldsymbol{N} \times \partial_v \boldsymbol{N}| = -K\,|\mathrm{d}h|^2\left(\left|\frac{1}{G_s}\right| + |G_s|\right)^2 = \frac{4}{(1+|G_s|^2)^2}\left|\frac{\mathrm{d}G_s}{\mathrm{d}w}\right|^2, \tag{B.72}$$

and on a minimal surface $K \leq 0$, which means [51]:

$$K = -\frac{4|G_s|^2}{|\mathrm{d}h|^2(1+|G_s|^2)^4}\left|\frac{\mathrm{d}G_s}{\mathrm{d}w}\right|^2 \tag{B.73}$$

C. Density Functional and Root-Mean-Square Calculations

C.1 Calculating Bloch States in Reciprocal Space

Let us begin with the Kohn–Sham equation [148, 153]:

$$\hat{H}\phi_i(\boldsymbol{r}) = [-\frac{\hbar^2}{2m_e}\nabla^2 + V_{\text{eff}}(\boldsymbol{r})]\phi_i(\boldsymbol{r}) = \varepsilon_i\phi_i(\boldsymbol{r}). \tag{C.1}$$

Both of $|\phi_i(\boldsymbol{r})\rangle$ and V_{eff} can be expanded in a Fourier series using vectors in $\boldsymbol{k}$-space:

$$\phi_{\boldsymbol{k},i}(\boldsymbol{r}) = \sum_{\boldsymbol{q}} c_{\boldsymbol{q},i}(\boldsymbol{k})\exp[\mathrm{i}(\boldsymbol{q}+\boldsymbol{k})\cdot\boldsymbol{r}], \qquad V_{\text{eff}}(\boldsymbol{r}) = \sum_{\boldsymbol{q}} V_{\text{eff}}(\boldsymbol{q})\exp(\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{r}). \tag{C.2}$$

The periodicity conditions $V_{\text{eff}}(\boldsymbol{r}+\boldsymbol{R}) = V_{\text{eff}}(\boldsymbol{r})$ require $\boldsymbol{q} = \boldsymbol{G}$, where $\boldsymbol{G}$ are reciprocal lattice vectors with:

$$\boldsymbol{G}\cdot\boldsymbol{R} = 2\pi n, \qquad n \in \mathbb{Z}, \tag{C.3}$$

where $\boldsymbol{R}$ is a lattice vector. Hence, we see:

$$V_{\text{eff}}(\boldsymbol{r}+\boldsymbol{R}) = V_{\text{eff}}(\boldsymbol{r}) \Rightarrow V_{\text{eff}}(\boldsymbol{r}) = \sum_{\boldsymbol{G}} V_{\text{eff}}(\boldsymbol{G})\exp(\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r}). \tag{C.4}$$

Therefore, in the Fourier space, let us write $|\boldsymbol{G}\rangle = \exp(\mathrm{i}\boldsymbol{G}\cdot\boldsymbol{r})$ and $\boldsymbol{q} = \boldsymbol{k}+\boldsymbol{G}_m$, we see that:

$$\begin{aligned}\langle \boldsymbol{G}'+\boldsymbol{k}|\hat{H}|\boldsymbol{G}+\boldsymbol{k}\rangle &= \langle \boldsymbol{G}'+\boldsymbol{k}| - \frac{\hbar^2}{2m_e}\nabla^2|\boldsymbol{G}+\boldsymbol{k}\rangle + \langle \boldsymbol{G}'+\boldsymbol{k}|V_{\text{eff}}|\boldsymbol{G}+\boldsymbol{k}\rangle \\ &= \frac{\hbar^2|\boldsymbol{G}+\boldsymbol{k}|^2}{2m_e}\delta_{\boldsymbol{G}',\boldsymbol{G}} + \sum_{\boldsymbol{G}''} V_{\text{eff}}(\boldsymbol{G}'')\delta_{\boldsymbol{G}'',\boldsymbol{G}'-\boldsymbol{G}}.\end{aligned} \tag{C.5}$$

By substituting Eq. C.4 into Eq. C.1 and left multiplying Eq. C.1 by $\langle \boldsymbol{G}' + \boldsymbol{k}|$, we have:

$$\sum_{\boldsymbol{G}} \langle \boldsymbol{G}' + \boldsymbol{k}|\hat{H}|\boldsymbol{G} + \boldsymbol{k}\rangle c_{\boldsymbol{G},i} = \sum_{\boldsymbol{G}} \varepsilon_i c_{\boldsymbol{G},i}(\boldsymbol{k}) \langle \boldsymbol{G}' + \boldsymbol{k}|\boldsymbol{G} + \boldsymbol{k}\rangle = \varepsilon_i \sum_{\boldsymbol{G}} \delta_{\boldsymbol{G}',\boldsymbol{G}} c_{\boldsymbol{G},i}(\boldsymbol{k}) = \varepsilon_i c_{\boldsymbol{G}',i}(\boldsymbol{k}). \tag{C.6}$$

Therefore, by substituting $\boldsymbol{G}$ and $\boldsymbol{G}'$ by $\boldsymbol{G}_m$ and $\boldsymbol{G}_{m'}$, where $m, m' \in [1, n_{\boldsymbol{G}}]$ are integers, we see that:

$$\sum_{m} H_{m'm}(\boldsymbol{k}) c_{m,i}(\boldsymbol{k}) = \varepsilon_i c_{m',i}(\boldsymbol{k}), \tag{C.7}$$

where $n_{\boldsymbol{G}}$ is the number of wave vectors needed to calculate the charge density, and:

$$H_{m'm} = \frac{\hbar^2 |\boldsymbol{G}_m + \boldsymbol{k}|^2}{2m_e} \delta_{m,m'} + V_{\text{eff}}(\boldsymbol{G}_m - \boldsymbol{G}_{m'}). \tag{C.8}$$

Now by solving Eq. C.7, one can obtain the coefficient $c_{m,i}(\boldsymbol{k})$.

C.2 RMS Data Sets

For all tables in this section, the unit of ρ is Å^{-3} and the unit of RMS is Å.

C.2.1 Na

BCC

64		80		96		112		140	
$\rho(\times 10^{-5})$	RMS	$\rho(\times 10^{-5})$	RMS	$\rho(\times 10^{-5})$	RMS	$\rho(\times 10^{-5})$	RMS	$\rho(\times 10^{-5})$	RMS
1.4136	0.1427	1.1367	0.1673	0.8323	0.1078	0.4731	0.1326	0.4434	0.1419
1.4073	0.1144	1.1235	0.1053	0.8250	0.1013	0.4655	0.1254	0.4284	0.1049
1.4057	0.1091	1.1192	0.0788	0.8177	0.0795	0.4582	0.1150	0.4209	0.1052
1.4043	0.0948	1.1159	0.0648	0.8103	0.0608	0.4507	0.0779	0.4179	0.0750
1.3999	0.1028	1.1147	0.0727	0.8036	0.599	0.4462	0.0784	0.4164	0.0745
1.3984	0.0884	1.1118	0.0898	0.8034	0.580	0.4447	0.0619	0.4149	0.0854
1.3969	0.0832	1.1091	0.0806	0.8033	0.591	0.4432	0.0533	0.4134	0.0604
1.3955	0.0729	1.1088	0.0774	0.8030	0.603	0.4430	0.0539	0.4119	0.0665
1.3940	0.0656	1.1087	0.0756	0.8028	0.602	0.4417	.0642	0.4104	0.0625
1.3932	0.0668	1.1081	0.0756	0.7986	0.637	0.4409	0.0665	0.4089	0.0682
1.3918	0.0590	1.1080	0.0750	0.7956	0.695	0.4387	0.0701	0.3985	0.0954
1.3903	0.0578	1.1077	0.0762	0.7927	0.837	0.4357	0.0801	0.3910	0.0832
1.3888	0.0541	1.1073	0.0811	0.7883	0.0949	0.4282	0.1101	0.3835	0.1541
1.3874	0.0593	1.1044	0.0848	0.7810	0.1274	0.4132	0.1720		
1.3859	0.0714	1.0971	0.1190	0.7736	0.1571				
1.3858	0.0781	1.0896	0.1177	0.7663	0.1774				
1.3830	0.0868	1.0830	0.1819						
1.3815	0.0952								
1.3800	0.1193								
1.3719	0.1087								

HCP

168		192	
$\rho(\times 10^{-7})$	RMS	$\rho(\times 10^{-7})$	RMS
3.9508	0.2022	3.7429	0.1976
3.7429	0.1955	3.5350	0.1884
3.5600	0.1896	3.3604	0.1855
3.3271	0.1810	3.2772	0.1795
3.2440	0.1785	3.2357	0.1766
3.1608	0.1766	3.1858	0.1740
3.0777	0.1751	3.1026	0.1723
2.9945	0.1774	3.0195	0.1777
2.9197	0.1797	2.8365	0.1811
2.8282	0.1808	2.6702	0.1850
2.7450	0.1834	2.5039	0.1888
2.6369	0.1855	2.3459	0.1853
2.3043	0.1903	2.1629	0.1879
1.9717	0.1893	1.9966	0.1870

C.2.2 NiTi

BCC

56		72		112		120	
$\rho(\times 10^{-6})$	RMS	$\rho(\times 10^{-6})$	RMS	$\rho(\times 10^{-6})$	RMS	$\rho(\times 10^{-6})$	RMS
6.8458	0.0611	6.8386	0.0610	6.9648	0.0612	6.5475	0.0588
6.3318	0.0561	6.4036	0.0592	6.5694	0.0560	6.1518	0.0594
5.8969	0.5420	6.0081	0.0586	6.1739	0.0534	5.7502	0.0526
5.5016	0.0556	5.5533	0.0502	5.7627	0.0550	5.3545	0.0513
5.1062	0.0494	5.1578	0.0456	5.3929	0.0483	4.9627	0.0367
4.7109	0.0434	4.7228	0.0439	4.9896	0.0473	4.5473	0.0409
4.3155	0.0378	4.3273	0.0427	4.5942	0.0434	4.1710	0.0411
3.9004	0.0354	3.9318	0.0401	4.1987	0.0360	3.7836	0.0335
3.5051	0.0364	3.4968	0.0336	3.7875	0.0350	3.3681	0.0325
3.1097	0.0344	3.1013	0.0303	3.3980	0.0340	2.9843	0.0321
2.7144	0.0313	2.7058	0.0294	3.0065	0.0308	2.5886	0.0265
2.3191	0.0271	2.3103	0.0319	2.6249	0.0262	2.1929	0.0210
1.9039	0.0247	1.9149	0.0241	2.2295	0.0257	1.7972	0.0231
1.5086	0.0220	1.5194	0.0199	1.8143	0.0239	1.3996	0.0210
1.1330	0.0225	1.1239	0.0221	1.4346	0.0224	0.9980	0.0212
0.7377	0.0227	0.7284	0.0212	1.0293	0.0224	0.6201	0.0199
0.2965	0.0202	0.3345	0.0174	0.6339	0.0174	0.2185	0.0189
0	0.0197	0	0.0165	0.2385	0.0171	0	0.0169
				0	0.0171		

C.2.3 Zr

BCC

96		112		120	
$\rho(\times 10^{-5})$	RMS	$\rho(\times 10^{-5})$	RMS	$\rho(\times 10^{-5})$	RMS
3.2439	0.1625	3.2258	0.1526	3.2220	0.1584
3.0070	0.1384	3.1322	0.1486	3.1501	0.1472
2.8902	0.1494	3.0821	0.1458	3.0773	0.1605
2.7962	0.1483	3.0378	0.1456	3.0058	0.1367
2.7723	0.1460	2.9889	0.1504	2.9082	0.1361
2.6552	0.1447	2.8920	0.1346	2.7415	0.1400
2.5348	0.1511	2.8013	0.1464	2.6700	0.1391
2.3014	0.1349	2.6567	0.1438	2.5777	0.1390
2.0613	0.1188	2.5592	0.1518	2.3392	0.1417
1.8243	0.1281	2.3231	0.1331	2.1000	0.1212
1.5878	0.1137	2.0842	0.1210	1.8606	0.1244
1.3469	0.1075	1.8520	0.1178	1.6253	0.1232
1.1139	0.0874	1.6115	0.1180	1.3862	0.1111
0.8742	0.0763	1.3765	0.1035	1.1508	0.0956
0.6301	0.0664	1.1338	0.0900	0.9123	0.0795
0.4012	0.0773	0.9004	0.0733	0.8145	0.0743
0.0319	0.0848	0.8006	0.0724	0.7679	0.0752
		0.7566	0.7180	0.7182	0.0763
		0.7041	0.0675	0.6755	0.0685
		0.6608	0.0742	0.6016	0.0744
		0.4214	0.0769	0.4560	0.0751
		0.0355	0.0844	0.3588	0.0744
				0.1679	0.0783

HCP

120		140	
$\rho(\times 10^{-5})$	RMS	$\rho(\times 10^{-5})$	RMS
0.7079	0.4727	0.7079	0.4751
0.7061	0.4681	0.7046	0.4768
0.7031	0.4660	0.7030	0.4777
0.6999	0.4650	0.7009	0.4782
0.6973	0.4606	0.6993	0.4746
0.6923	0.4630	0.6977	0.4680
0.6894	0.4677	0.6956	0.4707
0.6868	0.4762	0.6940	0.4638
0.6844	0.4774	0.6924	0.4629
0.6788	0.4707	0.6903	0.4660
0.6753	0.4765	0.6887	0.4691
0.6732	0.4882	0.6868	0.4738
0.6644	0.4880	0.6849	0.4762
0.6507	0.5043	0.6794	0.4816
0.6454	0.5013	0.6741	0.4883
0.6400	0.4926	0.6687	0.4822
0.6344	0.5032	0.6634	0.4902
0.6288	0.5033	0.6578	0.4890
0.6237	0.4961	0.6520	0.4877
0.6181	0.5007	0.6407	0.4957
0.6128	0.5117	0.6300	0.5009
0.5975	0.5332	0.6191	0.5126

Hexagonal

108		128	
$\rho(\times 10^{-8})$	RMS	$\rho(\times 10^{-8})$	RMS
9.2615	0.4225	9.1556	0.4279
8.4144	0.4359	8.0967	0.4401
7.1437	0.4541	7.1437	0.4498
6.1643	0.4244	6.1643	0.4615
5.3171	0.4069	5.7407	0.4389
5.0260	0.3989	5.3436	0.4168
4.7612	0.3906	4.9730	0.3942
4.6024	0.3887	4.4700	0.3731
4.4436	0.3831	4.0994	0.3843
4.3112	0.3856	3.8876	0.3876
4.0465	0.4009	3.7288	0.4037
3.7023	0.4081	3.0935	0.4247
2.9611	0.4263	2.1934	0.4445
2.2993	0.4331	1.3728	0.4472
0	0.4553	0	0.4573

C.2.4 Cu

FCC

84		100	
$\rho(\times 10^{-7})$	RMS	$\rho(\times 10^{-7})$	RMS
10.6114	0.1186	7.2043	0.1197
8.5451	0.1134	6.5740	0.1115
7.8772	0.1232	5.9175	0.0939
7.3554	0.1158	5.9175	0.1036
6.6457	0.1071	5.7088	0.1071
6.1406	0.0967	5.6555	0.1090
5.9319	0.0962	5.4468	0.1059
5.5604	0.1104	5.2829	0.1107
5.1430	0.1101	5.1128	0.1105
4.9760	0.1143	5.0137	0.1038
4.7673	0.1153	4.9041	0.1053
4.3707	0.1153	4.6954	0.1121
4.0785	0.1158	4.5075	0.1143
3.6297	0.1125	4.2988	0.1093

BCC

84		100	
$\rho(\times 10^{-6})$	RMS	$\rho(\times 10^{-6})$	RMS
4.9240	0.0945	4.9012	0.1187
4.4917	0.1029	4.4919	0.1372
4.1550	0.0932	4.0950	0.0822
4.0719	0.0909	4.0513	0.0915
3.9457	0.0911	4.0139	0.0954
3.8634	0.0867	3.9744	0.0912
3.7200	0.0862	3.8788	0.0905
3.6202	0.0808	3.8019	0.0884
3.5101	0.0960	3.6648	0.0809
3.4020	0.0904	3.5338	0.0932
3.2856	0.0644	3.4528	0.0897
3.1485	0.0806	3.3721	0.0941
3.0238	0.0763	3.3306	0.0901
2.8492	0.1022	3.2641	0.0692
		3.1281	0.0783
		3.0047	0.0797
		2.8280	0.1066

C.2.5 Al

FCC

60		80		100	
$\rho(\times 10^{-5})$	RMS	$\rho(\times 10^{-5})$	RMS	$\rho(\times 10^{-5})$	RMS
5.0941	0.2327	5.1130	0.2408	4.9801	0.1635
4.9408	0.1872	4.9611	0.2159	4.8268	0.1787
4.7891	0.1933	4.8101	0.2132	4.6755	0.1420
4.6373	0.1346	4.6584	0.1754	4.5230	0.1098
4.4855	0.1306	4.5066	0.1196	4.4775	0.0914
4.3337	0.0828	4.3548	0.1045	4.4331	0.0987
4.2882	0.0889	4.1552	0.0750	4.3727	0.0753
4.2730	0.0829	4.1383	0.0767	4.3246	0.0758
4.2305	0.0896	4.1264	0.0848	4.2791	0.0837
4.2154	0.0890	4.1006	0.0832	4.2194	0.0846
4.2002	0.0825	4.0854	0.0793	4.0646	0.0854
4.1902	0.0834	4.0793	0.1022	3.9129	0.0930
4.1820	0.0840	4.0701	0.1029	3.7573	0.1014
4.1637	0.0854	4.0687	0.1020	3.6075	0.1222
4.1516	0.0843	4.0516	0.1016	3.4525	0.1448
4.0302	0.0839	3.6009	0.1264	3.3053	0.1560
3.8807	0.0944				
3.7297	0.0969				
3.5779	0.1169				

BCC

72		100		112	
$\rho(\times 10^{-6})$	RMS	$\rho(\times 10^{-6})$	RMS	$\rho(\times 10^{-6})$	RMS
5.9189	0.2124	5.4252	0.3221	5.6718	0.2322
2.9704	0.1586	4.2514	0.2197	5.0776	0.1925
2.7497	0.1437	3.6425	0.1708	4.1967	0.1828
2.4555	0.1303	3.0424	0.1752	3.3142	0.1647
2.1612	0.1378	2.7600	0.1509	2.7106	0.1452
1.8743	0.1257	2.4526	0.1500	2.1193	0.1443
1.5802	0.1285	2.1584	0.1376	1.5237	0.1242
1.2713	0.1269	1.8554	0.1278	0.9354	0.1297
0.9858	0.1278	1.5568	0.1247	0.6412	0.1276
0.6916	0.1280	1.2553	0.1304	0.3409	0.1195
0.4120	0.1314	0.9758	0.1295	0	0.1200
0.1833	0.1237	0.6699	0.1314		
0	0.1278	0.3757	0.1230		
		0	0.1218		

D. Topology and Gauss–Bonnet Theorem

D.1 Mathematical Definition of Topology

An action of tearing, piecing and penetrating that breaks the continuity will make two connected points disconnected. The continuity of a deformation can be described using *neighbourhood*. A *neighbourhood* of $p \in S$ is defined as an **open subset** $A \subseteq S$ if there exist an **open set** U, such that $x \in U \subset A$ [124]. For example, given a set $S = \{a,b,c\}$ and define its topology[1] as $\{S, \emptyset, \{a\}, \{b\}, \{a,b\}\}$. Therefore, $\{a,c\}$ (or $\{b,c\}$, $\{c\}$) as a complement of $\{b\}$ (or $\{a\}$, $\{a,b\}$) thus is closed. Then in this case, the only neighbourhood of c is S, and the neighbourhoods of a can be $S, \{a,b\}$ and $\{a\}$. Now, the continuity in interpreted as following. Take a surface $S : \mathbb{R}^2 \to \mathbb{R}^3$ as an example. Let us consider a point $A \in S$ and the collection of all its neighbourhoods defined on S, $\{U_A^i\}|_S$. Define a map $\varphi : S \to S'$. Hence, we have $\forall A \in S$, $\varphi(A) \in S'$. If this deformation from S to S' is continuous, then we see:

$$\forall A \in S,\ \varphi(A) \in U_{\varphi(A)} \Rightarrow \exists U_A,\ A \in U_A,\ \varphi(U_A) \subset U_{\varphi(A)}. \tag{D.1}$$

In the language of neighbourhood, we have the topology defined as [230]:

1. $\forall A \in S$, $\exists U_A \subset S$, every point on S has neighbourhoods defined on S;
2. $\forall A \in S$, there exists a set of points within which we can move around A without leaving S, which means:
   a. $\forall U_A^1, U_A^2 \in \{U_A^i\}|_S, U_A^1 \cap U_A^2 \in \{U_A^i\}|_S$;
   b. $\forall U_A^1 \in \{U_A^i\}|_S, U_A^1 \subseteq U \Rightarrow U \in \{U_A^i\}|_S$;
   c. $\forall A \in S$, $\exists U_A^2 \subseteq U_A^1 \in \{U_A^i\}|_S$, such that $B \in U_A^2$, $U_A^2 \in \{U_B^i\}|_S$. This statements means A can find its path to B in its neighbourhood without leaving the surface S.

More general, the *topology* and *topological space* are defined from the perspective of the open set [124] according to the neighbourhood statements. Give any set S and the collection of its all subsets $\{U_i\}$, if $\{U_i\}$ satisfies [124]:

[1] See the definition of topology below.

1. $S, \emptyset \in \{U_i\}$;
2. Any $U_1, U_2, \cdots, U_n \in \{U_i\}$, $\cap_{q=1}^{n} U_q \in \{U_i\}$, $n \in \mathbb{Z}^+$ is a finite number;
3. Any $U_1, U_2, \cdots, U_m \in \{U_i\}$, $m \in \mathbb{Z}^+$ can be an infinite number, $\cup_{q=1}^{m} U_q \in \{U_i\}$;

$\{U_i\}$ is said to *give a topology to* S, and the pair $(S, \{U_i\})$ is said to be a *topological space*.

D.2 Embedding and Immersion

Roughly speaking, an *immersion* $\varphi : S \to W$ can be understood as a **bijective mapping between tangent spaces on** S **and** W, while an *embedding* requires further φ to be a **bijective mapping between points on** S **and** W. Take a 2D rope as an example. As shown in Fig. D.1, **tangent vectors** at any point on A [Fig. D.1(a)] can be mapped to **tangent vectors** at any point on the eight knot [Fig. D.1(b)]. At the cross-point O, since the tangent vectors pointing to two different directions, the map will be one-to-one and all images on the eight-knot can find a preimage on the circle. However, if we want to map **points** on a circle to **points** on an eight-knot [Fig. D.1(b)], the cross-point O will has two preimages, which means this map is not one-to-one. Therefore, a map from a circle to an eight-knot is an *immersion*, not an embedding. Hence, roughly speaking, if $\varphi : S \to W$ is an *embedding*, then **no self-intersection can be found on either** S **or** W.

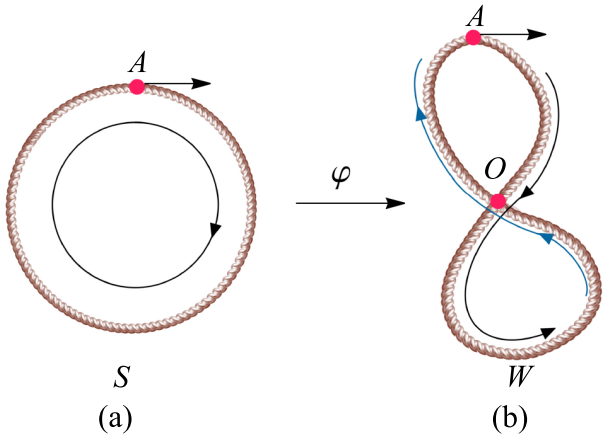


Figure D.1 Mapping from (a) a circle to (b) an eight-knot in a 2D plane.

D.3 Gaussian Curvature as Area Ratio

Consider a surface $S = \boldsymbol{x}(u,v)$, $(u,v) \in U \subset \mathbb{R}^2, S \subset \mathbb{R}^3$. Let us choose an arbitrary $p \in S$ and then take a unit normal vector $\boldsymbol{n}_p$ at this point to S. Given $\langle \boldsymbol{n}_p, \boldsymbol{n}_p \rangle = 1$, we have:

$$\langle \partial_u \boldsymbol{n}_p, \boldsymbol{n}_p \rangle = \langle \partial_v \boldsymbol{n}_p, \boldsymbol{n}_p \rangle = 0, \tag{D.2}$$

which means $\partial_u \boldsymbol{n}_p$ and $\partial_v \boldsymbol{n}_p$ are in the tangent plane $T_p(S)$ expanded by $\partial_u \boldsymbol{x}$, $\partial_v \boldsymbol{x}$ at p to S. In the basis $\{\partial_u \boldsymbol{x}, \partial_v \boldsymbol{x}\}$, we have:

$$\begin{pmatrix} \partial_u \boldsymbol{n}_p \\ \partial_v \boldsymbol{n}_p \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \partial_u \boldsymbol{x} \\ \partial_v \boldsymbol{x} \end{pmatrix}. \tag{D.3}$$

Hence, it can be seen, for instance, $\langle \partial_u \boldsymbol{n}, \partial_u \boldsymbol{x} \rangle = a_{11} \langle \partial_u \boldsymbol{x}, \partial_u \boldsymbol{x} \rangle + a_{12} \langle \partial_v \boldsymbol{x}, \partial_u \boldsymbol{x} \rangle$. Recalling the definition of the first (Eq. 3.38) and second (Eq. 3.42) fundamental forms, we have:

$$-\overline{\Pi}^2_p = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \Pi^1_p \Rightarrow \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \Phi_p = -\overline{\Pi}^2_p (\Pi^1_p)^{-1} \tag{D.4}$$

Taking determinant of Φ_p, recalling Eq. 3.45, we have:

$$\det(\Phi_p) = -K, \tag{D.5}$$

where K is the Gaussian curvature. Taking a Gauss map $N : U \subset S \to S^2$, with $\partial U \subset S$ being the boundary of U. Then the area before the map is:

$$A(U) = \int_U |\boldsymbol{x}_u \times \boldsymbol{x}_v| \mathrm{d}u \, \mathrm{d}v, \tag{D.6}$$

and the area of the image of U is:

$$A'(U) = \int_U |\partial_u \boldsymbol{N}_u \times \partial_v \boldsymbol{N}| \mathrm{d}u \, \mathrm{d}v \tag{D.7}$$

Referring to Eq. D.3 and Eq. D.5, we have:

$$|\partial_u \boldsymbol{N} \times \partial_v \boldsymbol{N}| = |(a_{11}\partial_u \boldsymbol{x} + a_{12}\partial_v \boldsymbol{x}) \times (a_{21}\partial_u \boldsymbol{x} + a_{22}\partial_v \boldsymbol{x})| = |K||\partial_u \boldsymbol{x} \times \partial_v \boldsymbol{x}|. \tag{D.8}$$

Taking the ratio then let U approach to a point $p \in U$:

$$\lim_{U \to p} \frac{A'(U)}{A(U)} = \lim_{U \to p} \frac{\int_U K |\boldsymbol{x}_u \times \boldsymbol{x}_v| \mathrm{d}u \, \mathrm{d}v}{\int_U |\boldsymbol{x}_u \times \boldsymbol{x}_v| \mathrm{d}u \, \mathrm{d}v} = K|_p. \tag{D.9}$$

Here we use the *mean value theorem for double integrals*, which states as following. Let $f(u,v)$ be a continuous function defined on a closed, connected region U, then there exist a pair of value $(u_0, v_0) \in U$, such that $\int_U f(u,v)\mathrm{d}u\,\mathrm{d}v = Cf(u_0,v_0)$, where C is the positive area of U. We remove the absolute sign here because Gaussian curvature is a map that takes $S \to \mathbb{R}$ and is continuous in our case, which means when U is small enough, K does not change its sign [47].

D.4 Spherical Triangle and Triangulation of Polygons

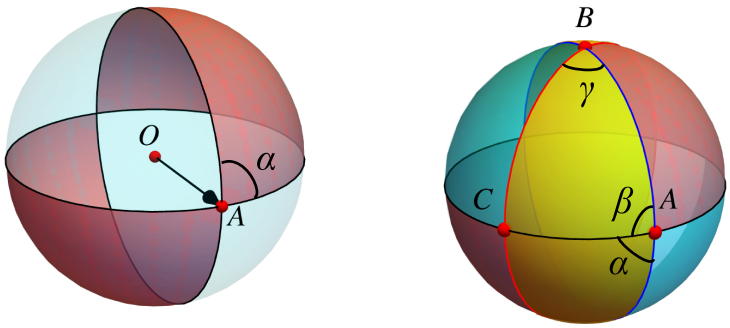
To calculate the area of a spherical triangle, we can begin with considering the *spherical lune*. As shown in Fig. D.2(a), the area coloured red that is **bounded by two great circles** is a pair of spherical lunes. Let the angle between these two circles be α and the radius of the sphere be R, then the area of the pair spherical lunes is:

$$A = \frac{4\pi R^2(2\alpha)}{2\pi} = 4\alpha R^2 \tag{D.10}$$

Let us cover the sphere with three pairs of spherical lunes and let the angle between their corresponding pair great circles be α, β and γ, respectively, we now see that the total area of these spherical lunes is:

$$A_T = 4R^2(\alpha+\beta+\gamma) \tag{D.11}$$

In comparison with the area of the sphere $4\pi R^2$, we see that A_T counts a spherical triangle ABC [Fig. D.2(b)] for extra four times thus becomes greater than $4\pi R^2$. Hence, the area of



(a) (b)

Figure D.2 Sketch of (a) a pair of spherical lunes on a sphere. O is the spherical centre and A is a point on this sphere. (b) Three pairs of spherical lunes covering a sphere, lunes belong to the same pair are coloured in the colour. A, B, C are points on the sphere.

the spherical triangle ABC is:

$$L = \frac{4R^2}{4}(\alpha+\beta+\gamma-\pi) = R^2(\alpha+\beta+\gamma-\pi) \tag{D.12}$$

Then, let us consider a triangulation of a polygon. Let this polygon have n_e edges. For a triangle having 3 edges, the triangulation gives 1 triangle.

Next, supposing when $n_e < m, m \in \mathbb{Z}^+$, the triangulation of this polygon requires $n_e - 2$ triangles. Now, cut an m-edge polygon along one of its **diagonals**. *Diagonal* here is defined as **a line linking two non-adjacent points of a polygon within it**. Let the obtained two new polygons have n, p edges respectively. It can be seen that:

$$n + p = m + 2 \tag{D.13}$$

Given the triangulation of these two new polygons requires $n-2$ and $p-2$ triangles, respectively, we can see that for the m-polygon, the number of triangles N_T is given by:

$$N_T = (n-2)+(p-2) = n+p-4 = m-2 \tag{D.14}$$

D.5 Euler Characteristic of a Planar Graph

A *graph* is defined as **a set of vertices and edges**, and a *region* in a graph is defined as **the area bounded by edges**. For example, the area coloured red in Fig. D.3(a) is a region and there are in total 6 regions, labelled by r_i, $i = 1,2,3,4,5,6$. When there are no *cycles* in a graph, such a graph is called a *tree graph* [Fig. D.3(b)], the number of region now is one [labelled by r_1 in Fig. D.3(b)]. A *cycle* is defined as **a path on which vertices are arranged in a cyclic sequence** [231]. Hence, for the tree graph, let the number of vertices, edges and

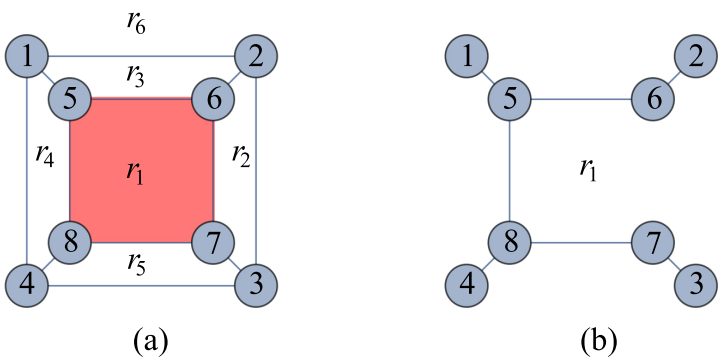


Figure D.3 Sketch of (a) a planar graph, (b) a tree graph.

regions be n_v, n_e and n_f. Then:

$$n_f + n_v - n_e = 2, \tag{D.15}$$

For a planar graph [Fig. D.3(a)], suppose when $n_e \geq m$, the relation $n_f + n_v - n_e = 2$ does not hold. Consider a planar graph with m edges, n_f regions and n_e vertices. Then, if we open a cycle by deleting an edge, we have:

$$n'_v = n_v, \qquad n'_f = n_f - 1, \qquad n'_e = n_e - 1 = m - 1. \tag{D.16}$$

Since:

$$n'_v + n'_f - n'_e = n_v + n_f - 1 - (n_e - 1) = 2, \tag{D.17}$$

we see:

$$n_v + n_f - n_e = 2 \tag{D.18}$$

This conflicts with the assumption. Hence for planar graphs having n_e, n_v, n_f edges, vertices and regions, we always have:

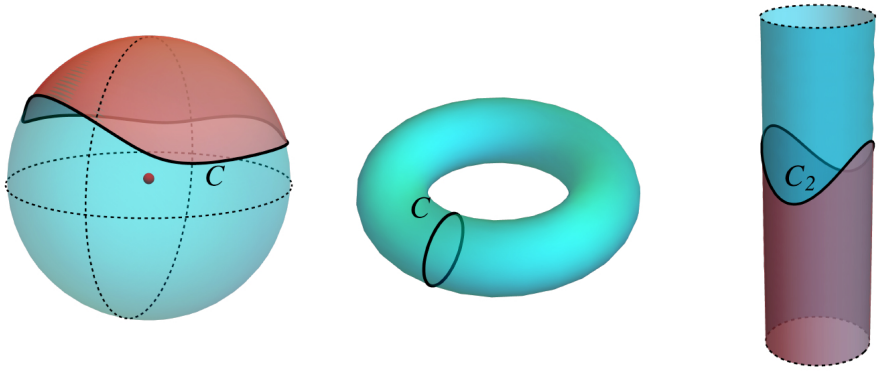
$$n_f + n_v - n_e = 2. \tag{D.19}$$

D.6 Euler Characteristic and Genus

Now, we will gain more insight into the definition of the genus. Genus on a connected surface can be roughly defined as the number of holes on a surface. Precisely, it is defined as **the maximum number of disjoint loops can be drawn on a surface without disconnecting it** [232]. For instance, there is no way for us to draw a loop without disconnecting a sphere [Fig. D.4(a)]. But for a torus, as shown in Fig. D.4(b,c), only the second loop C_2 can disconnect it. Hence the genus of a sphere is 0 while of a torus is 1.

Then let us come back to our Euler characteristic argument. Consider a closed and connected surface of genus g_N. Taking a triangulation on this surface, such that the number of vertices, faces and edges are n_v, n_f and n_e, with [47] (Euler's formula):

$$n_v + n_f - n_e = \chi \tag{D.20}$$



(a) (b) (c)

Figure D.4 Sketch of loops can be drawn on (a) a sphere and (b,c) a torus. There is no way for us to draw a loop on a sphere without disconnecting it. As shown in (a), the area coloured red inside the loop C is not connected with the rest part of the sphere. On a torus, we can first draw a loop C, as shown in (b). Opening the torus along this loop will give us a cylinder, which is a connected surface. But we cannot draw a loop on a cylinder without disconnecting it, as given in (c), the area coloured red is separated from the rest of the cylinder by C_2.

Next, drawing loops around each hole, such that by cutting the surface along these loops, we will not disconnect this surface. According to our definition, we can draw g_N loops. Also, by cutting the surface along these loops (Fig. D.5), we will have additional $2g_N$ faces. As for number of edges and vertices, since the cuts are around loops, the number of additional edges should equal to the number of additional vertices. Labelling the number of vertices, edges and faces of the new surface by n'_v, n'_e and n'_f, we have this no-hole surface satisfying (Eq. D.19):

$$n'_v - n'_e + n'_f = 2, \tag{D.21}$$

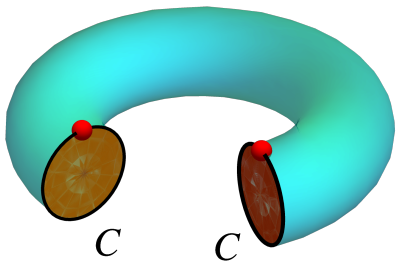


Figure D.5 Sketch of opening a hole (torus) around a loop C. When we open this torus, 2 addition edges (all labelled by C) together with the area enclosed by them that contribute to 2 additional faces will occur. Since C is a loop, then there will be 2 additional vertices (red spheres) in this case.

also, we have that:

$$n'_v - n'_e + n'_f = (n_v + 2g_N) - (n_e + 2g_N) + (n_f + 2g_N) = n_v - n_e + n_f + 2g_N. \tag{D.22}$$

Therefore, we see that:

$$n_v - n_e + n_f + 2g_N = 2 \Rightarrow n_v - n_e + n_f = 2 - 2g_N = \chi, \tag{D.23}$$

which is the relationship between the Euler characteristic and the genus on a closed surface.

D.7 Derivation of Gauss-Bonnet Theorem

Now let us switch $A(U) \to \mathrm{d}S$ in Eq. 3.46 since when U approaches $p \in S$, $A(U)$ will go to $\mathrm{d}S$, $A'(U)$ will go to $\mathrm{d}T$. On the left hand side, given the area of a spherical triangle on a unit sphere:

$$L = \sum_i \varphi_i - \pi, \qquad i = 1,2,3, \tag{D.24}$$

where φ_i are interior angles of this triangle (details are in Appendix D.4). Using Eq. 3.47, we have the original form of the Gauss-Bonnet theorem:

$$\sum_i \varphi_i - \pi = \int_T K \mathrm{d}S, \tag{D.25}$$

which states that **the excess over π, π is the sum of interior angles of a triangle on a plane, by the sum of interior angles equal to the integral of Gaussian curvature over this *geodesic triangle***. In other words, Gaussian curvature reveals the extent to which a vector is rotated after being parallel transported around a loop on a surface. As shown in Fig. D.6, if we parallel transport a vector $\boldsymbol{v}$ along the path $A \to B \to C \to D \to A$, $\boldsymbol{v}$ is flipped. **On a sphere the spherical triangle is a geodesic triangle**. More descriptions about geodesic can be found in Ref. [47]. Continuing on the unit sphere, but this time, we will project onto this sphere a polyhedron having n_f faces, n_v vertices and n_e edges placed at its centre (Fig. D.7). Take a *triangulation* of this projected image. A **triangulation** of a surface is **covering this surface by a finite collection of triangles such that**: 1) every point on this surface is in a triangle and 2) the intersection of two triangles is non-empty, if and only if the intersection is a common vertex or a common edge of these two triangles [47]. A *triangle* here is defined as **a region with no hole or singularity in it and has only three vertices with internal angles being $\alpha, \beta, \gamma \neq 0$** [47]. Since an n-edge polygon can be triangulated into $n-2$ triangles (Appendix D.4), we have the area of this image (equal to the surface area of this unit sphere)

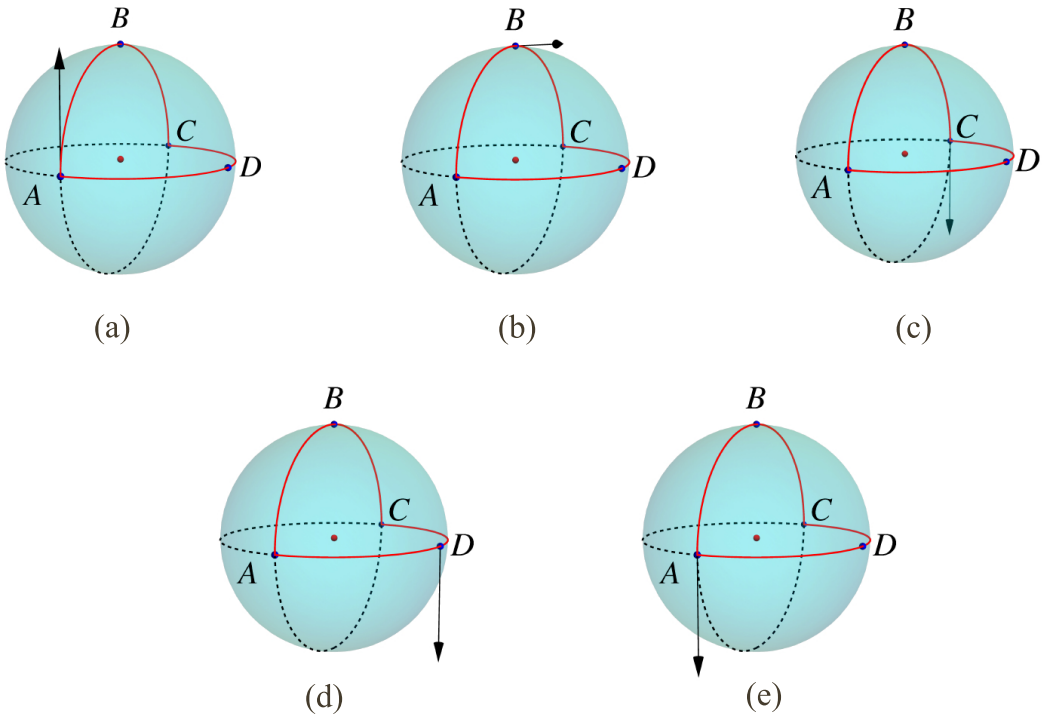


Figure D.6 Parallel transportation of a vector along a path $A \to B \to C \to D$ marked red on a sphere.

be:

$$A = \sum_{i=1}^{n_f} \sum_{j=1}^{e_i-2} (\alpha_j + \beta_j + \gamma_j - \pi) r^2 = 4\pi r^2 \tag{D.26}$$

Here we use the result in Eq. D.24 and e_i denotes the number of edges of the ith surface. Given we are talking about a unit sphere, we set $r = 1$. Now we have:

$$\sum_{i=1}^{n_f} \sum_{j=1}^{e_i-2} \pi = \sum_{i=1}^{n_f} (e_i - 2)\pi = 2n_e\pi - 2n_f\pi, \tag{D.27}$$

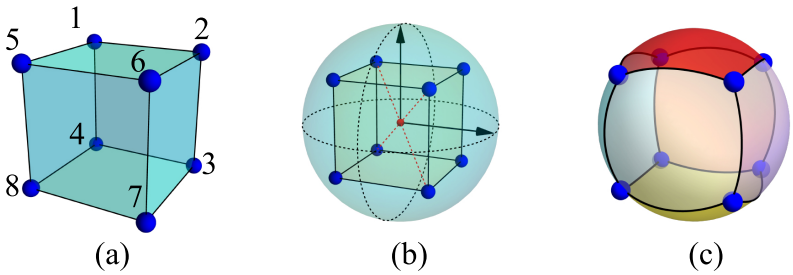


Figure D.7 Example of projecting a polyhedron onto a sphere. (a) A cube. Numbers label vertices as given in Fig. D.8(a). (b) Placing the cube at the centre of a sphere. (c) Projecting the cube onto the sphere. Blue spheres are vertices of the cube projected onto this sphere. Different faces are coloured differently.

because in the sum $\sum_{i=1}^{n_f} e_i$ every edge of each face is counted twice, we have it equal to $2n_e$. Also, since on a sphere, the sum of angles sharing the same vertex can be approximated to 2π, we have:

$$\sum_{i=1}^{n_f}\sum_{j=1}^{e_i-2}(\alpha_j+\beta_j+\gamma_j)\approx 2\pi n_v. \tag{D.28}$$

Now we see that:

$$2\pi(n_f+n_v-n_e)=4\pi \Rightarrow n_f+n_v-n_e=2, \tag{D.29}$$

The value given by $n_f+n_v-n_e$ is called the *Euler characteristic* [47] and is a **topological invariant**. A surface that is topologically equivalent to a sphere should have $\chi=2$. To see the topological invariance of the Euler characteristic, image we project this polyhedron onto a plane with no edge-crossing [Fig. D.8(a)], such that the image of this polyhedron can form a **planar graph**. A *graph* is a set of edges and points that can model relations between different objects [231]. A *planar graph* is **a graph without edge-crossings** [231]. The Euler

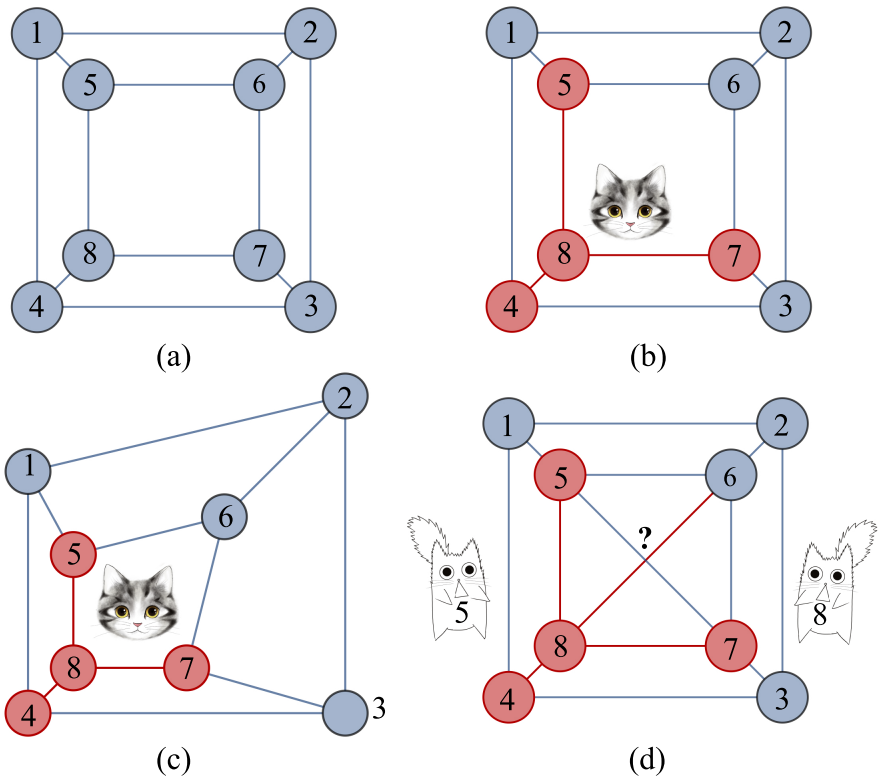


Figure D.8 (a) Planar graph formed by projecting Fig. D.7(c) onto a plane. Numbers label vertices. (b) A cat lives at vertex 8 and its neighbours vertices 4, 5 and 7. Vertices 4, 5, 7, 8 and edges linking them are coloured red. (c) Deformed planar graph without edge crossing. (d) New graph formed by adding edges connecting vertices 5 and 7, 8 and 6. Cats living at vertices 8 and 5 go to vertices 6 and 7 will meet at the centre labeled by the question mark.

characteristic of a planar graph is also 2 (derivation in Appendix D.5).

As shown in Fig. D.8, suppose that a 2D cat lives at a vertex, for instance, 8, in this plane, then his neighbours will live at vertices 4, 5 and 7. Then moving around any of these eight vertices around in this plane without causing edge-crossing or vertices-overlapping [Fig. D.8(d)], our cat will still have three neighbours living at vertices 4, 5 and 7 and the Euler characteristic remains 2. However, if we allow edge-crossings and add paths connecting $8 \leftrightarrow 6$ and $5 \leftrightarrow 7$, we now have eight vertices and fourteen edges. Suppose that the Euler characteristic formula still holds, then we will have eight faces. However, according to Fig. D.8(d), each edge is shared by two faces, which means eight faces require at least sixteen edges, and this contradicts to our condition. Therefore, a graph like Fig. D.8(d) cannot live in a plane. More intuitively, for our 2D cats, let one cat go from 5 to 7 and another one go from 8 to 6. Therefore, they are supposed to meet at the cross centre. However, what actually happens is that they arrive at the centre at the same time without seeing each other at all. That is because such a graph can only live on a torus but not a sphere (Fig. D.9). The action of adding paths changes the spatial relation between points such that Fig. D.8(d) cannot deform into either Fig. D.8(a,b,c) continuously (topologically inequivalent). And for Fig. D.8(d), the Euler characteristic is zero.

Therefore, we see **connected** graphs of the same Euler characteristic can deform into each other continuously, *viz.*, are topologically equivalent. A graph is *connected* if **each pair of vertices is connected by at least one edge in one graph** [231]. Hence, we call the Euler characteristic a topological invariant. Go back to Eq. D.25, Eq. D.26 and Eq. D.29, we have:

$$2\pi\chi = \int_S K \mathrm{d}S, \tag{D.30}$$

where the integral on the right-hand side is taken over the whole surface S, that is compact, connected and closed. Also, we can notice that the genus g_N of a sphere is zero while g_N of a torus is 1, both of them satisfies the relation [47] (a rough explanation can be found in

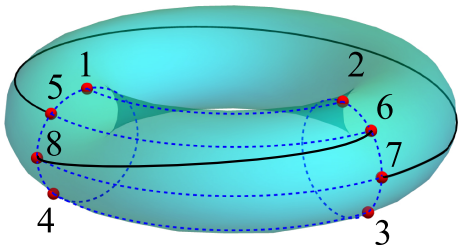


Figure D.9 Sketch of embedding Fig. D.8(d) on a torus. Vertices are labelled by numbers corresponding to those in Fig. D.8(d) represented by red spheres. Black solid lines are for edges linking vertices 5 and 7, 8 and 6. Blue dashed lines are for the rest of edges in Fig. D.8(d).

Appendix D.6):

$$2-2g_N=\chi. \tag{D.31}$$

Thus we have for a closed and connected surface, the *Gauss-Bonnet theorem* being expressed as [47]:

$$2-2g_N=\frac{1}{2\pi}\int\int_S K\mathrm{d}S. \tag{D.32}$$

D.8 Calculating Gaussian Curvature on Minimal Surfaces

D.8.1 P Surface

To calculate the Gaussian curvature K on the P surface, we first need to refer to the expression given in Sec. 3.3.5. For simplicity, we can take a rotation of the surface given by Eq. 3.61 by $\pi/4$, then find the Schwarz P surface expressed as [130]:

$$G_s=\frac{\theta_1(z,\tau)\theta_1(z-\frac{\tau}{2},\tau)}{\theta_1(z-\frac{1}{2},\tau)\theta_1(z-\frac{1}{2}-\frac{\tau}{2},\tau)} \tag{D.33}$$

where $\tau=0.781701\mathrm{i}$. Then, examining Eq. B.73, for $K=0$, we should either have $|G_s|=0$ or $|\mathrm{d}G_s/\mathrm{d}z|=0$. Given the point labelled by the pink sphere in Fig. 5.18(a) is at the centre of the surface, the coordinate of this point in $\mathbb{C}$ is $\tau/4$. Plugging $z=\tau/4$ into $G_s(z)$, we have $G_s(z)\neq 0$. Hence, we need to calculate the value of $|\mathrm{d}G_s/\mathrm{d}z|$. Because the θ function involves the summation of infinite terms, we can only calculate the K value numerically using:

$$f'(z)|_{z=z_0}=\lim_{r\to 0}\frac{f(z_0+r)-f(z_0-r)}{2r}. \tag{D.34}$$

Using the Mathematica code given below, and define G_s according to Eq. D.33, we see $|\mathrm{d}G_s/\mathrm{d}z|_{z=\tau/4}=0$:

```
Gs[z_] := θ1[z,τ] θ1[z-τ/2,τ]/θ1[z-1/2,τ] θ1[a-1/2-τ/2,τ]

gderive[z_?NumericQ] :=

N@With[z0 = SetPrecision[z,64], r = Power[10,-50], (Gs[z+r] - Gs[z-r])/2r]
```

In the oPa family expression, we have:

$$G_s(z) \equiv \sqrt{\frac{\theta_1(z-b_s,\tau)\theta_1(z+b_s,\tau)\theta_1(z-c_s,\tau)\theta_1(z+c_s,\tau)}{\theta_1(z-a_s,\tau)\theta_1(z+a_s,\tau)\theta_1(z-d_s,\tau)\theta_1(z+d_s,\tau)}}, \tag{D.35}$$

then the P surface has:

$$a_s = 0.123, \qquad b_s = 0.377, \qquad c_s = \frac{1}{2}(1-\tau) - a_s, \qquad d_s = \frac{1}{2}(1+\tau) - b_s, \tag{D.36}$$

and:

$$\tau = 0.893\mathrm{i}. \tag{D.37}$$

Using the code given above but define G_s according to Eq. D.35, we have the preimage of the flat point in $\mathbb{C}$ given by $(0,0)$, such that:

$$\left.\frac{\mathrm{d}G_s}{\mathrm{d}z}\right|_{z=0} = 0. \tag{D.38}$$

The oPb family expression is given by:

$$G_s(z) \equiv \sqrt{\frac{\theta_1(z-a_s,\tau)\theta_1(z+a_s,\tau)\theta_1(z-c_s,\tau)\theta_1(z+c_s,\tau)}{\theta_1(z-b_s,\tau)\theta_1(z+b_s,\tau)\theta_1(z-d_s,\tau)\theta_1(z+d_s,\tau)}}. \tag{D.39}$$

In this expression, to describe the P surface, we have:

$$a_s = 0.906, \qquad b_s = \frac{1}{2} - a, \qquad c_s = \tfrac{1}{2}(1-\tau) - a_s, \qquad d_s = \tfrac{1}{2}(1+\tau) - b_s, \qquad \tau = 0.800\mathrm{i}\,. \tag{D.40}$$

Using code given above but define G_s according to Eq. D.39, we see:

$$\left.\frac{\mathrm{d}G_s}{\mathrm{d}z}\right|_{z=0} = 0. \tag{D.41}$$

D.8.2 I-WP Surface

As for the I-WP surface, the Weierstrass representation is given as follow:

$$G_s(z) \equiv \exp(\mathrm{i}\frac{\pi}{2})\sqrt{\frac{\theta_1(z-a_s,\tau)^{\frac{3}{2}}\theta_1(z+b_s+\frac{\tau}{2},\tau)^{\frac{3}{2}}}{\theta_1(z+a_s,\tau)^{\frac{3}{2}}\theta_1(z-b_s+\frac{\tau}{2},\tau)^{\frac{3}{2}}}}, \tag{D.42}$$

where we have $a_s = 1/6$, $b_s = 1/3$, and the flat point locates at $(\frac{1}{6},0)$ in $\mathbb{C}$, with:

$$G_s(a_s) = 0 \tag{D.43}$$

D.8.3 F-RD Surface

The Weierstrass representation is:

$$G_s(z) \equiv \sqrt{\frac{\theta_1(z+a_s,\tau)\theta_1(z+b_s,\tau)^{\frac{3}{2}}\theta_1(z-c-\frac{\tau}{2},\tau)^{\frac{3}{2}}\theta_1(z-d-\frac{\tau}{2},\tau)}{\theta_1(z-a_s,\tau)\theta_1(z-b_s,\tau)^{\frac{3}{2}}\theta_1(z+c-\frac{\tau}{2},\tau)^{\frac{3}{2}}\theta_1(z+d-\frac{\tau}{2},\tau)}}, \tag{D.44}$$

where:

$$a_s = 0.108, \qquad a_s + d_s = \frac{1}{2}, \qquad c_s + b_s = \frac{1}{2}, \qquad 2a_s = 1 - 3b_s. \tag{D.45}$$

Use the code given in Sec. D.8.1 but define G_s according to Eq. D.44, we see:

$$\left|\frac{\mathrm{d}G_s}{\mathrm{d}z}\right|_{z=b_s} = 0, \qquad b_s = 0.261 \tag{D.46}$$

D.8.4 H′-T Surface

The Weierstrass representation is:

$$G_s(z) \equiv \sqrt{\frac{\theta_1(z-a_s,\tau)}{\theta_1(z+a_s,\tau)}}, \tag{D.47}$$

where:

$$a_s = \frac{1}{6}. \tag{D.48}$$

Use the code given in Sec. D.8.1 but define G_s according to Eq. D.47, we see for the hexagonal lattice:

$$\left|\frac{\mathrm{d}G_s}{\mathrm{d}z}\right|_{z=0} = 0\,. \tag{D.49}$$

And for the HCP lattice:

$$G_s(3a_s) = 0. \tag{D.50}$$

D.8.5 oPb Surface Describing B19′, 109-B19′ and R NiTi

The Weierstrass representation is given by:

$$\mathrm{d}h = 1, \qquad G_s(z) \equiv \sqrt{\frac{\theta_1(z-a_s,\tau)\theta_1(z+a_s,\tau)\theta_1(z-c_s,\tau)\theta_1(z+c_s,\tau)}{\theta_1(z-b_s,\tau)\theta_1(z+b_s,\tau)\theta_1(z-d_s,\tau)\theta_1(z+d_s,\tau)}}, \tag{D.51}$$

where:

$$a_s = 0.096, \qquad b_s = 0.198, \qquad c_s = \tfrac{1}{2}(1-\tau) - a_s\,, \qquad d_s = \tfrac{1}{2}(1+\tau) - b_s\,, \tag{D.52}$$

and:

$$\tau = 0.469\mathrm{i}\,. \tag{D.53}$$

The projection of a flat point is at $(0,0) \in \mathbb{C}$, and we see:

$$\left.\frac{\mathrm{d}G_s}{\mathrm{d}z}\right|_{z=0} = 0. \tag{D.54}$$

D.8.6 oPa Surface Describing B33 NiTi

The Weierstrass representation is given by:

$$G_s(z) \equiv \sqrt{\frac{\theta_1(z-b_s,\tau)\theta_1(z+b_s,\tau)\theta_1(z-c_s,\tau)\theta_1(z+c_s,\tau)}{\theta_1(z-a_s,\tau)\theta_1(z+a_s,\tau)\theta_1(z-d_s,\tau)\theta_1(z+d_s,\tau)}}, \tag{D.55}$$

where:

$$a_s = 0.153, \qquad b_s = 0.347\,, \qquad c_s = a_s + \tfrac{1}{2}\tau\,, \qquad d_s = b_s - \tfrac{1}{2}\tau\,, \tag{D.56}$$

with:

$$\tau = 2.235\mathrm{i}. \tag{D.57}$$

Using code given in Setion D.8.1 but define G_s according to Eq. D.55, we see at $(0,0) \in \mathbb{C}$:

$$\left.\frac{\mathrm{d}G_s}{\mathrm{d}z}\right|_{z=0} = 0. \tag{D.58}$$

E. Motion of Particles on A Manifold

This appendix is to provide more details about how the mathematical expression of the motion of a particle confined in a surface (Sec. 6.3.2), including more mathematical background related to differential geometry (Sec. E.1) and the generalised Hamiltonian method (Sec. E.2), as well as calculation details (Sec. E.3).

E.1 Metric Space and Christoffel Symbols

E.1.1 Christoffel Symbol and Connection

Let us consider a transformation between tow sets of coordinates $\{x^i\}$, $\{\bar{x}^i\}$ whose basis are $\{\boldsymbol{e}_i\}$ and $\{\boldsymbol{b}_i\}$, respectively, $i = 1,2,\cdots,n$. Then, these coordinates are related to each other by:

$$\bar{x}^i = \bar{x}^i(x^j), \qquad x^j = x^j(\bar{x}^i), \tag{E.1}$$

For instance, in Cartesian coordinate (x^1,x^2,x^3), a vector $\boldsymbol{v} = x^i\boldsymbol{e}_i, i = 1,2,3$, then in the spherical coordinate (r,θ,ϕ),

$$\begin{aligned} x^1 &= x^1(r,\theta,\phi) = r\sin\theta\cos\phi, \\ y^1 &= x^2(r,\theta,\phi) = r\sin\theta\sin\phi, \\ z^1 &= x^3(r,\theta,\phi) = r\cos\theta \end{aligned} \tag{E.2}$$

Hence the basis vectors can be calculated, according to the Jacobian matrix, as:

$$\begin{aligned} \boldsymbol{b}_i &= \frac{\partial x^j}{\partial \bar{x}^i}\boldsymbol{e}_j, \qquad \boldsymbol{b}^i = \frac{\partial \bar{x}^i}{\partial x^j}\boldsymbol{e}^j \\ \boldsymbol{e}_i &= \frac{\partial \bar{x}^j}{\partial x^i}\boldsymbol{b}_j, \qquad \boldsymbol{e}^i = \frac{\partial x^j}{\partial \bar{x}^i}\boldsymbol{b}^j \end{aligned} \tag{E.3}$$

Let $S = \boldsymbol{x}(u,v), (u,v) \in U \subset \mathbb{R}^2$, and S is a $2D$ subspace of the $3D$ Euclidean space. Therefore, for any point $p \in S$, $p = (x(u,v), y(u,v), z(u,v))$, (x,y,z) are Cartesian coordinates. The basis vectors in $T_p(S)$ are $\{\boldsymbol{x}_u, \boldsymbol{x}_v\}$.

Let us label $\{\boldsymbol{x}_u, \boldsymbol{x}_v\}$ by $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$, with corresponding coordinates $\{u, v\}$ being labelled by $\{x^1, x^2\}$. When the coordinates change, the direction and length of $\boldsymbol{e}_i$ should change accordingly and the variation can be written as components along $\boldsymbol{e}_1, \boldsymbol{e}_2$ and $\boldsymbol{N}$, the unit normal vector. From Eq. E.3, let the new coordinates be $\{\bar{x}^1, \bar{x}^2\}$, we have:

$$\boldsymbol{e}'_i = \frac{\partial x^j}{\partial \bar{x}^i}\boldsymbol{e}_j, \qquad \boldsymbol{e}_i = \frac{\partial \bar{x}^j}{\partial x^i}\boldsymbol{e}'_j, \tag{E.4}$$

where $i, j = 1, 2$. Recall Eq. 3.38, it can be seen:

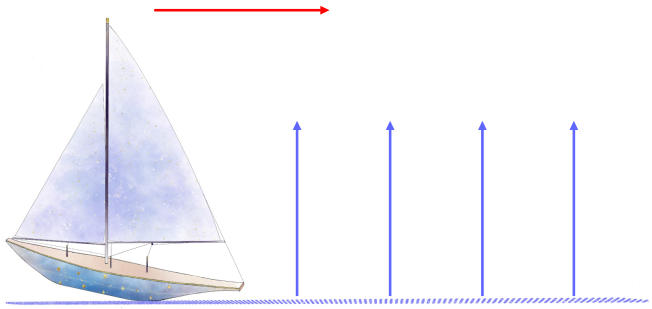
$$\begin{aligned} \frac{\partial \boldsymbol{e}_i}{\partial x^k} &= \frac{\partial}{\partial x^k}\left(\frac{\partial \bar{x}^j}{\partial x^i}\boldsymbol{e}'_j\right), \\ &= \frac{\partial^2 \bar{x}^j}{\partial x^k \partial x^i}\frac{\partial x^q}{\partial \bar{x}^j}\boldsymbol{e}_q + \frac{\partial \bar{x}^j}{\partial x^i}\frac{\partial \boldsymbol{e}'_j}{\partial x^k}, \\ &= \Gamma^q{}_{ki}\boldsymbol{e}_q + \langle \boldsymbol{N}, \frac{\partial \boldsymbol{e}_i}{\partial x^k}\rangle \boldsymbol{N}, \end{aligned} \tag{E.5}$$

$\Gamma^q{}_{ik}$ is defined as the *Christoffel symbol of the second kind* [229] and Eq. E.5 is the *Gauss–Weingarten relation* [228]. The *Christoffel symbol of the first kind* $\Gamma_{ik,j}$ can be obtained by using the metric to lower the indices of $\Gamma^q{}_{ik}$ (use the Einstein summation convention):

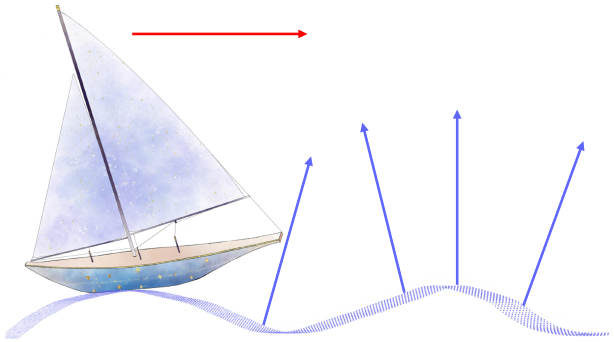
$$\Gamma_{ik,j} = g_{jq}\Gamma^q{}_{ik}, \tag{E.6}$$

$\{g_{ij}\}$ is the matrix form of the first fundamental form Π_1^p (or, equivalently, the metric) in basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$. The Christoffel symbol performs as a **connection**, which reveals the variation from point to point along a path of a vector when parallel transported on a manifold. Let us take a 2D space as an example, in which "parallel transport" means a vector is pointing in the same direction with its length unchanged during the transport. As shown in Fig. E.1(a), if a ship is travelling along a straight line, then the flag will always point upward in the same direction. However, if now we curve the line[Fig. E.1(b)], the flag will point in different directions. To make the transport a parallel transport, we need to allow the flag to rotate freely to make it always pointing upwards, as shown in Fig. E.1(c) by the orange vectors. The rotation angle along the curved path is determined by the geometry at each point on this path, and it is what a connection can tell us.

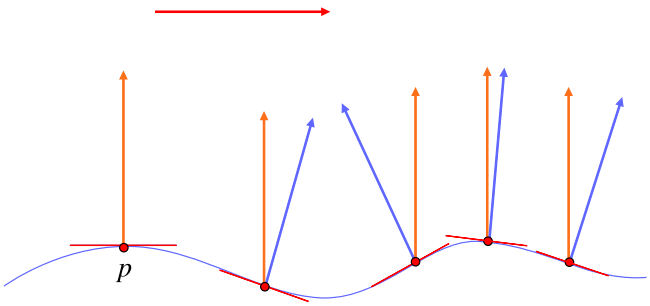
Extending to the 3D Euclidean space $\mathbb{E}^3$, then we do not need to consider the $h_{ij}\boldsymbol{N}$ term. Thus, for all $p \in \mathbb{E}^3$, recall Eq. E.3, let $\{\bar{x}_i\}, i = 1, 2, 3$ be Cartesian coordinates, which means



(a)



(b)



(c)

Figure E.1 Sketch of how the direction a flag pointing into is changed when the ship is moving along (a) a straight line, (b) a curved line in a 2D space. Red arrows shows the direction in which the ship is moving and blue arrows shows the direction into which the flag is pointing. (c) Sketch of how a parallel transport of the flag will be when the flag is allowed to rotate. Orange arrows shows the direction of the flag pointing into when it is parallel transported (and allowed to rotate) along the curve and blue arrows give the direction of the flag pointing into when it is fixed on the ship.

their corresponding basis vectors satisfy $\langle \boldsymbol{b}_i, \boldsymbol{b}_j \rangle = \delta_{ij}$, and we have:

$$\begin{aligned} g_{jq} &= \frac{\partial \bar{x}^i}{\partial x^j} \frac{\partial \bar{x}^k}{\partial x^q} \langle \boldsymbol{b}_i, \boldsymbol{b}_k \rangle \\ &= \frac{\partial \bar{x}^k}{\partial x^j} \frac{\partial \bar{x}^k}{\partial x^q} \end{aligned} \tag{E.7}$$

Hence, the Christoffel symbol of the second kind can be expressed as:

$$\begin{aligned}\Gamma_{ik,j} &= \frac{\partial^2 \bar{x}^l}{\partial x^k \partial x^i}\frac{\partial x^q}{\partial \bar{x}^l} g_{jq}, \\ &= \frac{\partial^2 \bar{x}^l}{\partial x^k \partial x^i}\frac{\partial x^q}{\partial \bar{x}^l}\frac{\partial \bar{x}^p}{\partial x^j}\frac{\partial \bar{x}^p}{\partial x^q}, \\ &= \frac{\partial^2 \bar{x}^l}{\partial x^k \partial x^i}\frac{\partial \bar{x}^p}{\partial x^j}\delta^p_l, \\ &= \frac{\partial^2 \bar{x}^l}{\partial x^k \partial x^i}\frac{\partial \bar{x}^l}{\partial x^j}.\end{aligned} \tag{E.8}$$

Now consider the differential of $g_{ij}, i, j = 1, 2$ with respect to the coordinate x^k:

$$\begin{aligned}\frac{\partial g_{ij}}{\partial x^k} &= \frac{\partial}{\partial x^k}\left(\frac{\partial \bar{x}^p}{\partial x^i}\frac{\partial \bar{x}^p}{\partial x^j}\right), \\ &= \frac{\partial^2 \bar{x}^p}{\partial x^i \partial x^k}\frac{\partial \bar{x}^p}{\partial x^j} + \frac{\partial \bar{x}^p}{\partial x^i}\frac{\partial^2 \bar{x}^p}{\partial x^j \partial x^k}, \\ &= \Gamma_{ik,j} + \Gamma_{jk,i}.\end{aligned} \tag{E.9}$$

Hence, denote $\partial g_{ij}/\partial x^k$ by $g_{ij,k}$:

$$\Gamma_{ik,j} = \frac{1}{2}(g_{ij,k} + g_{jk,i} - 2g_{ik,j}). \tag{E.10}$$

Similarly, we have:

$$\Gamma^q{}_{ik} = \frac{1}{2}g^{jq}(g_{ij,k} + g_{jk,i} - 2g_{ik,j}). \tag{E.11}$$

E.2 Generalised Hamiltonian Method

This method was first proposed by Dirac [218] to deal with generalised Hamiltonian dynamics when a system is under several constraints. In this section, we would like to introduce this method. Consider. a 3D manifold. Let the original Lagrangian be:

$$L(q_a, \dot{q}_a, t) = T - V = \frac{1}{2}\dot{q}_a \dot{q}^a - V(q), \tag{E.12}$$

where q_a are generalised coordinates and $\dot{q}_a = \mathrm{d}q_a/dt$, t is a parameter usually referred as time, $q = (q_1, q_2, q_3)$. If this system is under constraints that vary slowly with time:

$$f_i(q_a, t) = 0, \qquad 1 \le i \le 3, \qquad \frac{\partial f_i}{\partial t} \approx 0 \tag{E.13}$$

based on the Lagrangian method, the new Lagrangian is now expressed as:

$$L' = L + \sum_i \lambda_i f_i. \tag{E.14}$$

Here, λ_i are Lagrangian multipliers that are functions of q_a and t. Thus, the general momenta:

$$P_a = \frac{\partial L}{\partial \dot{q}_a}, \qquad P_\lambda^i = \frac{\partial L}{\partial \dot{\lambda}_i}, \tag{E.15}$$

where P_λ^i are canonical momenta contributed by constraints f_i. Thus, P_λ^i can impose constraints in the phase space by requiring $P_\lambda^i \approx 0$, given $L = 0$ is a strong condition independent of variations. More details can be found in Ref. [233]. Moreover, the P_λ^i are called *primary constraints*. Confinements mentioned above defines geometric structure of the phase space. The total Hamiltonian is thus expressed as:

$$H^* = \dot{q}_a P^a - L + \sum_i u_i P_\lambda^i = \frac{1}{2} P_a P^a + V - \lambda^i f_i + u_i P_\lambda^i = H - \lambda^i f_i + u_i P_\lambda^i, \tag{E.16}$$

where u_i are called *functions of velocities* and are required to weakly vanish ($u_i \approx 0$). Now, to find the geometric structure of this system, we need to find other constraints according to the *consistency conditions* using $f_i = 0$. That is to say, we expect the equation of motion of all constraints to weakly vanish. For example, the equation of motion of a constraint $f_i = 0$ should be $\dot{f}_i \approx \{f, H^*\}_P \approx 0$, where $\{,\}_P$ is the *Poisson bracket* that gives:

$$\{A, B\}_P = \sum_i \frac{\partial A}{\partial q_i}\frac{\partial B}{\partial P_i} - \frac{\partial A}{\partial P_i}\frac{\partial B}{\partial q_i}, \tag{E.17}$$

where A, B are functions of generalised coordinates q and momenta. The *Poisson bracket* has the following properties (A, B, C are all functions defined in the phase space, a, b are numbers):

- Anticommutativity:

$$\{A, B\}_P = -\{B, A\}_P \tag{E.18}$$

- Bilinearity:

$$\{aA + bB, C\}_P = a\{A, C\}_P + b\{B, C\}_P \tag{E.19}$$

- Leibniz's rule:

$$\{AB,C\}_P = A\{B,C\}_P + \{A,C\}_P B \tag{E.20}$$

- Jacobi identity:

$$\{A,\{B,C\}_P\}_P + \{B,\{C,A\}_P\}_P + \{C,\{A,B\}_P\}_P = 0 \tag{E.21}$$

Repeatedly taking the Poisson bracket until we arrive at terms involving u_i which is required to weakly vanish already. We obtain:

$$\begin{aligned}
&\phi_1 = f_i \approx 0, \\
&\phi_2 = P^i_\lambda = \{f_i, H^*\}_P \approx 0, \\
&\phi_3 = \dot{P}^i_\lambda = \{P^i_\lambda, H^*\}_P \approx 0, \\
&\phi_4 = \ddot{P}^i_\lambda = \{\dot{P}^i_\lambda, H^*\}_P \approx 0, \\
&\dots
\end{aligned} \tag{E.22}$$

Constraints other than ϕ_1 listed in Eq. E.22 are said to be *secondary constraints*. If a constraint obtained in Eq. E.22 satisfies:

$$\{\phi_i, \phi_j\}_P = 0, \tag{E.23}$$

for at least one ϕ_j, then ϕ_i is said to be of *second-class*. Otherwise, if for all ϕ_j Eq. E.23 holds, ϕ_i is said to be of *first-class*. Since we need the commutation to reveal the geometric structure, only second-class constraints that can contribute non-zero terms will be considered. Hence the Dirac bracket, as an extension to the Poisson bracket, of two functions A and B is now defined as:

$$\{A,B\}_D = \{A,B\}_P + \sum_{i,j} \{A,\phi_i\}_P M^{-1}_{ij} \{\phi_j,B\}_P, \tag{E.24}$$

where:

$$\{M_{ij}\} = \{\{\phi_i,\phi_j\}_P\} \tag{E.25}$$

E.3 Details of Hamiltonian Calculation

E.3.1 Simplification Using Gauss Normal Coordinates

Noticing that $F(\boldsymbol{x}) = q_0$ is an implicit function that describes the surface S. In Gauss normal coordinates in basis $\{\boldsymbol{g}_\mu\}$, $\mu = 0, 1, 2$ (coordinates $\{q^\mu\}$) and Cartesian coordinates in basis $\{\boldsymbol{e}_i\}$, $i = 1, 2, 3$ (coordinates $\{x^i\}$), we have the coordinate transformation (Appendix E.1.1):

$$\boldsymbol{e}_i = \frac{\partial q^\mu}{\partial x^i}\boldsymbol{g}_\mu, \qquad \boldsymbol{g}_\mu = \frac{\partial x^i}{\partial q^\mu}\boldsymbol{e}_i \tag{E.26}$$

where $\{\boldsymbol{g}_\mu\}$ are basis vectors and $\mu = 0, 1, 2$, $\{q^\mu\}$ are coordinates in the in the Gauss normal coordinate system. $\boldsymbol{g}_0$ coincides with the unit normal vector at each point of S. Hence we have:

$$\begin{aligned}
\langle \boldsymbol{g}_0, \boldsymbol{g}_0 \rangle = 1 = \langle \frac{\partial x^i}{\partial q^0}\boldsymbol{e}_i, \frac{\partial x^j}{\partial q^0}\boldsymbol{e}_j \rangle = \frac{\partial x^i}{\partial q^0}\frac{\partial x^i}{\partial q^0}, \\
\langle \boldsymbol{g}_0, \boldsymbol{g}_a \rangle = 0 = \langle \frac{\partial x^i}{\partial q^0}\boldsymbol{e}_i, \frac{\partial x^j}{\partial q^a}\boldsymbol{e}_j \rangle = \frac{\partial x^i}{\partial q^0}\frac{\partial x^i}{\partial q^a}, \quad a = 1, 2.
\end{aligned} \tag{E.27}$$

For the implicit function $F(\boldsymbol{x}) = q_0$, according to Eq. E.27, we see:

$$\partial_i F(\boldsymbol{x}) = \partial_i q^0 \Rightarrow \partial_i F(\boldsymbol{x}) \partial^i F(\boldsymbol{x}) = 1, \tag{E.28}$$

which means when using an implicit function, we can have the unit normal vector $n^i \boldsymbol{e}_i$ expressed as:

$$n^i \boldsymbol{e}_i = \partial^i F(\boldsymbol{x}) \boldsymbol{e}_i. \tag{E.29}$$

Thus, $\boldsymbol{g}_0$ in our Gauss normal coordinates basis is $\boldsymbol{g}_0 = n^i \boldsymbol{e}_i$. For the metric g in the Gauss normal coordinates, according to Eq. E.7, we have:

$$g_{\mu\nu} = \frac{\partial x^i}{\partial q^\mu}\frac{\partial x^i}{\partial q^\nu} \tag{E.30}$$

Referring to Eq. B.18, given our function satisfies Eq. E.29, which means the evolution field $\boldsymbol{F}$ of this surface in Eq. B.18 is a normal vector here, we have:

$$\partial_{q^0} g_{ab} = -2h_{ab}, \qquad a, b = 1, 2 \tag{E.31}$$

Also, as for the Christoffel symbol, we have:

$$\Gamma_{\mu\nu,\sigma} = \frac{\partial^2 x^i}{\partial q^\mu \partial q^\nu}\frac{\partial x^i}{\partial q^\sigma}, \tag{E.32}$$

we see:

$$n^i = \partial_i F, \qquad n^\mu = \frac{\partial x^i}{\partial q^\mu} n^i, \qquad n^i n_i = 1. \tag{E.33}$$

Moreover, in Cartesian coordinates:

$$n^i = n_i. \tag{E.34}$$

Hence, we have:

$$\begin{aligned} n_\mu \partial_\sigma n_\nu &= \frac{\partial x^i}{\partial q^\mu} n_i \frac{\partial}{\partial q^\sigma}\left(\frac{\partial x^j}{\partial q^\nu} n_j\right), \\ &= \frac{\partial x^i}{\partial q^\mu} \frac{\partial^2 x^j}{\partial q^\sigma \partial q^\nu} \delta_{ij}, \\ &= \Gamma_{\sigma\nu,\mu}. \end{aligned} \tag{E.35}$$

And:

$$\Gamma^\mu{}_{\sigma\nu} = \Gamma^\mu{}_{\nu\sigma} = \frac{\partial q^\mu}{\partial x^i} \frac{\partial^2 x^i}{\partial q^\sigma \partial q^\nu}. \tag{E.36}$$

E.3.2 Commutation Relation and Christoffel Symbol

Consider the following equation:

$$\begin{aligned} [P_i, P_j]|\psi\rangle &= (P_i P_j - P_j P_i)|\psi\rangle, \\ &= \frac{\hbar^2}{2}\partial^k[(n_i \partial_j n_k - n_j \partial_i n_k)|\psi\rangle] - \frac{\hbar^2}{2}(n_i \partial_j n_k - n_j \partial_i n_k)\partial^k|\psi\rangle, \\ &= \frac{\hbar^2}{2}[\partial^k(n_i \partial_j n_k - n_j \partial_i n_k)]|\psi\rangle. \end{aligned} \tag{E.37}$$

Thus, we have:

$$[P_i, P_j] = \frac{\hbar^2}{2}\partial^k(n_i \partial_j n_k - n_j \partial_i n_k) = \frac{\hbar^2}{2}\partial^k(\partial_i F \partial_j \partial_k F - \partial_j F \partial_i \partial_k F) \tag{E.38}$$

Then taking the expansion, we have:

$$\begin{aligned} &[P_i, P_j] \propto (\partial^k \partial_i F)(\partial_j \partial_k F) + \partial_i F \partial^k \partial_j \partial_k F - (\partial^k \partial_j F)(\partial_i \partial_k F) - \partial_j F \partial^k \partial_i \partial_k F, \\ &[P_i, P_j] = \frac{\hbar^2}{2}[(\partial^k \partial_i F)(\partial_j \partial_k F) - (\partial^k \partial_j F)(\partial_i \partial_k F)]. \end{aligned} \tag{E.39}$$

Here we use the relation that the mean curvature equal to one half of the trace of the second fundamental form if the basis is orthonormal (Sec. 3.3.1, and the basis we use here is the Cartesian coordinates), which gives [217]:

$$M = -\frac{1}{2}\partial^k n_k = -\frac{1}{2}\partial^k \partial_k F. \tag{E.40}$$

Also, we see:

$$\begin{aligned}\partial_i(n^k\partial_j n_k)-\partial_j(n^k\partial_i n_k) =&(\partial_i\partial^k F)(\partial_j\partial_k F)\\ &+\partial^k F\partial_i\partial_j\partial_k F-(\partial_j F\partial^k F)(\partial_i\partial_k F)-\partial^k F\partial_j\partial_i\partial_k F,\end{aligned} \tag{E.41}$$

which means:

$$\frac{i\hbar}{2}(P_i\Gamma^k{}_{jk}-P_j\Gamma^k{}_{ik})=\frac{\hbar^2}{2}[(\partial_i\partial^k F)(\partial_j\partial_k F)-(\partial_j F\partial^k F)(\partial_i\partial_k F)]. \tag{E.42}$$

Therefore, we have:

$$[P_i,P_j]=\frac{i\hbar}{2}[P_i(n^k\partial_j n_k)-P_j(n^k\partial_j n_k)]. \tag{E.43}$$

E.3.3 Coordinate Transformation of Commutation Relation

Let us consider the following equation that based on Eq. 6.20:

$$\{P_i,P_j\}_D=P_k(n_i\partial_j n^k-n_j\partial_i n^k). \tag{E.44}$$

On the left-hand side, we have:

$$\{P_i,P_j\}_D=\{\frac{\partial q^\mu}{\partial x^i}p_\mu,\frac{\partial q^\nu}{\partial x^j}p_\nu\}_D=\frac{\partial q^\mu}{\partial x^i}\frac{\partial q^\nu}{\partial x^j}\{p_\mu,p_\nu\}_D. \tag{E.45}$$

On the right-hand side, we have:

$$\begin{aligned}P_k n_i\partial_j n_k&=\frac{\partial q^\sigma}{\partial x^k}p_\sigma\frac{\partial q^\mu}{\partial x^i}n_\mu\frac{\partial q^\nu}{\partial x^j}\frac{\partial}{\partial q^\nu}\left(\frac{\partial q^\sigma}{\partial x^k}n_\sigma\right),\\ &=\frac{\partial q^\sigma}{\partial x^k}p_\sigma\frac{\partial q^\mu}{\partial x^i}n_\mu\frac{\partial q^\nu}{\partial x^j}\left(\frac{\partial^2 q^\sigma}{\partial q^\nu\partial x^k}n_\sigma+\frac{\partial q^\sigma}{\partial x^k}\partial_\nu n_\sigma\right).\end{aligned} \tag{E.46}$$

Hence, according to the constraint $P^i n_i=P_i n^i=0$ (ϕ_3 in Eq. 6.17):

$$\frac{\partial x^k}{\partial q^\sigma}p^\sigma\frac{\partial q^\mu}{\partial x^i}n_\mu\frac{\partial q^\nu}{\partial x^j}\frac{\partial^2 q^\sigma}{\partial q^\nu\partial x^k}n_\sigma\propto p^\sigma n_\sigma=0 \tag{E.47}$$

We see that:

$$P^k n_i\partial_j n_k=\frac{\partial x^k}{\partial q^\sigma}\frac{\partial q^\mu}{\partial x^i}\frac{\partial q^\nu}{\partial x^j}\frac{\partial q^\sigma}{\partial x^k}p^\sigma n_\mu\partial_\nu n_\sigma \tag{E.48}$$

Therefore,

$$\{p_\mu,p_\nu\}_D=p^\sigma(n_\mu\partial_\nu n_\sigma-n_\nu\partial_\mu n_\sigma) \tag{E.49}$$

According to Eq. E.43, we have:

$$\{p_\mu,p_\nu\}_D=p\mu(n^\sigma\partial_\nu n_\sigma)-p_\nu(n^\sigma\partial_\mu n_\sigma), \tag{E.50}$$

which can yield Eq.6.29.

Bibliography

[1] Lin Yutang. *The Importance of Understanding (古文小品译英)*. Foreign Language Teaching and Research Press, 2009.

[2] E. Pereloma and D. V. Edmonds, editors. *Phase Transformations in Steels: Diffusionless Transformations, High Strength Steels, Modelling And Advanced Analytical Techniques*. Elsevier, 2012.

[3] D. C. Lagoudas, editor. *Shape Memory Alloys: Modeling and Engineering Applications*. Springer Science & Business Media, 2008.

[4] H. Czichos. Adolf Martens and the research on martensite. In *European Symposium on Martensitic Transformations*, pages 3–14. EDP Sciences, 1989.

[5] K. Bhattacharya. *Microstructure of Martensite : Why It Forms and How It Gives Rise to The Shape-Memory Effect*. Oxford University Press, 2003.

[6] Y. A. Izyumov, V. M. Laptev, and V. N. Syromyatnikov. Phenomenological theory of martensitic and reconstructive phase transitions. *Phase Transit.*, 49(1-3):1, 1994.

[7] W. Campbell. Twenty-five years of metallography, 3rd Howe memorial award, 1926. *Trans. AIME*, 73:1135, 1926.

[8] A. B. Greninger and V. G. Mooradian. Strain transformation in metastable beta copper-zinc and beta copper-titanium alloys. *Trans. AIME*, 128:337, 1938.

[9] D. Dunne. Martens-ite. *Metals*, 8(6):395, 2018.

[10] F. Falk. Model free energy, mechanics, and thermodynamics of shape memory alloys. *Acta Mater.*, 28(12):1773, 1980.

[11] V. K. Wadhawan. Ferroelasticity and related properties of crystals. *Phase Transit.*, 3(1):3, 1982.

[12] V. Podzorov, B. G. Kim, V. Kiryukhin, M. E. Gershenson, and S-W. Cheong. Martensitic accommodation strain and the metal-insulator transition in manganites. *Phys. Rev. B*, 64:140406, 2001.

[13] M. Fremond and S. Miyazaki. *Shape Memory Alloys*. Springer-Verlag Wien GmbH, 1996.

[14] K. Amauchi, I. Ohkata, K. Tsuchiya, and S. Miyazaki, editors. *Shape Memory and Superelastic Alloys: Technolgies and Applications*. Woodhead Publishing, 2011.

[15] K. Otsuka and C. M. Wayman. *Shape Memory Materials*. Cambridge University Press, 1998.

[16] M. S. Hu, H. R. Wenk, and D. Sinitsyna. Microstructures in natural perovskites. *Am. Mineral.*, 77(3-4):359, 1992.

[17] X. X. Zhang and M. F. Qian. *Magnetic Shape Memory Alloys: Preparation, Martensitic Transformation and Properties*. Springer Nature, 2021.

[18] W. M. Kriven. Martensitic toughening of ceramics. *Mater. Sci. Eng.*, A 127(2):249, 1990.

[19] G. B. Olson and H. Hartman. Martensite and life: displacive transformations as biological processes. *J. phys., Colloq.*, 43(C4):855, 1982.

[20] L. Schetky. Shape-memory alloys. *SciAm*, 241:74–82, 1979.

[21] K. Otsuka and X. Ren. Recent developments in the research of shape memory alloys. *Intermetallics*, 7(5):511, 1999.

[22] A. Ölander. An electrochemical investigation of solid cadmium-gold alloys. *J. Am. Chem. Soc.*, 54(10):3819, 1932.

[23] L. C. Chang and T. A. Read. Plastic deformation and diffusionless phase changes in metals. *JOM*, 3:47, 1951.

[24] T. Tadaki, K. Otsuka, and K. Shimizu. Shape memory alloys. *Ann. Rev. Mater. Sci*, 18(1):25, 1988.

[25] H. Zhang. A theory for structural phase transitions in $BaTiO_3$ single crystal and $PbZrO_{3-x}PbTiO_3$ solid solution. *AIP Adv.*, 3:042118, 2013.

[26] Y. B. Wang and R.C. Liebermann. Electron microscopy study of domain structure due to phase transitions in natural perovskite. *Phys. Chem. Minerals*, 20(1):147, 1993.

[27] E. Fertman, E. Syrkin, V. Lykah, V. Desnenko, A. Beznosov, P. Pal-Val, L. Pal-Val, A. Fedorchenko, D. Khalyavin, and A. Feher. Structural phase transition in $La_{2/3}Ba_{1/3}MnO_3$ perovskite: Elastic, magnetic, and lattice anomalies and microscopic mechanism. *AIP Adv.*, 5(7):077189, 2015.

[28] N. Balke, T. Schenk, I. Stolichnov, and A. Gruverman. *Ferroelectricity in Doped Hafnium Oxide: Materials, Properties and Devices*. Woodhead Publishing, 2019.

[29] J. R. Greer and J. T. M. De Hosson. Plasticity in small-sized metallic systems: Intrinsic versus extrinsic size effect. *Prog. Mater. Sci.*, 56(654), 2011.

[30] T. Waitz, K. Tsuchiya, T. Antretter, and F. D. Fischer. Phase transformations of nanocrystalline martensitic materials. *MRS Bull.*, 34:814, 2009.

[31] E. C. Bain and N. Y. Dunkirk. The nature of martensite. *Trans. AIME*, 70(1):25, 1924.

[32] N. Hatcher, O. Yu. Kontsevoi, and A. J. Freeman. Role of elastic and shear stabilities in the martensitic transformation path of NiTi. *Phys. Rev. B*, 80(14):144203, 2009.

[33] G. Kostorz, editor. *Phase Transformations in Materials*. Weinheim: Wiley-vch, 2001.

[34] M. Cohen, G. B. Olson, and P. C. Clapp, editors. *Proc. ICOMAT '79. Cambridge (MA): MIT*, 1979.

[35] J. A. Krumhansl. Commentary on first order displacive phase transitions. Technical report, DOE-LINE Workshop, Berkeley, C.A., October 24-28 1989.

[36] Y. A. Izyumov and V. N. Syromaytnikov. *Phase Transitions and Crystal Symmetry*. Kluwer Academic Publishers, Dordrecht, Boston, London., 1990.

[37] P. Tol'edano and V. Dimitriev. *Reconstructive Phase Transitoins: In Crystals and Quasicrystals*. World Scientific, 1996.

[38] J. F. Scott. Soft-mode spectroscopy: Experimental studies of structural phase transitions. *Rev. Mod. Phys.*, 46(1):83, 1974.

[39] W. Cochran. Crystal stability and the theory of ferroelectricity. *Adv. Phys.*, 9:387, 1960.

[40] R. H. Lyddane, R. G. R. G. Sachs, and E. Teller. On the polar vibrations of alkali halides. *Phys. Rev.*, 59(8):673, 1941.

[41] D. de Fontaine. Cluster approach to order-disorder transformations in alloys. *Solid State Phys.*, 47:33, 1994.

[42] D. de Fontaine. Configurational thermodynamics of solid solutions. *Solid State Phys.*, 34:73, 1979.

[43] N. Hatcher, O. Yu. Kontsevoi, and A. J. Freeman. Martensitic transformation path of NiTi. *Phys. Rev. B*, 79:020202(R), 2009.

[44] S. T. Hyde and S. Andersson. Differential geometry of crystal structure descriptions, relationships and phase transformation. *Z. Kristallogr. Cryst. Mater.*, 170(1-4):225, 1985.

[45] S. T. Hyde and S. Andersson. The martensite transition and differential geometry. *Z. Kristallogr. Cryst. Mater.*, 174(1-4):225, 1986.

[46] S. Andersson, S. T. Hyde, K. Larsson, and S. Lidin. Minimal surfaces and structures: From inorganic and metal crystals to cell membranes and biopolymers. *Chem. Rev.*, 88(1):221, 1988.

[47] M. P. Do Carmo. *Differential Geometry of Curves and Surfaces*. Dover, 2016.

[48] H. G. von Schnering and R. Nesper. Die natürliche anpassung von chemischen strukturen an gekrümmte flächen. *Angew. Chem. Int. Ed.*, 99(11):1097 – 1119, 1987.

[49] H. G. von Schnering and R. Nesper. Nodal surfaces of fourier series: Fundamental invariants of structured matter. *Z. Physik B – Condensed Matter*, 83:407–412, 1991.

[50] R. Nesper and H. G. von Schnering. Periodische äquipotential flächen. *Z. Kristallogr*, 170:138, 1985.

[51] A. T. Fomenko and A. A. Tuzhilin. *Elements of the Geometry and Topology of Minimal Surfaces in Three-Dimensional Space*. American Mathematical Society, 2000.

[52] H. A. Schwarz. *Gesammelte Mathematische Abhandlungen*, volume 260. American Mathematical Soc., 1972.

[53] H. A. Schwarz. Bestimmung einer speciellen minimalfläche. In *Gesammelte Mathematische Abhandlungen*, pages 6–91. Springer, 1980.

[54] S. T. Hyde and S. Andersson. A cubic structure consisting of a lipid bilayer forming an infinite periodic minimum surface of the gyroid type in the glycerolmonooleat-water system. *Z. Kristallogr. Cryst. Mater.*, 168(1-4):213, 1984.

[55] W. Helfrich. Elastic properties of lipid bilayers: theory and possible experiments. *Z. Naturforsch. C*, 28(11):693–703, 1973.

[56] L. D. Landau and E. M. Lifshitz. *Statistical Physics: Volume 5*. Elsevier, 2013.

[57] L. Casella and A. Zaccone. Soft mode theory of ferroelectric phase transitions in the low-temperature phase. *J. Phys.: Condens. Matter*, 33:165401, 2021.

[58] S. Kamba. Soft-mode spectroscopy of ferroelectrics and multiferroics: A review. *APL Mater.*, 9(2), 2021.

[59] M. Maschek, S. Rosenkranz, R. Heid, A. H. Said, P. Giraldo-Gallo, I. R. Fisher, and F. Weber. Wave-vector-dependent electron-phonon coupling and the charge-density-wave transition in $TbTe_3$. *Phys. Rev. B*, 91:235146, 2015.

[60] C. Meingast, A. Shukla, L. Wang, R. Heid, F. Hardy, M. Frachet, K. Willa, T. Lacmann, M. L. Tacon, M. Merz, A. Haghighirad, and T. Wolf. Charge density wave transitions, soft phonon, and possible electronic nematicity in $BaNi_2(As,P)_2$. *Phys. Rev. B*, 106:144507, 2022.

[61] V. L. Ginzburg. *Ferroelectrics*, 267(1):23–32, 2002.

[62] V. L. Ginzburg. Theory of ferroelectric phenomena. *Usp. Fiz. Nauk*, 38(8):490, 1949.

[63] A. P. Levanyuk. Contribution to the theory of light scattering near the second-order phase-transition points, 1959.

[64] R. A. Cowley. The lattice dynamics of an anharmonic crystal. 12(48):421–480, 1963.

[65] R. A. Cowley. Structural phase transitions i. landau theory. *Adv. Phys.*, 29(1):1–110, 1980.

[66] H. T. Stokes and D. M. Hatch. *Isotropy Subgroups of The 230 Crystallographic Space Groupsroups*. World Scientific, 1988.

[67] V. Seetharaman. Deformation and martensitic transformation. *Bull. Mater. Sci*, 6(4):703, 1984.

[68] C. J. Bradley and A. P. Cracknell. *The Mathematical Theory of Symmetry In Solids: Representation Theory for Point Groups and Space Groups*. Oxford University Press, 2013.

[69] D. D. Vvedensky. *Transformations of Materials*. Morgan & Claypool Publishers, 2019.

[70] Z. S. Basinski and J. W. Christian. Crystallography of deformation by twin boundary movements in indium-thallium alloys. *Acta Metall.*, 2:101–116, 1954.

[71] K. Sotoodeh. *Case studies of material corrosion prevention for oil and gas valves*. Gulf Professional Publishing, 2022.

[72] L. Petrini and F. Migliavacca. Biomedical applications of shape memory alloys. *J. Metall.*, 2011, 2011.

[73] G. F. Andreasen and T. B. Hilleman. An evaluation of 55 cobalt substituted nitinol wire for use in orthodontics. *J. Am Dent Assoc.*, 82:1373–1375, 1975.

[74] E. Salzman M. Simon, R. Kaplan and D. Freiman. A vena cava filter using thermal shape memory alloy: experimental aspects. *Radiology*, 125:89–94, 1977.

[75] I. P. Lipscomb and L. D. M. Nokes. *The Application of Shape Memory Alloys in Medicine*. Paston Press Ltd, 1996.

[76] G. Lewis. Materials, fluid dynamics, and solid mechanics aspects of coronary artery stents: a state-of-the-art review. *J. Biomed. Mater. Res.*, 86:569–590, 2008.

[77] S. Tyagi, S. Singh, S. Mukhopadhyay, and U. A. Kaul. Self- and balloon-expandable stent implantation for severe native coarctation of aorta in adults. *Am. Heart J.*, 146:920–928, 2003.

[78] S. P. Hawkins and A. Al-Kutoubi. The Simon nitinol inferior vena cava filter: Preliminary experience in the UK. *Clin. Rad.*, 46(6):378–380, 1992.

[79] L. D. Landau and E. M. Lifshitz. *Course of Theoretical Physics, Statistical Physics*, volume 5. Pergamon Press, 3 edition, 1980.

[80] J. C. Tolédano and P. Tolédano. *The Landau Theory of Phase Transitions*. WORLD SCIENTIFIC, 1987.

[81] *Group Theory Application to the Physics of Condensed Matter*. Springer, 2007.

[82] M. Tinkham. *Grou Theory and Quantum Mechanics*. Dover, 2003.

[83] G. Ya. Lyubarskii. *The Application of Group Theory in Physics*. Elsevier, 2013.

[84] P. Tolédano. Symmetry aspects of structural transitions in liquid crystals. *Jap. Jour. Appl. Phys.*, 24(S2):121, 1985.

[85] D. J. Benson. *Polynomial Invariants of Finite Groups*. Cambridge University Press, 1993.

[86] W. Hergert and R. M. Geilhufe. *Group Theory in Solid State Physics and Photonics: Problem Solving with Mathematica*. John Wiley & Sons, 2018.

[87] S. Yue, B. Deng, Y. Liu, Y. Quan, R. Yang, and B. Liao. Phonon softening near topological phase transitions. *Phys. Rev. B*, 102:235428, 2020.

[88] C. V. Raman and T. M. K. Nedungadi. *Nature*, 145(3665):147, 1940.

[89] M. Born and K. Huang. *Dynamical Theory of Crystal Lattices*. Oxford University Press, 1954.

[90] G. P. Srivastava. *The Physics of Phonons*. CRC press, 2022.

[91] Jr. B. G. Dick and A. W. Overhauser. Theory of the dielectric constants of alkali halide crystals. *Phys. Rev.*, 112(1):90, 1958.

[92] R. A. Cowley and S. M. Shapiro. Structural phase transitions. *Adv. Phys.*, 75(11):111001, 2006.

[93] P. A. Fleury. The effects of soft modes on the structure and properties. *Annu. Rev. Mater. Sci.*, 6(1):157–180, 1976.

[94] I. Pallikara and J. M. Skelton. Phase stability of the tin monochalcogenides sns and snse: a quasi-harmonic lattice-dynamics study. *Phys. Chem. Chem. Phys.*, 23:19219, 2021.

[95] I. Pallikara, P. Kayastha, J. M. Skelton1, and L. D. Whalley. The physical significance of imaginary phonon modes in crystals. *Electron. Struct.*, 2022.

[96] S. Stackhouse, L. Stixrude, and B. B. Karki. Thermal conductivity of periclase (MgO) from first principles. *Phys. Rev. Lett.*, 104:208501, 2010.

[97] T. Tadano, Y. Gohda, and S. Tsuneyuki. Anharmonic force constants extracted from first-principles molecular dynamics: applications to heat transfer simulations. *J. Phys.: Condens. Matter*, 26:225402, 2014.

[98] D. Donadio and G. Galli. Atomistic simulations of heat transport in silicon nanowires. *Phys. Rev. Lett.*, 102:195901, 2009.

[99] L. Chaput. Direct solution to the linearized phonon Boltzmann equation. *Phys. Rev. Lett.*, 110:265506, 2013.

[100] A. Togo and I. Tanaka. First principles phonon calculations in materials science. *Scr. Mater.*, 108:1–5, 2015.

[101] F. VanGessel, J. Peng, and P. W. Chung. A review of computational phononics: the bulk, interfaces, and surfaces. *J. Mater. Sci.*, 53:5641–5683, 2018.

[102] B. Fultz. Vibrational thermodynamics of materials. *Prog. Mater. Sci.*, 55:247–352, 2010.

[103] Z. Nishiyama. *Martensitic Transformation*. Academic Press, 1978.

[104] G. Kurdjumov and G. Sachs. Über den mechanismus der stahlhärtung. *Z. Phys.*, 64(5):325, 1930.

[105] A. J. Bogers and W. G. Burgers. Partial dislocations on the {110} planes in the b.c.c. lattice and the transition of the f.c.c. into the b.c.c. lattice. *Acta Metall.*, 12:255–261, 1964.

[106] W. G. Burgers. On the process of transition of the cubic-body-centered modification into the hexagonal-close-packed modification of zirconium. *Physica.*, 1:561, 1934.

[107] A. Ojha A and H. Sehitoglu. Critical stress for the bcc–hcp martensite nucleation in ti–6.25at.%ta and ti–6.25at.%nb alloys. *Comput. Mater. Sci.*, 111:157–162, 2016.

[108] Z. Nishiyama. X-ray investigation of the mechanism of the transformation from face centered cubic lattice to body centered cubic. *Sci. Rep. Tohoku Univ.*, 637, 1934.

[109] C. Thiel, J. Voss, R. J. Martin, and P. Neff. Shear, pure and simple. *Int. J. Non-Linear Mech.*, 112:57–72, 2019.

[110] D. Hillel, editor. *Encyclopedia of Soils in The Environment*, volume 1. Academic Press, 2005.

[111] S. C. Han, J. W. Lee, and K. J. Kang. A new type of low density material: Shellular. *Adv. Mater.*, 27(37):5506–5511, 2015.

[112] K. Michielsen and D. G. Stavenga. Gyroid cuticular structures in butterfly wing scales: biological photonic crystals. *J. R. Soc. Interface*, 5:85–94, 2008.

[113] J. W. Galusha, L. R. Richey, J. S. Gardner, J. N. Cha, and M. H. Bartl. Discovery of a diamond-based photonic crystal structure in beetle scales. *Phys. Rev. E*, 77(5):050904, 2008.

[114] M. Lai, A. N. Kulak, D. Law, Z. Zhang, F. C. Meldrum, and D. J. Riley. Profiting from nature: macroporous copper with superior mechanical properties. *Chem. Commun.*, 34:3547, 2007.

[115] J. Sun, C. Bonneau nad Á. Cantín, A. Corma, M. J. Diaz-Cabanas, M. Moliner, D. Zhang, M. Li, and X. Zou. The ITQ-37 mesoporous chiral zeolite. 458(7242):1154–1157, 2009.

[116] F. S. Bates and G. H. Fredrickson. Block copolymers—designer soft materials. *Phys. Today*, 52(2):32–38, 1999.

[117] Y. Guo, M. I. N. Rosa, M. Gupta, B. E. Dolan, B. Fields, L. Valdevit, and M. Ruzzene. Minimal surface-based materials for topological elastic wave guiding. *Adv. Funct. Mater.*, 32(30):2004122, 2022.

[118] P. J.F. Gandy, S. Bardhan, A. L. Mackay, and J. Klinowski. Nodal surface approximations to the p. g, d and i-wp triply periodic minimal surfaces. *Chem. Phys. Lett.*, 336:187–195, 2001.

[119] P. J. F. Gandy and J. Klinowski. Geometric quantization of curvature energy in equipotential surfaces of ionic crystals. *J. Chem. Phys.*, 116:9431–9134, 2002.

[120] P. J. F. Gandy and J. Klinowski. Nodal surface approximation to the zero equipotetial surfaces for cubic lattices. *J. Math. Chem.*, 31(1):1–16, 2002.

[121] A. L. Mackay. Crystallographic surfaces. *Proc. R. Soc. Lond. A*, 442:47–59, 1993.

[122] S. T. Huang. *Solid X-Ray Analysis(固体 X 射线学)*. Beijing:Higher Education Press, 1985.

[123] K. Shinoda Y. Waseda, E. Matsubara. *X-Ray Diffraction Crystallography: Introduction, Examples and Solved Problems*. Springer Science & Business Media, 2011.

[124] M. Nakahara. *Geometry, Topology and Physics*. CRC press, 2018.

[125] S. M. Carroll. *Spacetime and Geometry*. Cambridge University Press, 2019.

[126] M. Spivak. *A Comprehensive Introduction to Differential Geometry*, volume 1. Publish Or Perish Inc., 1999.

[127] E. M. Stein and R. Shakarchi. *Complex Analysis*. Princeton University Press, 2003.

[128] Z. Nehari. *Conformal Mapping*. Courier Corporation, 2012.

[129] A. G. Weyhaupt. *New Families of Embedded Triply Periodic Minimal Surfaces of Genus Three in Euclidean Space*. PhD thesis, Indiana University, 2006.

[130] H. Chen and M. Weber. An orthorhombic deformation family of schwarz'h surfaces. *Trans. Am. Math. Soc* ., 374(3):2057, 2021.

[131] O. Bonnet. Note sur la théorie générale des surfaces. *CR Acad. Sci. Paris*, 37:529, 1853.

[132] A. H. Schoen. Infinite periodic minimal surfaces without self-intersections. Technical Report TN D-5541, NASA, Washington, DC, 1970.

[133] H. Chen and M. Traizet. Stacking disorder in periodic minimal surfaces. *SIAM J. Math. Anal.*, 53(1):855–887, 2021.

[134] P. Hohenberg and W. Kohn. Density functional theory (dft). *Phys. Rev.*, 136:B864, 1964.

[135] L. J. Sham W. Kohn. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133, 1965.

[136] P. J. Hasnip, K. Refson, M. I. J. Probert, J. R. Yates, S. J. Clark, and C. J. Pickard. Density functional theory in the solid state. *Phil. Trans. R. Soc. A*, 372(2011):20130270, 2014.

[137] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria. Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids. *Phys. Rev. Lett.*, 91:146401, 2003.

[138] G. Kresse and J. Hafner. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B*, 47:558, 1993.

[139] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54:11169, 1996.

[140] G. Kresse and J. Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mat. Sci.*, 6:15, 1996.

[141] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B*, 59:1758, 1999.

[142] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, and I. Dabo. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter.*, 21(39):39550, 2009.

[143] K. Schwarz and P. Blaha. Solid state calculations using WIEN2k. *Comput. Mater. Sci.*, 28:259–273, 2003.

[144] B. Santra. *Density-Functional Theory Exchange-Correlation Functionals for Hydrogen Bonds in Water*. PhD thesis, Technische Universität Berlin Berlin, 2010.

[145] R. O. Jones. In *Computational Nanoscience: Do It Yourself*, volume 31 of *NIC Series*, chapter Introduction to Density Functional Theory and Exchange-Correlation Energy Functionals, pages 45–70. John von Neumann Institute for Computing, Jülich, 2006.

[146] E. Engel and R. M. Dreizler. *Theoretical and Mathematical Physics: Density Functional Theory, And Advanced Course*. Springer, 2010.

[147] W. Koch and M. C. Holthausen. *A Chemist's Guide to Density Functional Theory*. Wiley-VCH, 2001.

[148] R. O. Jones. Density functional theory: Its origins, rise to prominence, and future. *Rev. Mod. Phys.*, 87(3):897, 2015.

[149] E. Cancès and C. Le Bris. Can we outperform the diis approach for electronic structure calculations? *Int. J. Quantum Chem.*, 79(2):82–90, 2000.

[150] P. Paulay. Convergence acceleration of iterative sequences. *Chem. Phys. Lett.*, 73(2):393–398, 1980.

[151] P. Pulay. Improved scf convergence acceleration. *J. Comput. Chem.*, 3(4):556–560, 1982.

[152] V. R. Saunders and I. H. Hillier. A "level–shifting"method for converging closed shell hartree–fock wave functions. *Int. J. Quantum Chem.*, 7(4):699–705, 1973.

[153] R. M. Martin. *Electronic structure: basic theory and practical methods*. Cambridge university press, 2020.

[154] D. Sholl and J. A. Steckel. *Density Functional Theory: A Practical Introduction*. John Wiley & Sons, 2011.

[155] D. H. Hamann, M. Schluter, and C. Chiang. Norm-conserving pseudopotentials. *Phys. Rev. Lett.*, 43(20), 1979.

[156] P. E. Blöchl. Projector augmented-wave method. *Phys. Rev. B*, 50:17953, 1994.

[157] A. Baldereschi. Mean-value point in the Brillouin zone. *Phys. Rev. B*, 7(12):5212, 1973.

[158] D. J. Chadi and M. L. Cohen. *Special points in the Brillouin zone*, 8(12):5747, 1973.

[159] J. Monkhorst H and J. D. Pack. Special points for Brillouin-zone integrations. *Phys. Rev. B*, 13:5188–5192, 1976.

[160] J. Monkhorst H and J. D. Pack. "special points for Brillouin-zone integrations" – a reply. *Phys. Rev. B*, 16:1748–1749, 1977.

[161] D. A. Young. *Phase Diagrams of The Elements*. Number UCRL-51902. Lawrence Livermore National Lab.(LLNL), Livermore, CA (United States), 1975.

[162] G. K. Straub and D. C. Wallace. Study of the martensitic phase transition in sodium'. *Phys. Rev. B*, 3(4):1234, 1971.

[163] T. Sjostrom, S. Crockett, and S. Rudin. *Phys. Rev. B*, 94:144101, 2016.

[164] C. Schmidt, F. Ernst, M. W. Finnis, and Vaclav Vitek. Prediction and observation of the BCC structure in pure copper at a Σ3 grain boundary. *Phys. Rev. Lett.*, 11:2160, 1995.

[165] C. Schmidt, M. W. Finnis, F. Ernst, and V. Vitek. Theoretical and experimental investigations of structures and energies of Σ= 3,[112] tilt grain boundaries in copper. *Philos. Mag. A*, 77:1161–1184, 1998.

[166] M. Sims, R. Briggs, T. J. Volz, S. Singh andS. Hamel, A. L. Coleman, F. Coppari, D. J. Erskine, M. G. Gorman, B. Sadigh, J. Belof, J. H. Eggert, R. F. Smith, and J. K. Wicks. Experimental and theoretical examination of shock-compressed copper through the fcc to bcc to melt phase transitions. *J. Appl. Phys.*, 132:075902, 2022.

[167] Y. J. Hao, L. Zhang, X. R. Chen, L. C. Cai, Q. Wu, and D. Alfè. Ab initio calculations of the thermodynamics and phase diagram of zirconium. *Phys. Rev. B*, 78(13):134101, 2008.

[168] W. J. Buehler, J. W. Gilfrich, and R. C. Wiley. Effect of lowtemperature phase changes on the mechanical properties of alloys near composition TiNi. *J. Appl. Phys.*, page 1475, 1963.

[169] G. T. Liu, K. T. Yin, J. K. Zhang, H. B. Wang, and H. Y. Liu. A new intermediate phase in compressed nitinol. *J. Alloys Compd.*, 817:153234, 2020.

[170] J. G. Niu and W. T. Geng. Anti-precursor effect of Fe on martensitic transformation in TiNi alloys. *Acta Mater.*, 104:18–24, 2016.

[171] K. G. Vishnu and A. Strachan. Phase stability and transformations in niti from density functional theory calculations. *Acta Mater.*, 58(3):745–752, 2010.

[172] S. Kibey, H. Sehitoglu, and D. D. Johnson. Energy landscape for martensitic phase transformation in shape memory NiTi. *Mater. Res. Soc. Symp. Proc.*, 57:1624–1629, 2009.

[173] Xu Yuanchong. *Selected Poems of Li Bai (李白诗选)*. Hu'nan People's Publishing House, 2007.

[174] I. M. James, editor. *History of topology*. Elsevier, 1999.

[175] G. Cantor. üeber unendliche, lineare punktmannichfaltigkeiten. *Math. Ann.*, 21:545–591, 1883.

[176] Y. Aharonov and D. Bohm. Significance of electromagnetic potentials in the quantum theory. *Phys. Rev.*, 115(3):485, 1959.

[177] A. Pressley. *Elementary Differential Geometry*. Springer Science & Business Media, 2010.

[178] S. Fujimori and M. Weber. Triply periodic minimal surfaces bounded by vertical symmetry planes. *Manuscr. Math.*, 129:29–53, 2009.

[179] M. Traizet. On the genus of triply periodic minimal surfaces[j]. journal of differential geometry. *J. Differ. Geom.*, 79(2):243–275, 2008.

[180] M. I. Aroyo, editor. *International Tables for Crystallography*, volume A, Space Group Symmetry. 2016.

[181] C. Kittel. *Introduction to Solid State Physics*. John Wiley & Sons, Inc., 2005.

[182] S. Gupta and A. Saxena, editors. *The Role of Topology in Materials*. Berlin/Heidelberg, Germany: Springer, 2018.

[183] P. Sittner, P. Lukás, D. Neov, V. Novák, and D. M. Toebbens. The martensite transition and differential geometry. *J. Phys. IV (France)*, 112:709–711, 2003.

[184] C. Cayron. What EBSD and TKD tell us about the crystallography of the martensitic B2-B19 transformation in NiTi shape memory alloys. 10(7):562, 2020.

[185] K. A. Persson, B. Waldwick, P. Lazic, and G. Ceder. Prediction of solid-aqueous equilibria: Scheme to combine first-principles calculations of solids with experimental aqueous states. *Phys. Rev. B*, 85:235438, 2012.

[186] I. Petousis, D. Mrdjenovich, E. Ballouz, M. Liu, D. Winston, W. Chen, T. Graf, T. D. Schladt, K. A. Persson, and F. B. Prinz. High-throughput screening of inorganic compounds for the discovery of novel dielectric and optical materials. *Scientific Data*, 4(1):160134, 2017.

[187] M. D. Jong, W. Chen, T. Angsten, A. Jain, R. Notestine, A. Gamst, M. Sluiter, C. Krishna Ande Chaitanya, S. V. D. Zwaag, J. J. Plata, C. Toher, S. Curtarolo, G. Ceder, K. A. Persson, and M. Asta. Charting the complete elastic properties of inorganic crystalline compounds. *Sci. Data*, 2(1):1–13, 2015.

[188] A. Wang, R. Kingsbury, M. McDermott, M. Horton, A. Jain, S. P. Ong, S. Dwaraknath, and K. A. Persson. A framework for quantifying uncertainty in DFT energy corrections. *Sci. Rep.*, 11(1):15496, 2021.

[189] M. Aykol, S. S. Dwaraknath, W. Sun, and K. A. Persson. Thermodynamic limit for synthesis of metastable inorganic materials. *Sci. Adv.*, 4(4), 2018.

[190] R. Tran, Z. H. Xu, B. Radhakrishnan, D. Winston, W. H. Sun, K. A. Persson, and S. P. Ong. Surface energies of elemental crystals. *Sci. Data*, 3:160080, 2016.

[191] M. D. Jong, W. Chen, H. Geerlings, M. Asta, and K. A. Persson. A database to enable discovery and design of piezoelectric materials. *Sci. Data*, 2(1):1–13, 2015.

[192] K. Latimer, S. Dwaraknath, K. Mathew, D. Winston, and K. A. Persson. Evaluation of thermodynamic equations of state across chemistry and structure in the materials project. *njp Comput. Mater.*, 4(1):40, 2018.

[193] Grain boundary properties of elemental metals. *Acta. Mater.*, 186:40–49, 2020.

[194] H. Ding, S. S. Dwaraknath, L. Garten, P. Ndione, D. Ginley, and K. A. Persson. Computational approach for epitaxial polymorph stabilization through substrate selection. *ACS Appl Mater Interfaces*, 8(20), 2016.

[195] Formation enthalpies by mixing GGA and GGA+U calculations, 2011.

[196] M. K. Horton, J. H. Montoya, M. Liu, and K. A. Persson. High-throughput prediction of the ground-state collinear magnetic order of inorganic materials using density functional theory. *npj Comput. Mater.*, 5(64), 2019.

[197] S. Singh, J. Nayak, A. Rai, P. Rajput, A. H. Hill, S. R. Barman, and D. Pandey. (3 + 1)d superspace description of the incommensurate modulation in the premartensite phase of ni_2mnga: a high resolution synchrotron x-ray powder diffraction study. *J. Phys.: Condens. Matter*, 25:212203, 2013.

[198] J. M. Munro, K. Latimer, M. K. Horton, S. Dwaraknath, and K. A. Persson. An improved symmetry-based approach to reciprocal space path selection in band structure calculations. *npj Comput Mater*, 6(112), 2020.

[199] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and A. K. Persson. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):011002, 07 2013.

[200] J. R. Morris, Y. Ye, M. Krcmar, and C. L. Fu. The role of phase stability in ductile, ordered B2 intermetallics. *Acta. Cryst. A*, 980:610, 2006.

[201] M. Krcmar, G. D. Samolyuk, and J. R. Morris. Stacking faults and alternate crystal structures for the shape-memory alloy niti. *Phys. Rev. M*, 4:103606, 2020.

[202] T. Chakraborty, J. Rogal, and R. Drautz. Martensitic transformation between competing phases in Ti–Ta alloys: A solid-state nudged elastic band study. *J. Phys. Condens. Matter*, 27(11):115401, 2015.

[203] D. R. Trinkle, D. M. Hatch, H. T. Stokes, R. G. Hennig, and R. C. Albers. Systematic pathway generation and sorting in martensitic transformations: Titanium α to ω. *Phys. Rev. B*, 72(1):014105, 2005.

[204] R. G. Hennig, D. R. Trinkle, J. Bouchet, S. G. Srinivasan, R. C. Albers, and J. W. Wilkins. Impurities block the α to ω martensitic transformation in titanium. *Nature*, 4(2):129–133, 2005.

[205] D. R. Trinkle, R. G. Hennig, S. G. Srinivasan, D. M. Hatch, M. D. Jones, H. T. Stokes, R.C. Albers, and J. W. Wilkins. New mechanism for the α to ω martensitic transformation in pure titanium. *Phys. Rev. Lett.*, 91(2):025701, 2003.

[206] P. S. Ghosh, A. Arya, R. Tewari, and G. K. Dey. Alpha to omega martensitic phase transformation pathways in pure zr. *J. Alloys Compd.*, 586:693–698, 2014.

[207] T. Suzuki and H. M. Ledbetter. Barrier energy for the b.c.c.-f.c.c. martensitic transition in sodium. *Phil. Mag.*, 48(1):83–94, 1983.

[208] S. V. Okatov, A. R. Kuznetsov, Yu. N. Gornostyrev, V. N. Urtsev, and M. I. Katsnelson. Effect of magnetic state on the γ-α transition in iron: first-principles calculations of the bain transformation path. *Phys. Rev. B*, 79:084111, 2009.

[209] L. H. Zhang, M. J. Cheng, X. H. Shi, J. W. Shuai, and Z. Z. Zhu. Bain and nishiyama–wassermann transition path separation in the martensitic transitions of fe. *RSC Adv.*, 11:3043, 2021.

[210] T. L. Wang, J. L. Du, S. Z. Wei, and F. Liu. Ab-initio investigation for the microscopic thermodynamics and kinetics of martensitic transformation. *Prog. Nat. Sci. Mater.*, 31:121, 2021.

[211] L. G. Wang and M. Łob. Structural stability of higher-energy phases and its relation to the atomic configurations of extended defects: The example of cu. *Phys. Rev. B*, 60:844, 1999.

[212] S. Jeong. Structural properties of bulk copper: Pseudopotential plane-wave-basis study. *Phys. Rev. B*, 53:13973, 1996.

[213] Y. Mishin, M. J. Mehl, D. A. Papaconstantopoulos, A. F. Voter, and J. D. Kress. Structural stability and lattice defects in copper: Ab initio, tight-binding, and embedded-atom calculations. *Phys. Rev. B*, 63:224106, 2001.

[214] Y. Mishin, D. Farkas, M. J. Mehl, and D. A. Papaconstantopoulos. Interatomic potentials for monoatomic metals from experimental data and ab initio calculations. *Phys. Rev. B*, 59:3393, 1999.

[215] P. Rangamani, A. Behzadan, and M. Holst. Local sensitivity analysis of the "membrane shape equation"derived from the helfrich energy. *Math. Mech. Solids*, 16(3), 2021.

[216] P. Helfrich and E. Jakobsson. Calculation of deformation energies and conformations in lipid membranes containing gramicidin channels. *Biophys. J.*, 57:1075–1084, 1990.

[217] M. Ikegami, Y. Nagaoka, S. Takagi, and T. Tanzawa. Quantum mechanics of a particle on a curved surface. *Prog. Theor. Phys.*, 88(2):229–249, 1992.

[218] P. A. M. Dirac. Note on exchange phenomena in the thomas atom. *Proc. Camb. Phil. Soc.*, 26:376, 1930.

[219] G.-T. Liu, H.-Y. Liu, X.-L. Feng, and S. A. T. Redfern. High-pressure phase transitions of nitinol niti to a semiconductor with an unusual topological structure. *Phys. Rev. B*, 97:140104, 2018.

[220] S. D. Prokoshkin, A. V. Korotitskiy, V. Brailovski, S. Turenne, I. Y. Khmelevskaya, and I. B. Trubitsyna. *Acta Mater.*, 52:4479–4492, 2004.

[221] Y. Kudoh, M. Tokonami, S. Miyazaki, and K. Otsuka. Crystal structure of the martensite in ti-49.2 at. %ni alloy analyzed by the single crystal x-ray diffraction method. *Acta Metall.*, 33:2049–2056, 1985.

[222] T. H. Nam, T. Saburi, Y. Nakata, and K. I. Shimizu. *Mater. Trans., JIM*, 31:1050–1056, 1990.

[223] T. Hara, T. Ohba, E. Okinushi, and K. Otsuka. Structural study of r-phase in ti-50.23 at. %ni and ti-47.75 at. %ni-1.50 at. %fe alloys. *Mater. Trans., JIM*, 38:11–17, 1997.

[224] Xu Yuanchong. *300 Tang Poems(唐诗三百首)*. China Translation & Publishing House, 2014.

[225] E. S. Fedorov. *Symmetry of Crystals*. American Crystallographic Association, 1971.

[226] M. I. Aroyo, editor. *International Tables for Crystallography*. John Wiley and Sons Limited, 2003.

[227] D. Feng and G. J. Jin. *Condensed Matter Physics (凝聚态物理学)*, volume 1. Higher Education Press, 2012.

[228] C. Mantegazza. *Lecture Notes on Mean Curvature Flow*, volume 290. Birkhauser, 2010.

[229] J. B. Hartle. *Gravity, An Introduction To Einstein's General Relativity*. Addison Wesley, 2003.

[230] R. Coghetto. Topology from neighbourhoods. *Formaliz. Math.*, 23(4):289–296, 2015.

[231] J. A. Bondy and U. S. R. Murty. *Graph theory with applications*. London: Macmillan, 1976.

[232] J. R. Munkres. *Topology*. Prentice Hall, 2000.

[233] P. A. M. Dirac. Generalized hamiltonian dynamics. *Can. J. Math.*, 2:129–148, 1950.